INAUGURAL - DISSERTATION

zur

Erlangung der Doktorwürde

der

Naturwissenschaftlich-Mathematischen Gesamtfakultät

der Ruprecht-Karls-Universität

Heidelberg

vorgelegt von

Iryna Charapitsa

aus Minsk, Weissrussland

Tag der mündlichen Prüfung: 22 – 06 – 2009

DISSERTATION

submitted to the
Combined Faculties for the Natural Sciences and for Mathematics
of the Ruperto-Carola University of Heidelberg, Germany
for the degree of

Doctor of Natural Sciences

Presented by

Iryna Charapitsa
Born in Minsk, Belarus
Date of oral examination: 22 - 06 - 2009

Biochemical and Functional Characterisation of the Novel Non Specific Lethal Complex

1st and 2nd examiners

Dr. Anne Ephrussi
Prof. Dr. Gabriele Peterson

Main Supervisor: Dr. Asifa Akhtar

To my family

ACKNOWLEDGMENTS

I am thankful to Asifa for introducing me into the world of chromatin and being a supportive inspiring supervisor, making the study presented in this thesis possible. It is a pleasure to know you and work together.

I am grateful to my Thesis Advisory Committee members: Dr. Anne Ephrussi, Dr. Lars Steinmetz and Prof. Gabriele Petersen for their useful input, guidance and helpful motivating feedback.

Special thanks to my lab and colleagues at EMBL for creating an enthusiastic working atmosphere that was driving our research.

I am deeply grateful to my friends and family, for their care and love.

Table of contents

SUMMARY

Genomic DNA is not freely accessible but it is compacted into chromatin by wrapping DNA around a histone octamer. Basic unit of chromatin is a nucleosome. Accessibility of nucleosomal DNA highly regulated and is orchestrated by many proteins that combinatorially alter the positional phasing of nucleosomes by chromatin remodeling enzymes, substitution of variant histones, post-translational modification of nucleosomes and the partitioning of chromatin into specific nuclear locations.

X chromosomal regulation by the process of dosage compensation provides an ideal model system to study the effect of chromatin and epigenetic factors on gene expression. In mammals, genes on the active X (Xa) chromosome are upregulated about twofold, with a corresponding inactivation of one of the two X chromosomes (Xi) ensuring equivalent sex chromosome expression in males and females. In *Drosophila*, dosage compensation is accomplished by the work of the MSL complex, which provokes a two-fold increase in the expression of genes on the male X chromosome. The MSL complex specifically binds to genes that require to be unregulated and, through the action of MOF, a histone acetyltransferase subunit within the complex, induces acetylation of H4K16, which is associated with an increase in the rate of transcription of genes.

In contradiction to the classic view that MOF was restricted to the male X chromosome, it has been found recently by our lab that MOF binds to multiple sites on the autosomes in both sexes. This suggests that MOF has a role in transcriptional regulation beyond dosage compensation. The work presented in this thesis shows the purification of a novel complex of evolutionary conserved proteins, which contains MOF. We termed the complex the NSL complex (Non-Specific Lethal), as mutation of proteins of the complex is lethal to both sexes. The NSL complex is composed of the evolutionary conserved proteins MOF, NSL1, NSL2, NSL3, MCRS2, MBDR2, WDS, Z4 and Chromator. These components of the NSL complex broadly decorate all chromosomes, and overlap with MOF on the X chromosome(s), as well as on all autosomes in males and females. Colocalization of NSL complex members with MOF occurs at the level of individual genes, with NSL associated with the promoters of

MOF-bound genes. Analysis of total RNA from fly lines expressing RNAi against NSLs specifically in salivary glands demonstrates that the binding of the NSL complex to promoters is functional, as there is a strong correlation between the absence of NSL and a decrease in transcription in males and females. Taken together, work performed in this thesis demonstrates that the NSL complex functions as a novel transcription regulator in Drosophila.

ZUSAMMENFASSUNG

Chromatin besteht aus DNA, die sich um Nukleosomen herumwindet. Die Nukleosomen haben voneinander einen mittleren Abstand von etwa 200 Basenpaaren (bp). Eine Regulation von Genexpression findet statt, indem der Zugang der RNA-Polymerase zur DNA durch Umbau der Chromatinstruktur ("chromatin remodeling") und kovalente Modifikation von Histonen moduliert wird. Diese Vorgänge werden durch eine Vielzahl an Proteinen gesteuert, die in kombinatorischer Weise die Position und den Abstand der Nukleosomen zueinander verändern, Histonvarianten substituieren, Nukleosomen posttranslationell modifizieren und das Chromatin in spezifischen Regionen des Zellkerns positionieren.

Der Prozess der Dosis-Kompensation ("dosage compensation") bietet ein Modellsystem, um epigentische Mechanismen der Expressionsregulation zu untersuchen. In Säugern werden Gene auf dem aktiven X-Chromosom (Xa) auf etwa die zweifache Dosis hochreguliert, korrespondierend zu einer Inaktivierung eines der beiden X-Chromosomen (Xi), wodurch eine äquivalente Expression der Gonosomen in Männchen und Weibchen sichergestellt wird. In *Drosophila* wird Dosis-Kompensation durch den MSL-Komplex erreicht, der eine Verdoppelung der Expression von Genen auf dem männlichen X-Chromosom bewirkt. Der MSL-Komplex bindet spezifisch an Gene, die hochreguliert werden müssen und induziert mittels seiner Komplexuntereinheit MOF, einer Histonacetyltransferase, die Acetylierung von H4K16, welche mit einer gesteigerten Transkriptionsrate assoziiert ist.

Im Gegensatz zu der klassischen Annahme, dass MOF auf das männliche X-Chromosom beschränkt sein solle, wurde vor kurzem gezeigt, dass MOF in beiden Geschlechtern an viele Regionen auf Autosomen bindet. Dies lässt vermuten, dass MOF eine über die Dosis-Kompensation hinausgehende Rolle bei der Transkriptionsregulation spielt. Die hier vorgestellte Arbeit beschreibt die Aufreinigung eines neuen MOF-enthaltenden Komplexes konservierter Proteine, Wir haben ihn NSL-Komplex genannt ("Non-Specific Letal"/unspezifisch letal), da Mutation von Komplexkomponenten in beiden Geschlechtern letal ist. Der NSL-Komplex setzt sich zusammen aus den konservierten Proteinen MOF, NSL1, NSL2,

NSL3, MCRS2, MBDR2, WDS, Z4 und Chromator. Die NSL-Komponenten interagieren über weite Strecken mit sämtlichen Chromosomen und überlappen dabei an vielen Stellen auf dem X-Chromosom (bzw. den X-Chromosomen) wie auch auf allen Autosomen mit MOF, sowohl in Männchen wie auch in Weibchen. Die Kolokalisation der NSL-Komponenten mit MOF geschieht auf der Ebene individueller Gene, wobei NSL mit dem Promotor MOF-gebundener Gene assoziiert. Die Analyse von Gesamt-RNA isoliert aus *Drosophila*-Linien, die eine RNAi gegen NSLs spezifisch in den Speicheldrüsen exprimieren, zeigt dass die Bindung des NSL-Komplexes an Promotoren funktionell relevant ist, da eine starke Korrelation zwischen der Abwesenheit von NSL und einer Abnahme der Transkription in Männchen und Weibchen beobachtet wurde. Diese Ergebnisse legen nahe, dass der NSL-Komplex eine allgemeine Rolle bei der Transkriptionsregulation spielt.

1. INTRODUCTION

1.1. Chromatin

Chromatin was named by Walter Flemming, who in 1882 observed the “nuclear scaffold” within nuclei, that could be visualized easily by staining with basophilic reagents. Chromatin literally means a “coloured, lightened material”, and today we know that it consists of the complex combination of DNA, RNA and protein that makes up chromosomes.

Chromatin is intra-nuclear within eukaryotic cells and present in the nucleoid in prokaryotic cells. It can be differentiated into heterochromatin (densely staining, condensed, inactive) and euchromatin (lightly staining, extended, active, generally found close to the nuclear periphery) (Frenster, 1965).

The major components of chromatin are DNA and histone proteins; however, many other chromosomal proteins have prominent roles, too. The essential functions of chromatin are to compact long molecules of DNA into a smaller volume to fit in the cell, to physically protect DNA, to strengthen and to compact the DNA to allow mitosis and meiosis to occur and, to provide a platform to regulate expression, DNA replication and DNA repair when needed. By physically achieving these complex roles, chromatin ensures a high fidelity of tra nsmission of genetic information from one generation to another.

1.2. Chromatin organization

The smallest brick of chromatin is a nucleosome (Kornberg, 1974). Nucleosomes are composed of two copies of canonical (core) histones, H2A, H2B, H3 and H4, which together comprise a histone octamers (Luger et al. 1997). Core histone proteins are small, highly basic molecules with two distinct domains: a globular compact core, and a flexible amino-terminal tail (Fig. 1, A). Their globular domains are composed of helix-turn-helix domains, which stack in the quaternary structure to promote oligomerisation. In contrast, the tail is unstructured, protrudes from the histone octamer and serves as a template on which are written a series of post-translational

covalent modifications, known as the histone code (Fig. 1, B). Nucleosome provides a scaffold structure, around which 147 bp of DNA make approximately 1.7 turns. Nucleosomes are separated from each other by linker DNA, which is generally 200 bp long. More recently, it has become apparent that nucleosome position is non-random and is intrinsically encoded within primary DNA sequence. Additionally, ATP-driven remodeling complexes act on nucleosome position to regulate access of proteins to cis-acting elements on DNA (Davey et al. 2002).

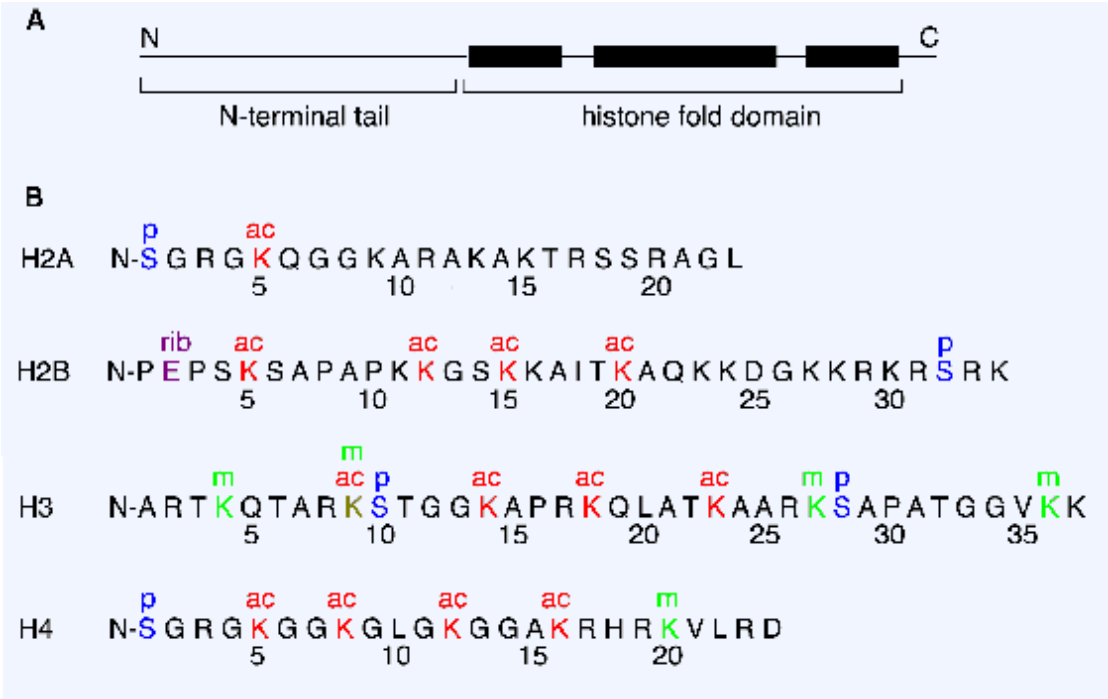


Figure 1. A. Structure of nucleosomal histones. **B.** Amino-terminal tails of core histones. The numbers indicate amino acid position. The post-translational modifications are indicated (red ac = acetylation sites; blue p = phosphorylation sites; green m = methylation sites; purple rib = ADP ribosylation) (Ridgway et al., 2002).

Chromatin at this initial level of a nucleosomal organization is 11 nm thick, has the appearance of “beads on a string” and is present as an accessible, active and largely unfolded interphase conformation (Fig. 2.). The interaction between nucleosomes and DNA is predominately altered by *cis*- and *trans*-effects of covalently modified histone tails.

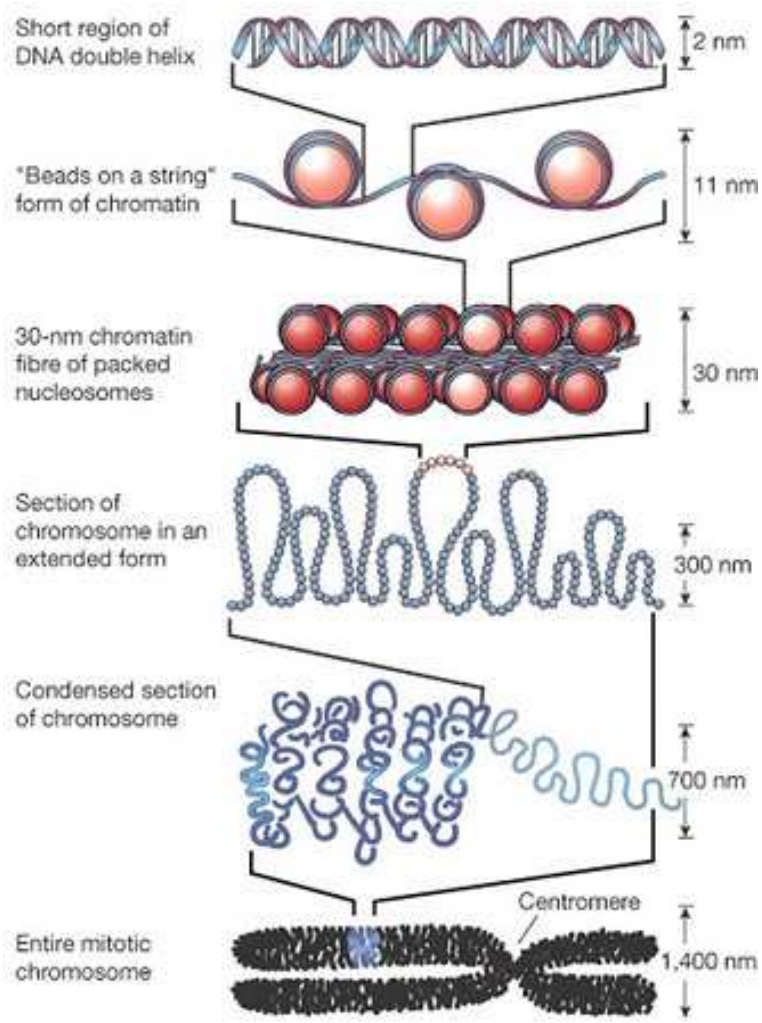


Figure 2. Each DNA molecule overgoes several levels of compaction, from a double helix into a mitotic chromosome that is 10.000 times shorter than its extended length (Felsenfeld and Groudine 2003)

cis-Effects are dictated through changes in physical properties of histone tails. Modulation of either the electrostatic charge or tail structure alters internucleosomal contacts. Acetylation of lysines and phosphorylation of serine and threonine residues are the most pronounced of such examples, when positive charges on highly basic histone tails are neutralized or negative charges are introduced to the histone tail; this generates a local expansion of chromatin fiber. In this relaxed state of chromatin, promoter elements are accessible to transcription factors and to the basal transcription machinery. In addition, an accumulation of negative charges can result on charge patches on chromatin, which can also alter nucleosomal packaging (Dou and Gorovsky 2000). *trans*-Effects result from the recruitment of modification-binding

elements to the chromatin. Many proteins, usually through discrete binding domains, have a specific affinity to particular histone modifications. Such recognition provides a platform for other proteins, frequently members of large enzymatic complexes, to associate with and further modify chromatin. For example, a bromodomain recognizes acetylated histone residues, and is often a part of histone acetyltransferase (HAT) enzymes, which in turn are associated with chromatin-remodeling complexes, that increase local DNA accessibility (Dhalluin et al. 1999; Jacobson et al. 2000). Methylated lysine residues are read by chromodomains, or similar domains, such as MBT or tudor (Bannister et al. 2001; Lachner et al. 2001). In this way, the addition and removal of specific post-translational histone modifications result in concerted regulatory effects on chromatin function. ATP-dependent remodeling complexes play a particular physical role in regulating gene expression. Their activity can result in octamer sliding, alteration of nucleosomal structure by DNA looping, or replacement specific canonical histones by their variants. These non-covalent modifications change positions of nucleosomes to expose or conceal DNA sequences, thereby regulating their physical exposure to molecular complexes, such as the basal transcription machinery (Narlikar et al. 2002).

The next level of chromatin compaction is to 30nm fibers. Nucleosome are stabilized by a linker histone H1, that associates at the entry and exit point of DNA on the core nucleosome, and/or by chromatin associated factors, such as heterochromatin protein H (HP1) or Polycomb (PC) (Fan et al. 2005). At this stage of organization, chromatin is looped and compressed about approximately 50 fold. Further compaction then results in 300-700 nm fibers that are fixed through anchoring to the nuclear periphery via chromatin associated factors, such as nuclear lamins. There is evidence that this high-order geography of chromatin is associated with distinct functional nuclear sub-domains, such as the clustering of active chromatin sites to RNA polymerase II transcription factories, or around replicating DNA and DNA polymerase, or to “silent” chromatin domains, such as pericentromeric foci. The dynamics and correlation between active or silent chromatin configuration with particular nuclear positioning remains poorly defined and subject to intense research activity.

DNA is at its most compact in metaphase chromosomes, both during meiosis or mitosis. This high condensation of DNA achieves equal distribution of sister

chromosomes between daughter cells. The 10.000 fold compaction is promoted by hyperphosphorylation of histones H1 and H3, the action of ATP-dependent complexes of condensin and cohesin, and super-coiling driven by the activity of topoisomerase II.

1.3. Chromatin dynamics and regulation of gene expression

Chromatin is subject to many different modifications and changes in its structure. It is achieved through tuned work of many regulatory proteins and results in chromatin remodeling, as well as appearing of covalent marks on it (Fig. 3). Following part of the introduction is focused on nucleosome remodeling and histone modifications.

1.3.1. Nucleosome remodeling

As discussed, DNA is wrapped around nucleosomes and is then further condensed at several levels. However, for processes such as gene expression, DNA repair and replication, large molecular assemblies have to gain access to DNA. This is achieved by dynamic alteration to the structure of chromatin that is fulfilled by several complexes, which either do not require energy, or are dependent on ATP hydrolysis. Energy independent processes generally act to covalently modify the amino terminal histone tails. Energy-dependence is a property of chromatin remodeling complexes, with the result of this work being the movement of histone octamers relative to DNA. Chromatin remodeling is consequently used to regulate access to specific DNA sequences. Both mechanisms are functionally interconnected, and both are required for opening chromatin structure to achieve activation of transcription, DNA repair and replication (Gangaraju and Bartholomew 2007)

There are five families of ATP-dependent remodeling complexes, classified upon the nature of their ATPase unit: the SWI/SNF, ISWI, Mi-2/NuRD, INO80, and SWR1families (Bao and Shen 2007).

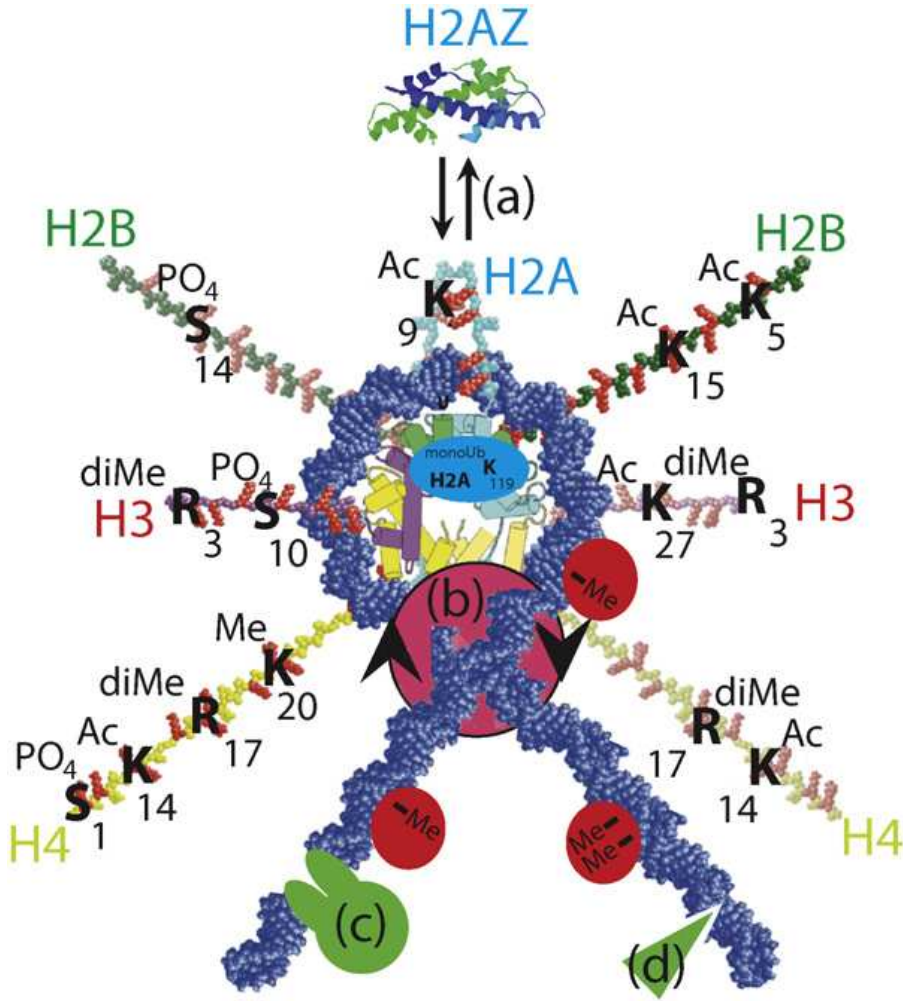


Figure 3. Chromatin provides a structural platform that is subject to extensive post-translational modifications: methylation, acetylation, phosphorylation and ubiquitination of specific histone residues; methylation of CpG dinucleotides; exchange of histones (a); changes in the relative position of the nucleosomemediated byATP-dependent remodeling complexes (b); induction of double-stranded DNA breaks by topoisomerase II (c) and the generation of single-stranded DNA breaks by topoisomerase I (d) (Reid et al. 2009).

1.3.1.1. SWI/SNF family

The 11-subunit SWI/SNF complex was the first chromatin remodeling factor to be discovered (Gangaraju and Bartholomew 2007). It was identified genetically as a regulator of mating type switching (SWI) and as required for growth on nutrient sources other than sucrose – thus SNF, sucrose nonfermenting (Peterson and Herskowitz 1992; Sudarsanam and Winston 2000). In *S. cerevisiae*, *Drosophila* and humans, there are two versions of the SWI/SNF complex: RSC and SWI/SNF. RSC is

more abundant and is essential for the cell growth, where as SWI/SNF is less present and is not critical for the growth (Du et al. 1998).

SWI/SNF is required for telomeric silencing and for silencing transcription of rRNA genes by RNA polymerase II (Dror and Winston 2004). It is also involved at an early step in homologous recombination, where RSC also acts at the stage of strand invasion (Chai et al. 2005; Huang et al. 2005). SWI/SNF also participates in sister chromatid cohesion and chromosome segregation (Huang et al. 2004; Huang and Laurent 2004) (Chang et al. 2005).

In Drosophila, the two forms of SWI/SNF are called BAP (Brahma associated proteins) and PBAP (Polybromo-associated BAP), and both share the same catalytic subunit (Brahma) (Mohrmann and Verrijzer 2005). In humans, the homologues are known as BAF (BRG1/hBRM-Associated Factors) and PBAF (Polybromo-associated BAF). However, there are many forms of human SWI/SNF that acquire tissue-specific subunits (Wang 2003) or additional sub-complexes, where the SWI/SNF-type remodelers become associated with other factors, such as BRCA1 (Bochar et al. 2000) (Decristofaro et al. 2001), components of the histone deacetylase Sin3 complex (Sif et al. 2001) and histone methylases (Pal et al. 2003; Pal et al. 2004). The action of SWI/SNF increases nucleosome mobility, through propagating DNA loops around the nucleosome surface by provoking a transitory disruption of DNA-nucleosome contacts (Aoyagi et al. 2002). SWI/SNF makes nucleosomal DNA accessible by creating loops on nucleosome surface. This does not alter nucleosome but brings DNA sequences into linker regions. This results in DNA becoming accessible to either transcription activators or to repressors (Gangaraju and Bartholomew 2007).

1.3.1.2. ISWI family

An *in vitro* assay for activities allowing transcriptional factor access to sites in nucleosomal arrays (Tsukiyama et al. 1995; Varga-Weisz et al. 1997) identified two chromatin remodeling enzymes, dNURF and dCHARC, the founding members of a growing ISWI family. Additionally chromatin remodelers belonging to this group have been identified in yeast, humans, mouse and *Xenopus*. Because of the similarity of their ATPase subunit to the SWI2 ATPase of the SNF2 subfamily, this class of

remodelers became known as Imitation SWItch (ISWI). The ATPase of ISWI type is characterized by the presence of a SANT (SWI3, ADA2, N-CoR and TFIIIB B) domain and by the absence of a bromodomain (Grune et al. 2003). There are indications that SANT domains might be responsible for the nonspecific binding of ISWI complexes to DNA and their resulting preferential biding to de-condensed nucleosomes with exposed linker DNA over nucleosomes associated with histone H1 (Langst et al. 1999). In general, ISWI complexes are smaller (300-800 kDa) and have 2-4 subunits, in comparison to larger complexes (up to 2mDa) from three other families which may contain up to 15 units.

In *Drosophila*, there are three complexes in the ISWI family: NURF, ACF and CHRAC. NURF (Nucleosome Remodeling Factor) was first identified by its requirement to induce accessibility of the *hsp70* heat shock promoter in the presence of the GAGA transcription factor (Tsukiyama et al. 1995). The complex is composed of four subunits: BPTF/Nurf301, ISWI, Nurf-38 and Nurf-55 (Tsukiyama and Wu 1995). NURF interacts with the histone H4 N-terminal tail and this interaction is essential for its ATPase and nucleosome mobilization activity (Georgel et al. 1997). The ATPase activity is stimulated by nucleosomes but not by DNA, in contrast to the SWI/SNF complex, where nucleosomes and DNA equally stimulate ATPase activity. NURF activates transcription *in vivo* and *in vitro* (Mizuguchi et al. 1997), and is achieved by mobilizing nucleosomes along the DNA. This requires the largest subunit of NURF – NURF301.

ACF (ATP-utilizing chromatin factor) is another NURF multisubunit complex. It processively deposits histone octamers along DNA to form long periodic arrays of nucleosomes (Ito et al. 1997; Fyodorov and Kadonaga 2002). ACF is a major chromatin assembly protein in *Drosophila*. Cells lacking it proceed more rapidly through S phase due to the lack of resistance from chromatin, as these complexes are involved in the formation of repressive chromatin.

CHRAC (chromatin accessibility complex) is a further ISWI containing complex, that additionally contains Acf1 and two histone fold containing proteins, CHARC-14 and CHRAC-16 (Varga-Weisz et al. 1997). Both CHARC subunits are involved in early

Drosophila development (Corona et al. 2000). CHRAC can also generate nucleosome arrays with regular spacing.

Human orthologues of ISWI (SNF2H and SNF2L) are incorporated into many complexes (Corona and Tamkun 2004) that work in transcription initiation, repression, elongation, termination and sister chromatid cohesion. The smooth functioning of these complexes are important for many developmental programmes to be fulfilled: defined examples include muscle (de la Serna et al. 2001; Simone et al. 2004; de la Serna et al. 2005; Ohkawa et al. 2006), heart (Lickert et al. 2004), blood (Vradii et al. 2006), skeletal (Young et al. 2005), neuron (Battaglioli et al. 2002; Olave et al. 2002; Seo et al. 2005; Matsumoto et al. 2006), adipocyte (Salma et al. 2004), liver (Inayoshi et al. 2006) and immune system/Tcell development (Gebuhr et al. 2003; Mudhasani and Fontes 2005).

In general, ISWI family chromatin remodeling complexes modulate nucleosomal DNA accessibility, by moving the entire nucleosome to either place the DNA site into the linker DNA region to increase accessibility or to move DNA onto the surface of the nucleosome, to decrease ease of access. ISWI is mostly involved in establishing a repressive chromatin environment (Gangaraju and Bartholomew 2007).

1.3.1.3. CHD family

CHD-1 (chromodomain-helicase DNA binding protein) was isolated from mouse and contains features of both the Swi2/Snf2 family of ATPase and of the Polycomb/HP1 chromodomain family of proteins (Delmas et al. 1993; Tsukiyama and Wu 1997). CHD1 has a minor groove DNA binding motif (Stokes and Perry 1995). In *Drosophila* it is found on polytene chromosomes, and is localized to interbands and puffs, which are regions of high transcriptional activity (Stokes et al. 1996). In contrast to the polycomb/HP1 complex, it is not localized to condensed chromatin. Both the chromo- and helicase domains of CHD1 are required for association with chromatin.

1.3.1.4. INO80 and SWR1 family

INO80 and SWR1 are both large complexes, containing 14 and 15 units, of which four are common between both complexes. INO80 and SWR1 are involved in transcription activation and DNA repair. The largest subunits of both complexes contain a conserved ATPase/helicase domain that is divided by a large spacer, whereas similar domains in other members of the SNF2 superfamily (Swi2/Snf2 and ISWI) are continuous (Shen et al. 2000). Yeast strains lacking INO80 mis-regulate transcription and are also hypersensitive to DNA-damaging agents, suggesting that INO80 regulates transcription and is involved in DNA repair (Morrison et al. 2004; van Attikum et al. 2004).

1.3.2. Histone modifications

Histones are highly conserved proteins; however, chromatin is not a uniform structure. Extensive post-translational modification, particularly to the unstructured N-terminal tail, generates local diversity in histone structure. It was initially shown that histones carry acetyl, methyl and phosphate groups; later, histones were the first discovered proteins substrates for ubiquitination (Robzyk et al. 2000; Wang et al. 2004). At first, a correlation between histone posttranslational modification (HPTMs) and their role in the regulation of gene expression was not obvious, with, many believing that there could not be a link between nucleosome modification and the transcriptional state of chromatin. However, a direct connection between gene regulation and nucleosome modification has been established. Indeed, post-translational modification of nucleosomes and nucleosome positioning can be maintained through cellular division, giving rise to an epigenetic role for the information content of nucleosomes in chromatin function.

All histone PTMs can be divided into two groups, dependent on the size of the covalent modification. Either small residues, such as acetyl, methyl and phosphate groups, can be added, often in combination, to nucleosomes or larger peptides, such as ubiquitin and SUMO may be added. The influence of PTM in gene regulation differs for each modification. They can directly affect the conformation of chromatin, through structural changes affecting nucleosomal or even higher-order organizations. HPTMs may also disrupt binding of chromatin or histones associated proteins. HPTMs also generate alternative binding surfaces, and by so doing, provide

interaction surfaces that can be interrogated by transcription factors. The information content generated by combinatorial covalent modifications at multiple sites on each of the four histones, known as the histone code, is interpreted by proteins that interact with each modification. The consequence of these interactions can be gene activation, gene repression or further sequential modification of the local histone information content. This results in changes in transcription. In the following part of the introduction, more details of HPTMs will be given.

1.3.2.1. Acetylation

The presence of acetylayed histones correlates with transcriptionally active regions, which usually have an open chromatin configuration that is accessible to large molecular probes, such as DNase and MNase. In the mid-90's, the first nuclear histone acetylation and deacetylation enzymes were identified, providing the first direct evidence that these enzymes play a role in transcription. The first nuclear histone acetyltransferase was isolated form *Tetrahymena macronucleus* (Brownell et al. 1996), and was found to be homologous to a previously isolated transcriptional coactivator in *S. cerevisiae*, Gcn5. In turn, Gcn5 was known to interact with transcriptional activators. Following this, the first histone deacetylase (HDAC) was isolated by biochemical purification (Taunton et al. 1996). This enzyme was homologous to the yeast transcriptional regulator Rpd3p, which was defined previously as having a cofactor role in gene expression. Collectively, these discoveries established a model, where DNA-bound activators recruit HATs to acetylate nucleosomal histones, with repressors targeting HDACs to de-acetylate histones. These alterations change the charge and structure of the nucleosome and regulate gene expression.

Many coactivators and corepressors possess HAT or HDAC activity, or associate with such enzymes (Sterner and Berger 2000; Roth et al. 2001); this enzymatic activity is crucial in gene activation. HATs and HDACs are often components of complexes, and the histone-modifying activity of them is just one function, and others include, for example, the recruitment of TBP (Grant et al. 1998). Some nuclear hormone receptors, for example, when bound to ligand, function as DNA-binding transcriptional activators, and when not bound, as transcriptional repressors. This is

predominately achieved by recruiting HATs to acetylate target chromatin regions when liganded, and by recruiting HDAC activity when not (Baek and Rosenfeld 2004).

There are three families of HAT proteins. They are distinguished by their targets. For the major HAT family, GNAT, (Gcn5 related acetyltransferase), histone H3 is the main target. CBP/p300 family is able to target both H3 and H4. Another large family, MYST, characterized by the presence of a chromodomain, targets histone H4. Depending on their specificity, enzymes of the MYST subfamily are divided into two groups: those that exclusively acetylate H4K16 *in vivo* (MOF and hMOF) (Smith et al. 2005), and those that acetylate all four terminal lysines on H4, such as Eas1, an essential SAS-related acetyltransferase 1 protein in yeast (Smith et al. 1998). Many HATs contain bromodomains which reinforce their association with acetylated histones (Hassan et al. 2002).

1.3.2.1.1. MYST family

A large part of the work described in this thesis is related to the histone acetyltransferase MOF from the MYST family. This family was first described in 1996 and originally named so by the name of its four founding members in yeast and mammals: MOZ, YBF2/SAS3, SAS2 and TIP60 (Borrow et al. 1996; Reifsnyder et al. 1996). The defining feature of HATs in this family is the presence of the highly conserved MYST domain composed of an acetyl-CoA binding domain and a zinc finger; some members of this family also have additional structural features such as chromodomains, plant homeodomain-linked (PHD) zinc fingers (Utley and Cote 2003; Yang 2004). They are parts of evolutionary conserved multisubunit complexes which play key roles in post-translational modifications of histones and by doing this influence on chromatin structure. Malfunctions of MYST HATs are linked to a number of human diseases including cancer (Avvakumov and Cote 2007). One of the members of the MYST family is a histone acetyltransferase MOF, which is an essential part of the dosage compensation machinery in *Drosophila*. Since the major part of the thesis is dedicated to this process, the next paragraph describes this protein in more details.

1.3.2.1.1.1. MOF

MOF, histone acetyltransferase from the MYST family, was first described in *Drosophila* screen for mutations that affect only male flies (Hilfiker et al. 1997). It is a key component of the dosage compensation complex (Akhtar and Becker 2000; Smith et al. 2000), and is a catalytic subunit of this complex with an enzymatic specificity to H4K16 residue (Turner et al. 1992); consequently, male X chromosome is hyperacetylated (Hilfiker et al. 1997). Although MOF is also capable of acetylating MSL3 (Buscaino et al. 2003) and MSL1 (Morales et al. 2004), its preferred substrate is histone 4 (Akhtar and Becker 2000; Smith et al. 2000), and the substrate specificity to H4 acetylation is increased upon integration into the dosage compensation complex (Morales et al. 2004). Solving the structure of MOF revealed that a putative chromo domain of it is organized by five beta strands, which are different from the alpha+beta fold of the canonical chromo domain, and was named a chromo-barrel domain, CBD. The domain shares a common fold with several other chromatin-associated modules, such as MB-like repeat, Tudor, and PWWP domains (Nielsen et al. 2002), which might mean that a chromo-barrel is an intermediate structure in the evolution of canonical chromo domains to these other modules, or vice versa (Nielsen et al. 2002). CBD and its adjacent lysine-rich region are engaged in RNA binding activity *in vivo* and *in vitro*, and a conserved tyrosine is important for this interaction (Akhtar and Becker 2000; Akhtar et al. 2000; Nielsen et al. 2005). Apart from the CBD, MOF also has a HAT domain, which activity is stimulated upon the interaction with MSL3 (Morales et al. 2004) and is required for specific acetylation of H4K16 (Smith et al. 2000; Akhtar and Becker 2001). Point mutation in HAT domain causes male specific lethality (Hilfiker et al. 1997). The third domain that MOF has is a zinc finger, a domain known to bind DNA.

1.3.2.2. Deacetylation

There are many HDACs that remove acetyl mark on histones, with at least 10 histone deacetylases been identified in *S. cerevisiae* and 19 in humans (Yang and Seto 2003; Keogh et al. 2005). They are categorized into three groups that are conserved from *S. cerevisiae* to mammals. Type I and Type II are hydrolases that contain Zn^{2+} at their catalytic site, whereas Type III, the Sir2-related enzymes, require the cofactor NAD^{+}

as to achieve hydrolytic cleavage of the acetyl group. Many HDACs are found in large multisubunit complexes that target enzymes to promoters to induce transcriptional repression. For example, Rpd3 is a part of a complex which includes the HDAC Sin3. Rpd3 is also a part of a complex that binds to open reading frames through a chromodomain mediated association with H3K36me. This results in histone deacetylation, which suppresses DNA pol II initiation (Carrozza et al. 2005; Joshi and Struhl 2005).

1.3.2.3. Phosphorylation

Phosphorylation is a very well characterized post-translational modification frequently involved in regulatory pathways and in signal transduction from the cell surface, through the cytoplasm and into the nucleus, resulting in changes to gene expression. Correspondingly, it was discovered that when cells were stimulated to proliferate, a set of immediate-early genes were induced to become transcriptionaly active. Increased gene expression correlated with histone H3 phosphorylation (Mahadevan et al. 1991).

Serine 10 of histone H3 is an important phosphorylation site regulating transcription from yeast to human, including *Drosophila*. A high density of H3S10 phosphorylation correlates with chromosome segregation during mitosis and meiosis. Although remaining rather unclear, or may be that phosphorylation induced charge neutralization of residues around S10H3 thereby allowing compaction to occur (Nowak and Corces 2004).

The mechanistic role of histone phosphorylation remains largely undefined. There are currently three views on what the influence of this modification is. In line with the proposal for chromosome condensation, work in *Tetrahymena* has demonstrated that the patch of negative charge induced by phosphorylation influences nearby residues, including linker histone H1, to decrease the affinity between the nucleosome and DNA. This increases the transcriptional potential of the local chromatin environment (Dou and Gorovsky 2002). Secondly, proteins bound to chromatin can be dislodged by phosphorylation, as shown by the lowered binding affinity of HP1 during mitosis subsequent to mitosis-specific H2S10 phosphorylation (Fischle et al. 2005; Hirota et

al. 2005). Thirdly, transcriptional regulation may be influenced by recruitment of 14-3-3 adaptor proteins to phosphorylated H2S10 (Macdonald et al. 2005).

1.3.2.4. Methylation

Histone methylation is diverse and complex and can be present on either lysine or arginine residues. The consequence of methylation upon transcriptional regulation can be either positive or negative, depending on the position of methylated residue within the histone. A further layer of complexity is that each residue can be multiple methylated, with lysines either mono (me1), di- (me2) or tri- (me3) methylated; whereas arginines can be mono- or dimethylated; dimethylation can either be symmetrical or asymettrical. As 24 lysine and arginines are available on H2A, H2B, H3 and H4 in total, one can imagine that there is a huge combinatorial potential of methylated nucleosomes. This diversity allows fine tuning of complex and dynamic processes, such as the regulation of gene expression regulation (Jenuwein and Allis 2001; Zhang and Reinberg 2001; Lee et al. 2005).

The number of all theoretically possible combinations of different methylated states in a given protein, provided all lysines and arginines can be methylated, is:

$$4^{K} \cdot 4^{R} - 1$$

where K is a number of lysines, and R – number of arginines in the protein.

It has long been known that histones are methylated; however the biological role of methylation was elucidated only recently, following the discovery of the first methyltransferase that uses histones as substrate (Rea et al. 2000). Today, many more histone methyltransferases, along with their sites of modification on histones, have been characterized (Martin and Zhang 2005). The common feature of all methyltransferases is the occurrence of a SET domain, with the only exception being Dot1. The SET domain contains a catalytic active site to which the methyl donor S-adenosyl-L-methionine cofactor binds.

Table 1. Histone lysines.

Modified residue	Effects	References
H3K4	di-methylation occurs on inactive and active euchromatic genes; tri-methylation is present exclusively on active genes; can result in recruitment of specific factors; H3K4 is prevented *in vitro*, when H3K4 is methylated and H3S10 is phosphorylated, it might function to block repressive H3K9 methylation on actively transcribed genes.	(Litt et al. 2001) (Noma et al. 2001) (Santos-Rosa et al. 2002) (Sims et al. 2005) (Li et al. 2006) (Zhang and Reinberg 2001)
H3K36	present on the coding region of active genes and is thought to be necessary for efficient elongation of transcripts; mono-ubiquitylation of Lys 123 of H2B represses H3K36 methylation; may repress transcriptional initiation when is present on promoters of inducible genes.	(Zhang and Reinberg 2001) (Carrozza et al. 2005) (Joshi and Struhl 2005) (Keogh et al. 2005) (Zhang and Reinberg 2001)
H3K79	present in euchromatic regions and in the transcribed region of active genes in yeast; restricts recruitment of the transcriptional repressors and by enhancing their concentration at repressive chromatin regions contributes to establishment and maintenance of silent heterochromatin; yeast H3K79 lysine methyltransferasse Dot1 is involved in the DNA repair checkpoint.	(Martin and Zhang 2005) (Huyen et al. 2004) (Okada et al. 2005)
H3K9	involved in silencing chromatin; creates a binding platform for HP1.	(Rea et al. 2000) (Bannister et al. 2001) (Lachner et al. 2001; Nakayama et al. 2001)
H3K27	repressive mark, present at pericentromeric heterochromatin, at the inactive X chromosome in mammals, and in euchromatic gene loci that contain, in case of *Drosophila*, polycomb response elements, PREs; is a binding site for a Polycomb.	(Cao and Zhang 2004) (Cao et al. 2002) (Czermin et al. 2002) (Muller et al. 2002) (Kuzmichev et al. 2002) (Fischle et al. 2003; Min et al. 2003)
H4K20	one of the less studied modifications; involved in the maintenance of heterochromatin and cell-cycle control; linked to DNA repair in budding yeast.	(Karachentsev et al. 2005) (Julien and Herr 2004) (Sanders et al. 2004)

The best characterized lysine sites of histone methylation are five on histone H3 (lysines 4, 9, 27, 36 and 79) and one on histone H4 (lysine 20) (Table 1.). Modification of three of these sites (H3K4, H3K36 and H3K79) induces transcriptional activation, with the remainder imposing repression (for review see Martina and Zhang, 2005). In addition to regulation of gene expression, methylation of H3K79 and H3K20 has been shown to be involved in the DNA repair.

1.3.2.5. Deimination

Arginine methylation is reversed through four activities: (i) the peptidylarginine desiminase PADI4 converts mono-methylated arginines to citrulline (Cuthbert et al., 2004; Wang et al., 2004); (ii) LSD1 (lysine-specific demsethylase 1) is an amine oxidase that demethylates H3K4me1 and H3K4me2 (Shi et al., 2004); (iii) di-oxygenases, characterised by a JmjC domain, demethylate mono- and di-methylated histones (Trewick et al., 2005; Tsukada et al., 2006); and (iv) the protein JMJD2C acts to demethylate H3K9me2 and me3 through a hydroxylation reaction requiring iron and alpha-ketoglutarate as cofactors (Cloos et al., 2006).

No enzyme directly capable of demethylating methylated arginine was found, which lead to a suggestion that probably there are other types of enzymatic reactions that may antagonize arginine methylation (Bannister et al. 2002). One of them is deimination. It is a process by which an arginine can be converted to citrulline via the removal of an imine group. Deimination of monomethylated arginine would result in the removal of methylamine group from arginine. In recent studies they have demonstrated the presence of citrulline in histones and identified the enzyme that converts arginines within histones into citrulline (Cuthbert et al. 2004; Wang et al. 2004; Wang et al. 2004). Also, the appearance of citrulline on histones H3 and H4 coincides with the disappearance of arginine methylation *in vivo*. Analysis of estrogen-regulated promoter, where arginine methylation correlates with the active state of transcription, has shown that citrulline appears with the promoter is shut off.

1.3.2.6. Ubiquitylation, deubiquitylation and sumoylation

In contract to previously described post translational modifications of acetyl or methyl groups, histones can also be modified with peptides, such as ubiquitin (Ub) and SUMO. Addition of these molecules increases the mass of histones by up to two-thirds. Ub and SUMO are 18% identical in sequence, and have a similar 3D structure and mechanism of ligation to substrates, although their surface charges are different, and so are the functional consequences upon ligation to substrates (Shiio and Eisenman 2003). Histones were the first example of proteins that are monoubiquitylated (polyubiquitylated substrates undergo proteosome mediated degredation), although the modified lysine residue (e.g. K119 of H2A) was discovered several years later (Robzyk et al. 2000; Wang et al. 2004).

Depending on the modified residue and histone, ubiquitylation can be, as with methylation, either repressive or activating. H2B monoubiquitylation activates transcription (Kim et al. 2005) and leads to H3K4 methylation (Henry et al. 2003). A monoubiquitylation mark on lysine 119 of H2A, in contrast, is repressive (Wang et al. 2004). Many ubiquitin interaction domains which bind to non-histone ubiquitilated substrates have been identified; however, to date, there are no known proteins that bind specifically to ubiquitulated histones.

Deubiquitylation of H2BK123Ub promoted both gene activation and heterochromatic silencing, achieved through the action of two different proteases: Ubp8 and Ubp10. Ubp 8 is a part of the SAGA histone acetylation complex (Sanders et al. 2002) and acts following ubiquitylation by Rad6 (Henry et al. 2003; Daniel et al. 2004). It may look strange on the first glance, that in order to achieve the result, first the mark has to be established, and then erased, but the sequence of H2B ubiquitylation and deubiquitylation is necessary to establish the right levels of methylation marks on lysines 4 and 36 of histone H3: H3K4 is dependent on H2Ub (and H3K36 does not require it) (Henry et al. 2003). Ubp10 works at silenced regions and is important to keep low levels of H3K4me and H3/H4, which are markers of transcription repression (Gardner et al. 2005). Small ubiquitin-related modifier (SUMO) is a member of a growing family of ubiquitin-like proteins involved in HPTM (for reviews see (Melchior 2000; Hay 2001; Johnson and Gupta 2001).

Histone sumoylation has a generally negative-acting role by preventing activation of HPTMs, which can be done via two mechanisms: sumoylated histone directly blocks lysine substrate sites (which are otherwise targets of acetylation), or they can also mediate transcriptional repression through recruitment of histone deacetylases and heterochromatin protein 1 (HP1) (Shiio and Eisenman 2003).

1.4. Dosage compensation as a model of chromatin regulation of gene expression

Gene expression is a very complex process, having several levels of regulation. It is orchestrated by many regulatory proteins, which lead to a diverse range of events, including chromatin remodeling, DNA and histone modifications, as well as positioning the chromatin into specific genomic loci.

One of the systems which allow study of gene expression regulation is a dosage compensation. Dosage compensation is a regulatory mechanism that ensures equal expression of X chromosome linked genes despite the difference in copy number of the sex chromosome between males and females. Different organisms have evolved diverse ways of compensating unequal distribution of sex chromosomes. In the following part of the introduction, evolution of dosage compensation and various ways of compensation the unequal amount of genes between sexes are discussed.

1.4.1. Evolution of dosage compensation

Many organisms have different sexes which are distinguished by having a different number of sex chromosomes. The defining influence for sex determination can be either genetic, and consequently heritable in the species, or external to it. In evolutionary terms, males and females had identical chromosomes with sex determined by environmental factors, such as temperature. Examples of this regulation are seen in some fish and reptile species today, where sex is determined by the incubation temperature of the egg, which directly affects sociosexual behavior and brain measures (Crews 2003). Environmentaly dependent sex determination has an advantage that better-adapted offspring arise under differing environmental conditions. However, the existence of the whole species can be compromised upon

sudden changes in environment conditions. Indeed, it has been postulated that inbalance in sex distribution through the lack of a temperature-independent checkpoint might have contributed to the demise of long-extinct reptiles, notably the dinosaurs, as a consequence of temperature deviations forcing production of predominantly one sex, eventually leading extinction (Miller et al. 2004).

In contrast to environmental sex determination, genetic sex determination defines gender at fertilization. Depending on which of the two sexes is homogametic, that is, possessing two identical sex chromosomes, two major classes of organisms can be distinguished. In most mammals, males are heterogametic (XY) with females being homogametic (XX). In birds and some reptiles, in particular snakes, females are heterogametic and have Z and W chromosomes, whereas males have two Z chromosomes.

With time, the additional accumulation of sex-specific mutations and genes and further divergence of the sex chromosomes has lead to a progressive degradation of the sex chromosome specific to the heterogametic sex (W in birds, Y in mammals); this may eventially result in the disappearence of the heterogametic chromosome (Graves 2006); Ohno 1967).

In consequence of the genetic inbalance arising from the loss of genes on the heterogametic sex chromosome, there is the potential that differential gene expression occurs between males and females. Inbalanced gene dosage is compensated by restricted expresssion of one of the homogametic sex chromosomes, with a number of dosage compensation mechanisms evolved in different organisms to deal with unequal gene dosage between sexes (Payer and Lee 2008), summarized in the table 2. and discussed in details further in this introduction.

Table 2. Dosage compensation in different organisms.

	Birds	***C. elegans***	**Mammals**	***Drosophila***
Sex determination	ZW/ZZ	XX/XO	XX/XY	X/A ratio
Dosage compensation	not known	Xx=XO X repression	Xx=XY Xi inactivation Xa activation	XY=xx X hypertranscription
Mechanism	gene by gene	Condensins Polycomb complex	Polycomb complex	MSL complex
Protein component	not known	DPY, SDCs, MIX-1, MES proteins	BED/Enx1 BRCA1	MSLs
RNA component	not known	not known	XIST	roX1, roX2

1.4.2. Dosage compensation in birds

The mechanism that birds use for dosage compensation is not entirely clear. In ZW females, the Z chromosome dosage compensation is incomplete and there are many Z-linked genes that have higher expression levels in males compared to females (Ellegren et al. 2007; Itoh et al. 2007; Mishra et al. 2007). So far, birds are the only example of an organism with a lack of global dosage compensation, providing a case of a viable system with large-scale imbalance in gene expression between sexes.

Dosage compensated genes in birds belong mostly to functional groups that differ from those of noncompensated. It suggests, that recruitment of dosage compensation machinery to genes depends on how important it is to maintain regulation of expression levels (Melamed and Arnold 2007). Such genes are mostly localized within the male hypermethylated region (MHR) on the Z chromosome. The corresponding regions on the female Z chromosome are coated by the noncoding MHR RNA and are enriched in H4K16ac mark (Teranishi et al. 2001; Bisoni et al. 2005). This resembles the situation in *Drosophila*, where the male X chromosome is bound by noncoding roX RNA-containing MSL complex in which induces H4K16 acetylation. This, in turn, leads to a transcriptional upregulation of the X chromosome

(discussed further in this thesis). Correspondingly, a similar model could also be true for birds, where MHR RNA recruits a histone acetyltransferase, which evokes a local hypertranscription of key genes on the single Z chromosome in female. In addition to MHR, DMRT1 also influences dosage compensation and sex determination (Teranishi et al. 2001). DMRT1 resides on the Z chromosome, resulting in a double dose in ZZ males. The double expression of DMRT1 in gonads of males may induce male-specific development (Raymond et al. 1999; Teranishi et al. 2001). This could have happened during evolution, as the MHR is adjacent to DMTR1 (Teranishi et al. 2001). MHR is likely to have been the first gene to become differentiated between the sexes and therefore required dosage compensation. Interestingly, DMRT1 itself is not hyperacetylated (Bisoni et al. 2005). This could be a mechanism to escape compensation to allow DMRT1 to function as a dosage-dependant determinant.

DMRT1 homologues are also involved in male sex determination in *Drosophila*, *C. elegans* and in vertebrates, including mice and humans (Raymond et al. 1999). Temperature-dependent sex determination in turtles and alligators is dependent on the expression levels of DMRT1, which is higher in males compared to females gonads, implying that DMRT1 links environmental and genetic sex-determination (Smith et al. 1999; Kettlewell et al. 2000).

1.4.3. Dosage compensation in *Caenorhabditis elegans*

In *Caenorhabditis elegans*, like in flies and mammals, heterogametic XO embryos become males, and homogametic XX turn into hermaphrodites. Hermaphrodite worms maintain both X chromosomes active, but represses transcript levels from each X chromosome by half, to match the expression from the single X in males (Meyer and Casson 1986). Several of the proteins (MIX-1 and DPY-27) that comprise the dosage compensation complex in *C. elegans* are similar to the conserved 13S condensin complex, which is required for both mitotic and meiotic chromosome resolution and condensation (Meyer 2005). DCC members also perform double duty as members of canonical meiotic and mitotic condensin complexes, and play role in regulating the number and distribution of crossovers during meiosis. These studies provide a nice example of how the protein function can be generalized through

evolution, in this case, from constraining and resolving topological features of DNA to the regulation of gene expression.

DCC in *C. elegans* is composed of at least eight proteins encoded by *sdc-1*, *sdc-2*, *sdc-3*, *dpy-21*, *dpy-26*, *dpy-27*, *dpy-28* and *mix-1* (Hodgkin and Brenner 1977; Hodgkin 1980; Meneely and Wood 1984; Meyer and Casson 1986; Meneely and Wood 1987; Villeneuve and Meyer 1987; Nusbaum and Meyer 1989; Plenefisch et al. 1989; Lieb et al. 1998). Each one of the proteins is localized to both X chromosomes of hermaphrodites (Chuang et al. 1994; Chuang et al. 1996; Davis and Meyer 1997; Dawes et al. 1999). Mutations in the corresponding genes lead to XX-specific lethality with few exceptions: *dpy-21* and *sdc-1* are not essential for XX survival, and *mix-1* is essential for both XX and XO animals, as MIX-1 is a shared subunit between the DCC and condensin complexes (Lieb et al. 1998; Hagstrom et al. 2002).

DCC is recruited to the X specifically by the action of SDC-2 and DNA sequence. During sex determination, the ratio between X and autosomes (X:A) is sensed by a set of X signal elements (XSEs) and autosomal signal elements (ASEs), which regulate the expression of *xol-1*. In XX hermaphrodites, *xol-1* is repressed, in XO males *xol-1* expression is induced, which in turn represses *sdc-2*. As a consequence, SDC-2 is present only in hermaphrodite embryos. It is the only protein of the DCC members that can recognize the X in the absence of all other DCC components (Nusbaum and Meyer 1989; Dawes et al. 1999; Chu et al. 2002).

The DCC recognizes the X through a limited number of recruitment sites called *rex* (recruitment elements on X), and spreads from them afterwards (Csankovszki et al. 2004; McDonel et al. 2006; Ercan et al. 2007). High-resolution ChIP mapping of DCC binding identified approximately 50 putative *rex* sites and a single 10 bp motif that encompasses information from previously identified motifs was found to be in common between them (Ercan et al. 2007). However, are not exclusive to the X, and not all the motifs on the X are bound by the DCC. The motif is more clustered on the X than other chromosomes, which suggests that the presence of multiple motifs provides a high-affinity binding site for the complex (McDonel et al. 2006; Ercan et al. 2007). DCC members are found preferentially bound near the transcription start sites, which implies that transcription initiation might be affected, although no

enzymatic function is known for the DCC members, except for the possible ATPase activity of DPY-27 and MIX-1 (Ercan et al. 2007). After the DCC has been targeted and spread, a repressed chromatin state is established over the chromosome, thereby maintaining the global, epigenetic regulation of X chromosomes that is maintained throughout the lifetime of hermaphrodites (Meyer 2005; Ercan and Lieb 2009).

1.4.4. Dosage compensation in mammals

Classically known and most studied process in mammalian dosage compensation is X inactivation in females: one of the two X chromosomes in females gets inactivated. This process can be divided into tree steps: (i) determination of the number of sex chromosomes followed by commitment to undergo dosage compensation; (ii) initiation of the inactivation process and spreading of silenced chromatin along the chromosome and (iii) maintaining the inactive state of the Xi (Avner and Heard 2001).

The initial steps in X inactivation are achieved by Xic, the X chromosome inactivation center. Prior to inactivation, determination of the number of X chromosomes relative to the cell ploidy has to be achieved, with only one X chromosome per cell eventually left functional. It is hypothesized that a blocking factor is produced in limiting amounts such that there is sufficient to bind only one Xic per diploid cell. The choice which of the two female X chromosomes will be inactivated depends on the tissue. In embryonic tissues, this choice is random, inactivated can be either paternal (Xp), or maternal (Xm) X chromosome. And in extraembryonic tissues it is always paternal X chromosome that gets silenced (Avner and Heard 2001). In consequence, in mammals, females have mosaic of X chromosome inactivation.

Suppression of X chromosome expression is initiated at X inactivation center, a locus known as Xic in mouse and XIC in human that encodes the X inactive specific transcript, Xist (Morey et al. 2004). Xist is a polyadenylated, spliced non-coding RNA transcribed only from the inactive X chromosome, which it binds to and coats. This induces recruitment of Polycomb group proteins, Eed and Enhancer of Zeste, that maintain the selected X chromosome in an inactive state (Czermin et al. 2002; Muller

et al. 2002; Silva et al. 2003). The action of Xist produces an inactive state that is initially labile; however, additional epigenetic marks, such as methylation, act to permanently silence Xi activity (Plath et al. 2003). Later, additional heterochromatic marks appear on the Xi soon after: hypoacetylation of histone H4, methylation of H3K27, methylation of the CpG islands, incorporation of the histone variant macroH2a. Late replication timing is also conferred on Xi (Heard 2004).

Interestingly, X inactivation has recently been found, in mice, during early pre-implantation development, to be much more dynamic than previously thought. Paternally inherited X chromosome is initially inactivated in all cells of early embryos, but then is selectively reactivated in the subset of cells that will form the embryo, with random X inactivation occurring afterwards (Heard 2004).

Dosage compensation in mammals, however, is not achieved by X inactivation only. Studies of single genes found that there are X-linked genes that escape the X inactivation. Thus, it was shown that X-linked in *Mus spretus* Clcn4 gene is expressed two-fold higher as compared with its autosomal ortholog in *Mus susculus* (Adler et al. 1997). Recently, due to the development of microarray technique, it became possible to measure the average levels of X-linked and autosomal expressions. These studies demonstrate that the gene upregulation on the active X chromosome is involved in dosage compensation along with inactivation, and the upregulation of the single active X is independent of the process of X inactivation (Nguyen and Disteche 2006).

In summary, mammals have developed two compensation mechanisms to counteract the imbalance of X-linked genes. Genes on the active X (Xa) chromosome are upregulated about twofold by a mechanism that still remains to be fully explained, resulting in a balance of X chromosome and autosomal expression in males. In females, upregulation of Xa expression is counteracted by inactivation of one of the two X chromosomes (Xi). This achieves a balance of X chromosome expression between sexes, although it remains unclear which of these mechanisms developed first, or if they co-evolved (Payer and Lee 2008).

1.4.5. Dosage compensation in *Drosophila*

Drosophila melanogaster, often used as a model organism to dissect epigenetic regulation, provides another example of dosage compensation. The extensively characterized genetics of the fruit fly makes it possible to address many questions on the mechanisms of dosage compensation, and has provided an excellent model to study epigenetic regulation.

Unlike in other organisms where dosage compensation is dependent on the restriction of X chromosome expression, the single X chromosome in heterogametic males is upregulated two fold, to achieve equal levels of transcripts in XY males and XX females. The X chromosome in males has no sequence difference to that of females, so males had to develop certain ways to a) make dosage compensation happen only in males and not in females, b) distinguish the X chromosome from autosomes, and c) maintain same level of gene expression compared to the other sex (Taipale and Akhtar 2005).

Genetic screens in *Drosophila* directed to find male-specific lethal mutations identified several genes, collectively named MSLs, standing for *m*ale *s*pecific *l*ethals (Bashaw and Baker 1997). They act together as a dosage compensation complex (Lucchesi 1998), which binds to multiple sites on the single male X chromosome and restores the level of gene transcripts to that of females. However, not all X chromosomal genes are dosage compensated (Ghosh et al. 1989; Baker et al. 1994; Kelley et al. 1995; Legube et al. 2006). One example of escape is the larval serum protein LSP1α. It is not compensated in males and consequently females have higher levels (Ghosh et al. 1989). Additionally, some genes are compensated in an MSL-independent way. These genes are most likely compensated by Sex-lethal, the master sex-determining gene in *Drosophila* (Baker et al. 1994; Kelley et al. 1995; Cline and Meyer 1996).

DCC includes 5 MSL proteins – MSL1, MSL2, MSL3, MLE (maleless), MOF (males-absent on the first) – and two non-coding RNAs, *roX1* and *roX2* (RNA on X) (Lucchesi 1998; Stuckenholz et al. 1999), and references therein; Fig. 4). The presence of this complex on the male X chromosome correlates with the occurrence

of acetylated lysine 16 on histone H4 (Turner et al. 1992; Bone et al. 1994). H4K16 is a distinguishing feature of the male X chromosome in *Drosophila*.

Figure 4. The dosage compensation complex contains five proteins (MSL1-3, MOF and MLE) and two non-coding RNAs (roX1 and roX2). The members of the complex are defined by their male specific lethality in respective mutant flies. An additional protein, Jil-1, is shown to interact with components of the DCC.

MSL1, MSL2 and MSL3 are required for the DCC complex to associate with the X chromosome (Lucchesi 1998). They appear to mediate binding of the whole complex to chromatin, although none of these proteins contain a distinguishable DNA-binding domain (Kelley et al. 1995; Copps et al. 1998; Gu et al. 1998).

MSL1 provides the assembly basis for the complex, as it interacts with all the other DCC members, except MLE (Scott et al. 2000). Interaction between MSL1 and MSL2 occurs through amino-terminal leucine zipper like motif of MSL1 and the RING finger domain of MSL2 (Copps et al. 1998; Scott et al. 2000; Li et al. 2005). Carboxyl-terminus of MSL1 binds MOF with its PEHE domain, and further to the C-terminus MSL1 binds to MSL3 (Scott et al. 2000; Morales et al. 2004).

MSL2 protein has a RING finger domain and a cysteins rich motif at the C-terminus (Zhou et al. 1995). This RING domain has two zinc finger clusters and mutations in it result in disruption of interaction between MSL2 and MSL1 (Copps et al. 1998). It is through MSL2, that the DCC complex associates with the X chromosome, resulting in a very stable interaction between MSL2 and chromatin (Straub et al. 2005).

MSL3 has a C-terminus MRG domain, which is responsible for mediating the interaction with MSL1 (Morales et al. 2005). MRG domains are thought to be interaction platforms in large complexes that are usually chromatin related (Bowman et al., 2006). MSL3 belongs to a family of proteins that coevolved with the chromodomain-bearing HATs (Pannuti and Lucchesi 2000) and may have a function in spreading of the MSL complex over the X chromosome (Taipale and Akhtar 2005). MSL3 interacts with roX2 in immunoprecipitation experiments and is tethered to the X chromosome via RNA (Buscaino et al. 2003). It is also discovered in the same study that association of MSL3 with the X chromosome is sensitive to RNase treatment. MSL3 is regulated by acetylation by MOF as a consequence of a direct interaction between these proteins. The interaction of MSL3 with *roX2* RNA, as well as localization to the X chromosome, are acetylation sensitive *in vitro* (Buscaino et al. 2003). This findings show that MOF is important not only for acetylation of the X chromosome, but also for regulation of other DCC members.

MOF is a histone acetyltransferase with specificity for lysine 16 acetylation on histone H4 (Akhtar and Becker 2001). It is an important enzymatic component of the dosage compensation complex and is discussed in the 3.2.1.1.1 part if this thesis.

Another protein with enzymatic activity in the DCC is **MLE**. It has an ATP-dependent RNA- and DNA-helicase activity in vitro, with the ATP-binding domain critical for its function *in vivo* (Lee et al. 1997). Most probably, however, that MLE functions in the DCC by altering the structure of the non-coding RNA, rather than by remodeling chromatin. Its localization to the X chromosome is RNAse sensitive, and the fact that MLE has a weak interaction with the rest of the DCC suggests that the binding may occur through *roX* RNAs (Richter et al. 1996; Copps et al. 1998).

An additional protein, **Jil-1**, interacts with components of the DCC. While it does associate with all chromosomes in males and females, it is enhanced at the MSL binding sites in males (Jin et al. 2000) with enrichment dependent on the MSL complex. JIL-1 maintains chromatin in an open configuration in transcriptionally active regions in the genome through phosphorylation of histone H3 (Wang et al. 2001). However, whether this protein plays a general or specific role in dosage compensation is not determined.

The DCC consists of not only proteins, but also of two noncoding RNAs, known as ***roX1*** and ***roX2***. These were discovered as male-specific RNAs in the adult brain (Amrein and Axel 1997; Meller et al. 1997). Although they differ in size (3.7 kb and 0.5-1.4 kb) and sequence, their functions of targeting MSL complex to the male X chromosome are redundant. Most double mutants die with the very few males that do survive exhibiting profound mislocalization of their MSL complex, and binding to a number of ectopic autosomal sites are detected (Meller and Rattner 2002; Deng et al. 2005). In contrast, males with a single *roX* knockout or mutation have no known phenotype (Meller et al. 1997; Meller and Rattner 2002). Overexpression of MSL1 and MSL2 can partially compensate for the lack of either *roX*, through promoting the assembly of the MSL complex on the X chromosome and increasing viability of *roX1 roX2* mutant males. This suggests that proteins of the DCC have sufficient capacity to effect dosage compensation, and that *roX* RNAs enhance either complex assembly or localization (Oh et al. 2003).

roX RNAs share a 30 nt similarity between themselves (Franke and Baker 1999). Deletion of it – along with another ~110 nt stretch of a similar sequence in the two RNAs (the DNAseI hypersensitive sites, DHS) – results in no obvious phenotype (Kageyama et al. 2001; Park et al. 2003; Stuckenholz et al. 2003). *roX1* has a putative stem loop at the 3' end, which might be responsible for *roX1* functions, as deletions of other parts of the RNA of approximately 400 bases does not affect the rescue of male lethality by truncated RNAs in a roX^- double-mutant background (Stuckenholz et al. 2003). Experiments with *roX2* give similar results: deletion of 17 nt from each of four regions with evolutionary conserved sequences and expressing the constructs in a roX^- double-mutant background also does not affect the rescue of male lethality (Park et al. 2008). However, it is shown that x-linked expression is reduced in roX^- double mutant male larvae (Deng and Meller 2006). *roX* RNA functions might be maintained by their secondary structures, which so far have been difficult to be predicted computationally.

1.4.6. Mechanism of dosage compensation

1.4.6.1. Choosing the sex

Prior to implementation of dosage compensation activity, the embryo must determine the ploidy of X chromosomes, in order to decide whether to implement dosage compensation or not. Failure to reach the correct decision results in lethality. In the fruit fly, sex determination, including estimating ploidy of sex chromosomes, occurs early in the development (Cline and Meyer 1996).

Phenotypic sex is determined by the number of X chromosomes per nucleus: XX embryos are females, and XY are males. Unlike in mammals, the Y chromosome does not play role in phenotypic sex determination. The ratio between the number of X chromosomes and autosomes controls both sex determination and dosage compensation. It does so by regulating the master regulator of sexual differentiation, sex lethal (Sxl). Sxl is encoded on the X chromosome and is up-regulated by transcription factors transcribed from the X chromosome such that embryos with two X chromosomes initiate transcription from *sxl* promoter, and embryos with a single copy of X do not. This initial difference in *sxl* expression is further propagated by a positive feedback autoregulatory loop; Sxl protein self regulates its own mRNA splicing from the promoter that is constitutively expressed and in addition, *Sxl* regulates splicing of the *transformer* (*tra*) gene in a sex-specific manner, thus initiating phenotypic differentiation into females. Together with *transformer2* (*tra2*), which is expressed in both sexes, *tra* directs the splicing of the *doublesex* (*dsx*) transcript, whose translated product represses male-specific genes, resulting in female sexual differentiation. In male embryos, the *dsx* transcript undergoes alternative splicing to result in a protein that represses female-specific genes, thereby inducing male sexual differentiation (Cline and Meyer 1996).

Splicing of *msl2* mRNA is under direct control of Sxl. Sxl-binding sites are located in both the 5’ and 3’ UTRs of the *msl2* mRNA (Bashaw and Baker 1997; Kelley et al. 1997). Normally, Sxl is present only in females, where it represses translation of the *msl2* mRNA. Correspondingly, when Sxl is absent in females, dosage compensation is induced, resulting in the death of females. Conversely, ectopic expression of SXL in

males abrogates dosage compensation, resulting in the death of genetically determined males. Ectopic expression of MSL2 in females induces DCC assembly on both X chromosomes, indicating that all MSL components are induced and/or are stabilized by the presence of MSL2 (Duncan et al. 2006).

1.4.6.2. Targeting the X

MSL1 and MSL2 are the core components of the dosage compensation complex. Both are required for nucleation of DCC function, and depletion of them results in disassembly of the complex (Kelley et al. 1995; Zhou et al. 1995; Lyman et al. 1997). Other members of the complex are responsible for the subsequent activity of X chromosome inactivation. After targeting and assembly, the MSL complex spreads throughout chromatin. As a consequence of DCC spreading, histones of genes subject to dosage compensation become hyperacetylated at H4K16ac, which is linked to an increase in gene expression.

Despite of many studies, the question of how the MSL complex distinguishes the X from the autosomes is still not answered. It is targeted to the X chromosome through *trans*-acting factors – *roX* RNAs, MSL1 and MSL2 – in conjunction with *cis*-acting DNA sequences. It was observed in mutants for MSL3, MLE or MOF flies, that MSL1 and MS3 are targeted to a limited number (35-100) of sites on the X chromosome, which have been named as chromatin entry sites (Lyman et al. 1997; Kelley et al. 1999). Not much is known about these sites; they function as nucleation sites, where the MSL complex enters and gets spread afterwards. Interestingly, *roX1* and *roX2* genomic loci themselves are chromatin entry sites for the assembly of the DCC (Kelley et al. 1999; Meller et al. 2000) and this function is independent of their transcription (Kageyama et al. 2001; Park et al. 2003). The DCC initially assembles at over a hundred entry sites on the X chromosome, and thereafter, propagates over the entire chromosome to spread to all target genes. The spreading, however, rarely occurs from autosomal *roX* transgenes (Kelley et al. 1999), which implies that chromosomal context also plays an important role. It was also shown, that large X chromosomal translocations are able to recruit DCC complex even if they do not have a previously mapped entry site (Fagegaltier and Baker 2004; Oh et al. 2004).

Early observations were done on the polytene chromosomes, which can not provide high resolution. Development of chromatin immunoprecipitation technique (ChIP) made possible studying DCC binding to chromatin on a new level. Analysis of data from high resolution MSL1 and MSL3 binding profiles has not revealed any universal targeting sequence, although short degenerative sequences have been identified (Dahlsveen et al. 2006; Gilfillan et al. 2006), and it still remains unanswered if chromatin entry sites are restricted to a DNA sequence, or dependent on chromatin structure. In general, summarized data from ChIPs of MSLs show that despite of different immunoprecipitated proteins, cell types and embryonic stages, MSL share several similar features in their profiles (Alekseyenko et al. 2006; Gilfillan et al. 2006; Legube et al. 2006; Kind et al. 2008). Interestingly, not all genes on the X chromosome are bound by the DCC, however, there are few autosomal sites of clear binding (Alekseyenko et al. 2006; Gilfillan et al. 2006). In addition, there is no correlation found between the expression level and MSL abundance (Alekseyenko et al. 2006; Legube et al. 2006). Although many genes that are compensated are actively transcribed, transcription by itself does not explain the MSL binding. There are many genes that are bound by elongating form of RNA pol II and elongation factors, but are devoid of MSLs (Gilfillan et al. 2006; Legube et al. 2006). Another observation from these studies is the MSLs distribution on the genes: the MSLs' profiles indicate an enrichment of proteins on the body and towards the end of the genes and affinity for targets correlates with their dosage compensation state (Gilfillan et al. 2006; Legube et al. 2006).

MOF profile clearly stands out from those of other MSLs. First, MOF has binding targets throughout the whole genome in both males and females; second, it also shows a strong peak at promoters of bound genes. 3' end enrichment of the genes on the X chromosome is MSL dependent, however promoter binding of both X chromosomal and autosomal genes is not (Kind et al. 2008). This finding suggests that MOF is involved in more general regulation at promoters in both sexes, and the MSL complex prolongs MOF's functioning towards the 3' end of the dosage compensated genes to up-regulate male X chromosome. Discovery of 3' bias in binding of MSLs led to looking at positioning of known epigenetic marks. Thus, it was found that more than 90% of MSL targets are enriched with H3K36me3 and it is a high correlation of MSLs and H3K36me3 position on the gene (Larschan et al. 2007).

It was recently shown in our lab, that ongoing transcription and polymerase passage through the gene is a prerequisite for target recognition, whereas the type of promoter and direction of transcription are not important (Kind and Akhtar 2007). Blocking of transcription by α-amanitin greatly reduces binding of the MSL complex to X chromosomal genes. However, transcription is not sufficient enough by itself as there are many genes on the X chromosome that are transcriptionally active, but not recognized by MSLs (Alekseyenko et al. 2006; Gilfillan et al. 2006; Legube et al. 2006). Targeting occurs independently of the neighboring context, as it is shown that MSLs can be recruited on X-liked genes, translocated to autosomes. Most probably, the recruitment signal lies in the coding region of the gene, which is exposed during transcription.

Recent work demonstrated that MOF is bound to promoters of numerous genes, that the distribution of MOF is not restricted to the maternal X chromosome and that MOF functions on all chromosomes in both sexes. However, how MOF is targeted to chromatin, the distribution of it over genes and the functional implications of association were not defined. The work presented in this thesis demonstrates that MOF is a constituent of a novel regulatory complex, termed NSL, that is targeted to the promoters of autosomal genes and that this NSL complex up-regulates expression of targeted genes.

2. AIMS AND OBJECTIVES

Gene expression is a highly regulated, complex process that requires to overcome multiple levels of restriction to productively engage DNA dependent RNA polymerases (Woolfe and Elgar 2008). Whether a gene is expressed is dependent not only on the information content within DNA, but also on multiple epigenetic and regulatory effects acting on chromatin, the functional template upon which gene expression is regulated. These events are orchestrated by numerous proteins, which achieve a diverse range of events on the regulatory elements of gene promoters including alteration of the positional phasing of nucleosomes, substitution of variant histones, post-translational modification of nucleosomes and positioning of chromatin into specific nuclear locations.

One attractive system to study the regulation of gene expression is dosage compensation in *Drosophila*. Dosage compensation ensures equal expression of X linked genes in males and females. In *Drosophila*, it is achieved by the MSL complex, which specifically recognizes the male X chromosome and doubles the expression of genes that require to be produced in equivalent levels in both sexes. Through the MOF acetyltransferase subunit, association of the MSL complex with the male X chromosome induces local acetylation of H4K16. This correlates with an increase in the expression of X linked genes, which the MSL complex fine tunes to two-fold higher than compared to females.

Recently, enzymatically active MOF-containing complexes have been purified from *Drosophila* embryos, Schneider SF4 cells and from human HeLa cells expressing tagged constituents of MSL complex: MOF and MSL3 (Mendjan et al. 2006). Mass spectrometric analysis revealed a diverse spectrum of evolutionary conserved proteins associated with MSLs in flies and humans. These include components of the nuclear pore complex (Mtor, Nup153, Nup154, Nup160 and Nup98), the nuclear exosome (Dis3 and Rrp6) and chromatin-interacting proteins that are enriched at polytene chromosome interbands (Z4, Chromator and MBDR2) (Mendjan et al. 2006). In addition, four novel and uncharacterized proteins CG1135, CG4699, CG18041 and CG10081 were found in purification of TAP tagged MOF. CG1135 was named

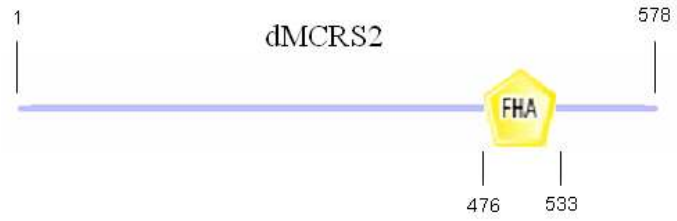
dMCRS2 by the name of its human ortholog hMCRS2, and CG4699, CG18041 and CG10081 were named NSL1, NSL2 and NSL3 correspondingly, standing for *n*on-*s*pecific *l*ethals, as P-element insertions in corresponding genes in *Drosophila* result in a general lethal phonotype.

The aim of my PhD project was to functionally and biochemicaly characterize CG1135/MCRS2, a novel protein identified in MOF purification. In the course of this project it was found out that this protein is a member of a distinct complex which is composed of evolutionary conserved proteins. To characterize it and gain further insight into a functional role of this novel complex and MCRS2 in particular, multiple approaches were applied, including biochemical, cytological and genetic methods.

3. RESULTS

3.1. Domain structure of MCRS2

Drosophila MCRS2 protein is a novel poorly characterized protein composed of 578 aa (63 kDa) (Fig.4). It has a fork-head associated (FHA) domain, which belongs to a class of signaling modules able to recognize phosphorylated epitopes on proteins (Hofmann and Bucher 1995; Yaffe and Smerdon 2001). This domain has been found in many regulatory proteins in eubacterial and eukaryotic genomes. They include kinases, phosphatases, kinesins, transcription factors, RNA-binding proteins and metabolic enzymes, which bind to phosphopeptides and take part in many different cellular processes, such as DNA repair, signal transduction and vesicular transport (Durocher et al. 2000). Mammalian homologues of *Drosophila* MCRS2, hMCRS1 and its splice variant, have been reported to be involved in transforming, nucleolar sequestration, ribosomal gene regulation, signaling between telomere maintenance and cell-cycle regulation (Song et al. 2004; Wu et al. 2009). Recently, cDNA encoding residues 126-475 of hMCRS2 from the HeLa cell cDNA library has been found in yeast two-hybrid screening assay to identify Nrf1-interacting proteins (Nrf1 [p45 nuclear factor-erythroid 2 (p45 NF-E2)-related factor 1] is a transcriptional activator), showing that hMCRS2 has a repression effect on Nrf1-mediated transcriptional activation (Wu et al. 2009).




```
human_MCRS1_isoform1 ---------MDK-------------DSQGLLDSSLMASGTASRSEDEESLAGQKRAS--- 35
human_MCRS1_isoform2 ---------MTRGTGGTAQRGRSGPDSQGLLDSSLMASGTASRSEDEESLAGQKRAS--- 48
mouse_MCRS1          ---------MDK-------------DSQGLLDSSLMASGTASRSEDEESLAGQKRAS--- 35
drosophila_MCRS2     MEASRITAIASSAVSVTAPNPPTVSTIPTAAASTLIQVGVSPATTTMPTPAATTTTTTIG 60

human_MCRS1_isoform1 ---SQALGTIP-----------------KRRSSSRFIKRKKFDDELVESSLAKSSTRAKG 75
human_MCRS1_isoform2 ---SQALGTIP-----------------KRRSSSRFIKRKKFDDELVESSLAKSSTRAKG 88
mouse_MCRS1          ---SQALGTIP-----------------KRRSSSRFIKRKKFDDELVESSLAKSSTRVKG 75
drosophila_MCRS2     STASSAVGISTPIRNPISNLQIEQQNDQKRRSSSRTIKRKRFDDEIVEYNIAVPTNRSGT 120

human_MCRS1_isoform1 ASGVEPGRCSGSE------------------PSSSEKKKVS-----------------KA 100
human_MCRS1_isoform2 ASGVEPGRCSGSE------------------PSSSEKKKVS-----------------KA 113
mouse_MCRS1          AGGVESGRCSGSE------------------PSSSEKKKVS-----------------KA 100
drosophila_MCRS2     DANRSSRPRTTSQNYPALVGVPHTTLAPLNIPTSTPQTPLSVDSLLPGTPSTVASLSLAT 180

human_MCRS1_isoform1 PSTPVPPS-PAPAPGLTKRVKKSKQP-------------------------------LQV 128
human_MCRS1_isoform2 PSTPVPPS-PAPAPGLTKRVKKSKQP-------------------------------LQV 141
mouse_MCRS1          PSTPVPPS-PAPTPGLTKRVKKSKQP-------------------------------LQV 128
drosophila_MCRS2     PTTPAPLATPLPVAPIVTAVAHPKPPAMERSTTSERRSRPVRPASKKAQRRNGRPMGQMA 240

human_MCRS1_isoform1 TKDLGRWKPADDLLLINAVLQTNDLTSVHLGVKFSCRFTLREVQERWYALLYDPVISKLA 188
human_MCRS1_isoform2 TKDLGRWKPADDLLLINAVLQTNDLTSVHLGVKFSCRFTLREVQERWYALLYDPVISKLA 201
mouse_MCRS1          TKDLGRWKPADDLLLINAVLQTNDLTSVHLGVKFSCRFTLREVQERWYALLYDPVISKLA 188
drosophila_MCRS2     TKDLGRWKPIDDLALIIGIQQTNDLRIIHRGVKFSCKFTLQELQQRWYALLYEPAVSRIA 300

human_MCRS1_isoform1 CQAMRQLHPEAIAAIQSKALFSKAEEQLLSKVGSTSQPTLETFQDLLHRHPDAFYLARTA 248
human_MCRS1_isoform2 CQAMRQLHPEAIAAIQSKALFSKAEEQLLSKVGSTSQPTLETFQDLLHRHPDAFYLARTA 261
mouse_MCRS1          CQAMRQLHPEAIAAIQSKALFSKAEEQLLSKVGSSSQPTLETFQDLLHTHPDAFYLARTA 248
drosophila_MCRS2     VSAIRNLHPELVESVQRKALYSVQEEDLLGTIKSSEQPKLEQFQELLDKNASVFYCARTA 360

human_MCRS1_isoform1 KALQAHWQLMKQYYLLEDQTVQPLPKGDQ-VLNFSDAEDLIDDSKLKDMRDEVLEHELMV 307
human_MCRS1_isoform2 KALQAHWQLMKQYYLLEDQTVQPLPKGDQ-VLNFSDAEDLIDDSKLKDMRDEVLEHELMV 320
mouse_MCRS1          KALQAHWQLMKQYYLLEDQTVQPLPKGDQ-VLNFSDAEDLIDDSKLKDMRDEVLEHELTV 307
drosophila_MCRS2     KSLQNHWLLLKQYTLLPDQSVKPIYGTDQQPLSFSDAEDQIFEHDLNEPRDEALEMERAL 420

human_MCRS1_isoform1 ADRRQKREIRQLEQELHKWQVLVDSITG-MSSPDFDNQTLAVLRGRMVRYLMRSREITLG 366
human_MCRS1_isoform2 ADRRQKREIRQLEQELHKWQVLVDSITG-MSSPDFDNQTLAVLRGRMVRYLMRSREITLG 379
mouse_MCRS1          ADRRQKREIRQLEQELHKWQVLVDSITG-MGSPDFDNQTLAVLRGRMVRYLMRSREITLG 366
drosophila_MCRS2     ADRRNKRNIRLLENELSRWAVLVDSVLSPTAASEFDNQTLACLCGRHVRYLMRSKEITFG 480

human_MCRS1_isoform1 RATKDNQIDVDLSLEGPAWKISRKQGVIKLKNNGDFFIANEGRRPIYIDGRPVLCGSKWR 426
human_MCRS1_isoform2 RATKDNQIDVDLSLEGPAWKISRKQGVIKLKNNGDFFIANEGRRPIYIDGRPVLCGSKWR 439
mouse_MCRS1          RATKDNQIDVDLSLEGPAWKISRKQGVIKLKNNGDFFIANEGRRPIYIDGRPVLCGSKWR 426
drosophila_MCRS2     RDAKDCVVDVDLGLEGPAAKISRRQGTIKLRSNGDFFIANEGKRAIFIDGTPLLSANKAR 540

human_MCRS1_isoform1 LSNNSVVEIASLRFVFLINQDLIALIRAEAAKITPQ-- 462
human_MCRS1_isoform2 LSNNSVVEIASLRFVFLINQDLIALIRAEAAKITPQ-- 475
mouse_MCRS1          LSNNSVVEIASLRFVFLINQDLIALIRAEAAKITPQ-- 462
drosophila_MCRS2     LGHNCTVEISGLRFTFLVNYELINAIRQESAKTSNPLN 578
```

Figure 4. Scheme of *Drosophila* MCRS2 and alignment of its homologues in human and mouse. Underlined residues are conserved amino acids in the FHA domain (Durocher and Jackson 2002).

3.2. Raising antibodies against MCRS2

Overall, there was very limited information of published literature about this protein. As a first step towards characterization of the MCRS2 protein, polyclonal antibodies against 1-319 aa of MCRS2 protein were raised in rats and rabbits, in a project that was initiated by Anan Ragab and Herbert Holz in the lab. N-terminally GST tagged first 319 aa of MCRS2 protein were expressed in BL21 Rosetta (EMD Biosciences) using the pET41a vector system (EMD Biosciences). The protein was recovered from inclusion bodies and affinity purified on glutathione agarose (Fig. 5; for detailed information refer to Materials and Methods). This material was then mixed with Titre-Max (Sigma) adjuvant and used to immunize 3 rats and 2 rabbits at three week intervals for a series of six injections.

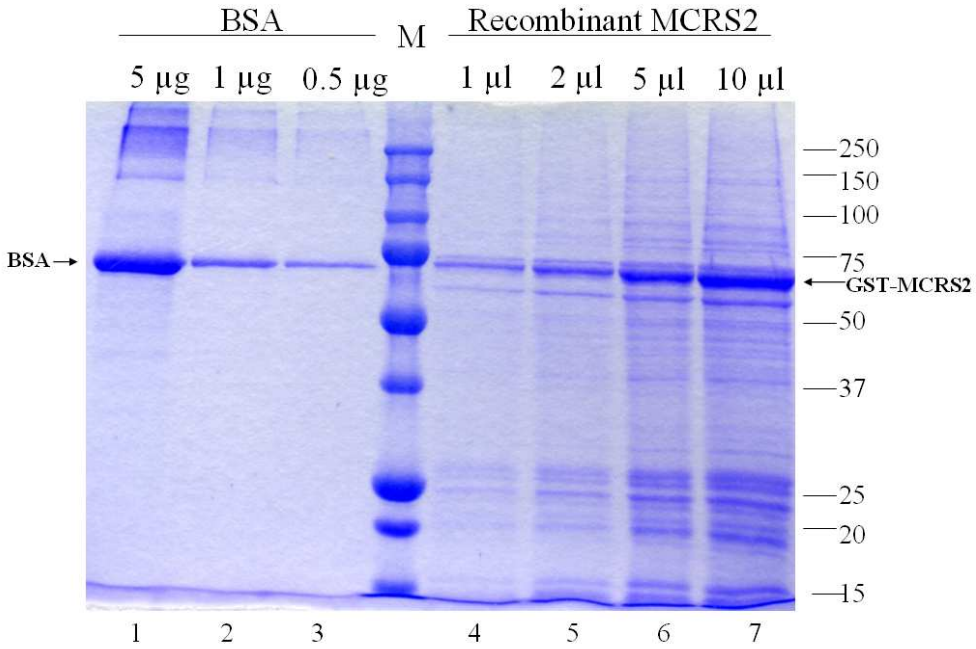


Figure 5. Purified, recombinant GST-tagged MCRS2, used for antibody production, loaded as 1, 2, 5, and 10 µl from the purification (lanes 4-7), compared to 0.5, 1 and 5 µg of BSA (lanes 1-3). Calculated molecular weight of GST-MCRS2 is 61 kDa. Additional bands are representing degraded protein as well as unspecifically copurified proteins.

3.2.1. Rat αMCRS2 antibodies

Pre- and post- immunization sera of rats were evaluated for specificity against *Drosophila* embryo nuclear extract and recombinant MCRS2. Western blot analysis (Fig. 6B) shows that immune sera of all three rats specifically recognize both recombinant GST-MCRS2 and the endogenous MCRS2 from *Drosophila* embryo nuclear extract. Incubation of membranes with preimmune sera results in a clean background (Fig. 6A); preimmune serum of the rat 2 shows in unspecific band above 75 kDa in embryo nuclear extract, which is of a different size of the band recognized by immune sera.

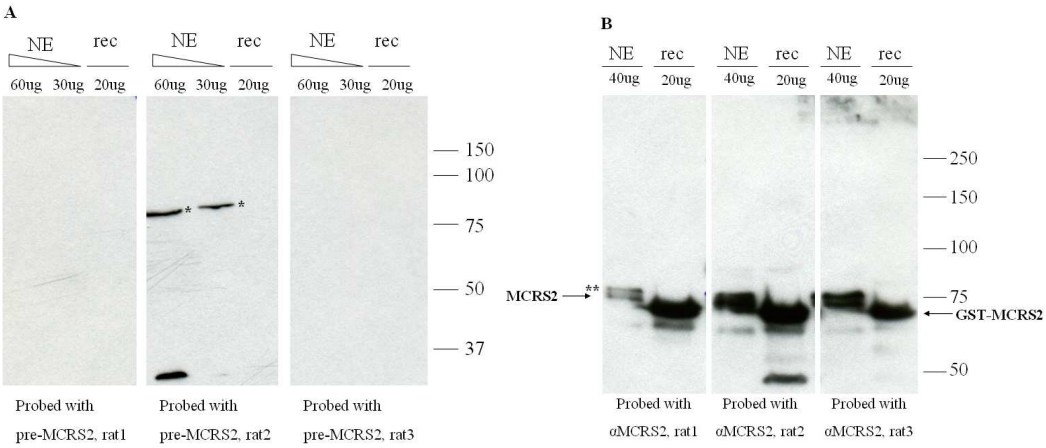


Figure 6. A. Western blots of *Drosophila* embryo nuclear extract (NE) and recombinant MCRS2 protein, probed with pre-immune from three rats. (*) indicates a nonspecific recognition of a band above 75 kDa in *Drosophila* embryos nuclear extract by the pre-immune serum of rat 2. **B**. Western blots of *Drosophila* embryo nuclear extract (NE) and recombinant MCRS2 protein, probed with immune sera from three rats. (**) shows a recognized band above MCRS2 size from a cytoplasmic contamination of a nuclear extract. Western blots were performed by Anan Ragab (A) and Herbert Holz (B).

3.2.2. Rabbit αMCRS2 antibodies

Similarly, pre- and post- immunization sera of rabbits were evaluated for specificity against *Drosophila* embryo nuclear extract and recombinant MCRS2. The rabbit α-MCRS2 anti-sera (Fig. 7B) recognize specifically both recombinant GST-MCRS2 and the endogenous MCRS2 from *Drosophila* embryo nuclear extract. Unlike the anti-sera, preimmune sera does not recognize specific proteins (Fig. 7A).

In conclusion, antibodies against MCRS2 were generated and characterized in three rats and two rabbits. All of them are able to specifically recognize MCRS2 in *Drosophila* embryo nuclear extract and can be used for multiple biochemical applications.

At the same time, antibodies against a number of proteins, identified in MOF TAP purification, were raised in the lab, allowing further immunochemical characterization of MCRS2 and its potential interactors.

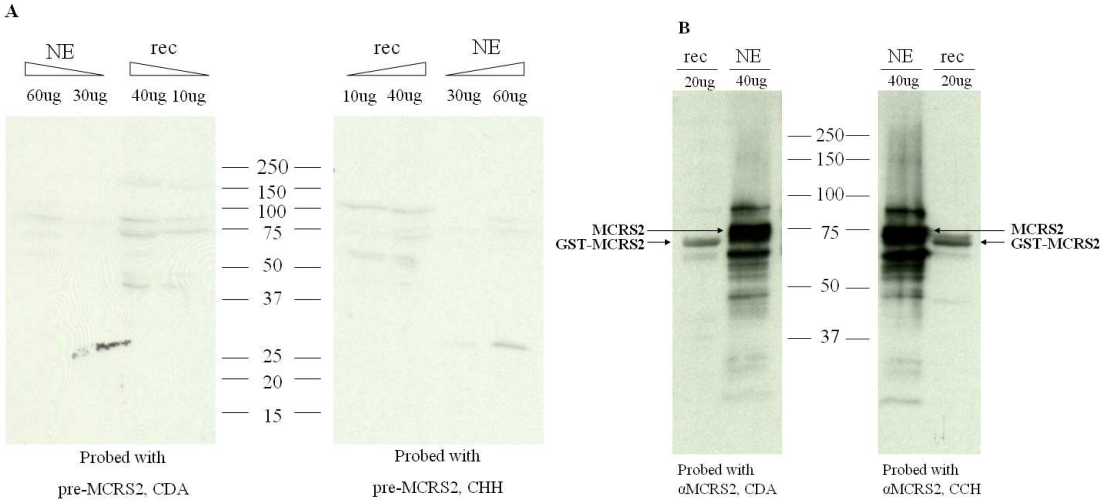


Figure 7. Western blot of *Drosophila* embryo nuclear extract (NE) and recombinant MCRS2 protein, probed with (A) pre-immune sera from two rabbits – CDA and CCH; and (B) their corresponding post-immune sera.

3.3. Immunoprecipitation analysis reveals biochemical partners of MCRS2

MCRS2 was found as one of the proteins purified with TAP-tagged MOF. In order to address if the interaction of MCRS2 and MOF can be reproduced, an immunoprecipitation experiment using newly produced antibodies against MCRS2 was performed. In addition to MOF antibodies, western blot membranes were probed with antibodies against proteins, which were found in the MOF TAP purification. Surprisingly, MSLs were not coimmonuprecipitated by MCRS2 (Fig. 8, lane 4). In order to compare MCRS2 interacting proteins to the MSL complex, an additional immunoprecipitation experiment using MSL1 antibodies was done and the membrane

was probed with the same set of antibodies (Fig. 8, lane 6). Again, none of the MSLs were found in the immunoprecipitate of MCRS2. Among other proteins, that were tested for the presence in MCRS2 immunoprecipitate, were Mtor and NXF1. Mtor has also been identified in the MOF TAP purification, however, immunoprecipitation with MCRS2 showed that it is not residing in one complex. NXF1 was used as a negative control protein, which was not present in MOF TAP purification.

The experiment shows that MCRS2 coimmunoprecipitates MOF, NSL2, NSL3, WDS and MBDR2 in *Drosophila* embryo nuclear extracts, whereas MSL1 coimmunoprecipitates only MSL3 and MOF (Fig. 8, lane 4 versus 6). These results indicated that MOF exists in two separate complexes: one complex is the MSL complex, and another one groups MCRS2 and NSLs. The experiment was performed together with Herbert Holz.

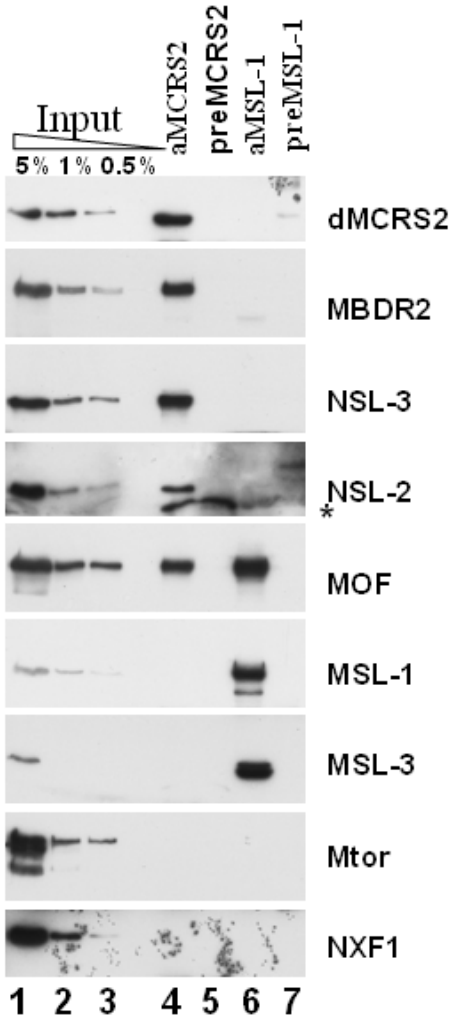


Figure 8. Western blot analysis of MOF interacting proteins. Immunoprecipitation was performed with MCRS2 and MSL1 antibodies using *Drosophila* embryos nuclear extract (lanes 4 and 6), corresponding pre-immune sera was used as a negative control (lanes 5 and 7), eluates were resolved by SDS-PAGE and probed with MCRS2, MBDR2, NSL3, NSL2, MOF, MSL1, MSL3, Mtor and NXF1 antibodies.

3.4. Purification of a novel MCRS2 containing complex

The results of coimmunoprecipitation experiment revealed a subgroup of proteins, distinct from the MSL complex, which were interacting with MOF. In order to identify all the interacting partners of MCRS2, tandem affinity purification of N-terminally TAP-tagged MCRS2 was performed. To facilitate these analyses, a stable *Drosophila* SL2 cell line was established that expresses TAP-tagged MCRS2. The scheme of MCRS2 tandem affinity purification (TAP) is shown on figure 9.

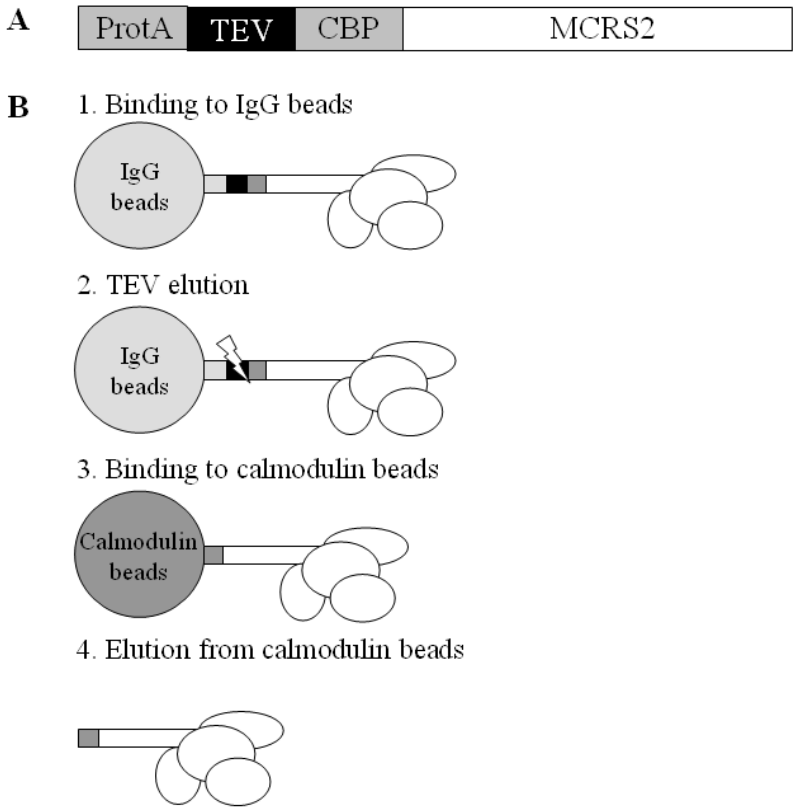


Figure 9. **A**. Schematic representation of TAP-MCRS2 protein. MCRS2 is tagged with TAP tag, which consists of two units of protein A, separated by a TEV protease cleavage site from a calmodulin binding site.
B. Sequential affinity purification of TAP MCRS2. The purification is performed under conditions that do not disrupt complexes containing the tagged protein. It is achieved by (1) retaining the TAP-tagged protein on IgG beads, washing non-binding proteins from the beads and then (2) releasing the complex from the beads by digestion with TEV protease. The resulting eluate is subjected to (3) a second round of purification on beads covalently coated with calmodulin. After (4) elution the resulting material is highly enriched in the TAP tagged protein and in the components of complexes containing the TAP tagged protein (adapted from (Puig et al. 2001)).

3.4.1. TAP-tagging of the MCRS2 protein and generation of stable *Drosophila* SL2 cell line

The full-length open reading frame of MCRS2 protein was cloned into the multiple-cloning site of the pBSactshort-N-TAP vector. Expression of the tagged protein is driven from a shortened Actin5C-promoter, which results in low-level ubiquitous expression of TAP tagged MCRS2. A stable *Drosophila* TAP-MCRS2 producing SL2 cell line was established by co-transfecting the pBSactshort-N-TAP-MCRS2 expression plasmid and the pUC-NEO resistance vector. Geneticin was used for the selection, with a range of concentrations between 0.8 and 1.2 mg/ml. Selection was monitored by the complete death of mock transfected cells and colony formation of stably expressing cells in cells transfected with the selection vector. Heterogeneous populations of transfected cells were then used to derive cell lines which were evaluated for MCRS2 expression (Fig. 10). For this, whole cell extracts from in the same amounts were loaded on a acrylamide gel and checked in western blot with PAP (anti-TAP) antibodies for the expression of the TAP tagged MCRS2. The result showed no difference in cell lines established either in 1.2 or 0.8 mg/ml of geneticin, the lower concentration of 0.8 mg/ml was used to establish and maintain stable cell line. Probing with antibodies against MCRS2 antibodies revealed a low level of expression of TAP-MCRS2 as compared to endogenous protein, as it can be hardly detected in the cell extract. This, however, precludes forcing the formation of aberrant complexes through over-expressing MCRS2 and is of a benefit.

The expression of TAP-MCRS2 was checked in the whole cell extract. However, fractionation of cells into nuclear and cytoplasmic extracts (for the procedure, refer to Materials and Methods part of this thesis) indicated that MCRS2 is predominately present in the nucleus (Fig. 11, line 1 versus 2-3). It was therefore prudent to perform tandem affinity purification from nuclear extract preparations, thereby reducing contamination with cytoplasmic proteins.

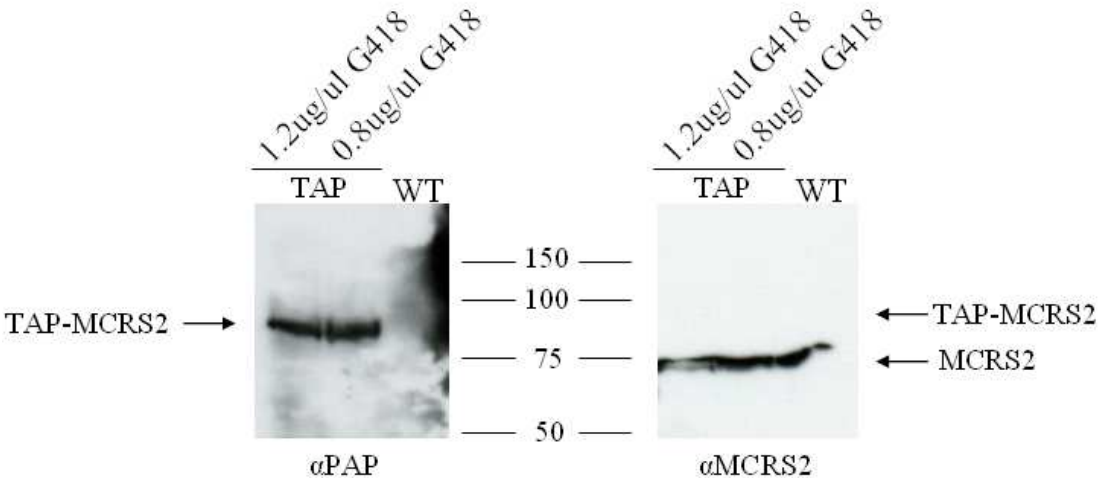


Figure 10. Expression of endogenous and TAP-tagged MCRS2 under 1.2 and 0.8 mg/ml of geneticin treatment in transfected SL2 cells, comparing to the wild type control (WT). Whole cell extracts from identical cell numbers ($5x10^5$) are loaded in each lane. Western blots were probed with peroxidase-anti-peroxidase (PAP, Sigma) diluted 1:2000 for detection of the protein A region within the TAP tag, and αMCRS2 antibodies for detection of both endogenous and tagged MCRS2. TAP-MCRS2 is expressed in the stably transfected cell lines as shown on membranes probed with PAP antibodies, but at a low level as compared to endogenous MCRS2 and can be hardly detected.

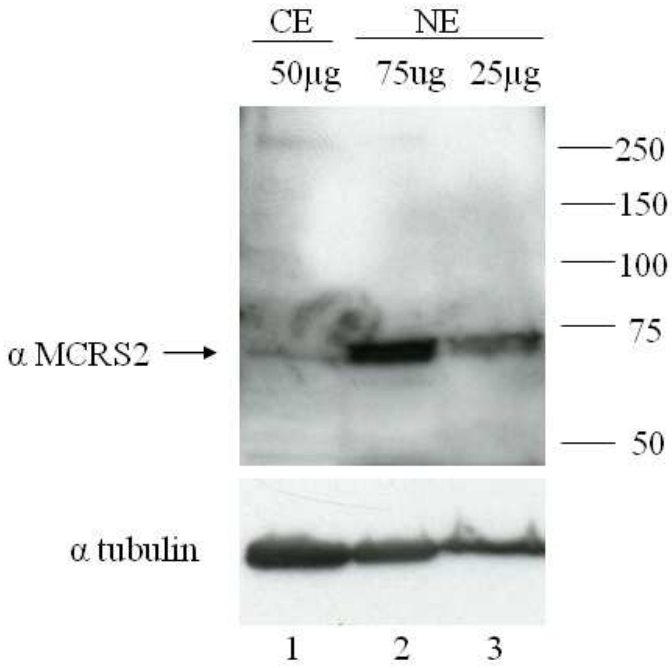


Figure 11. Nuclear localization of MCRS2. The presence of MCRS2 in nuclear (nuclear extract, NE) and cytoplasmic (cytoplasmic extract, CE) fractions was determined by western blotting. The blot was also probed for the cytoplasmic protein tubulin, to establish that fractionation had been achieved.

3.4.2. Optimization of the purification procedure

Prior to isolating TAP tagged MCRS2 containing complexes from the stable cell line, conditions were established to determine that the procedures used to isolate the complex would function. Figure 9 shows the efficiency of TAP-MCRS2 binding to IgG beads. Nuclear extract from TAP-MCRS2 expressing cell line was incubated with IgG beads, after which washed beads were boiled in SDS loading buffer and supernatant loaded on a gel (Fig. 12, lanes 3 and 5). Nuclear extract from wild type SL2 cells was taken as a negative control and treated in the same way (Fig. 12, lanes 4 and 6). The results show, that despite very low amount of TAP-MCRS2 in the starting material that prevents easy visualization of tagged protein, TAP-MCRS2 efficiently binds to IgG beads.

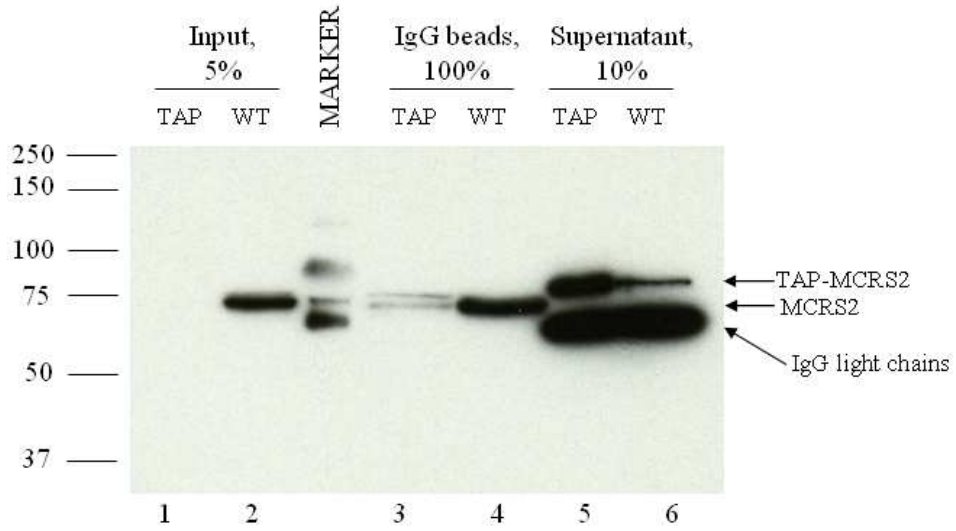


Figure 12. IgG binding of TAP-MCRS2. The western blot is probed with αMCRS2 antibodies. Because protein concentrations of starting materials of nuclear extract from TAP-MCRS2 cell line and wild type SL2 cells were not equalized, endogenous MCRS2 protein is not detected in the input of TAP-MCRS2 extract, but is present, however in very low amounts, in the supernatant after incubation of the extract with IgG beads due to the lager volume that was loaded (5% or the extract loaded as input, comparing to the 10% of the extract after incubation with IgG beads). TAP-MCRS2 is specifically enriched on IgG beads (100% of eluted from beads material is loaded).

Next, the efficiency of TAP-MCRS2 binding to calmodulin beads was determined. Nuclear extract from TAP-MCRS2 expressing cells as well as from the wild type SL2 cell line were incubated with calmodulin beads. Afterwards, beads were boiled in SDS loading buffer and 100% of material was loaded on an acrylamide gel (Fig. 13,

lanes 3-6), together with 10% of input (Fig. 13, lanes 1-2) and 10% of supernatant after incubation of calmodulin beads (Fig. 13, lanes 7-9). The binding of TAP-MCRS2 to calmodulin was not very efficient and could not be increased by modifications to the protocol (changing salt concentrations in buffers, trying different amounts and batches of calmodulin beads), indicating that it would be necessary to start with a large amount of material to produce sufficient protein for determination of protein constituents of the complex by mass sequencing.

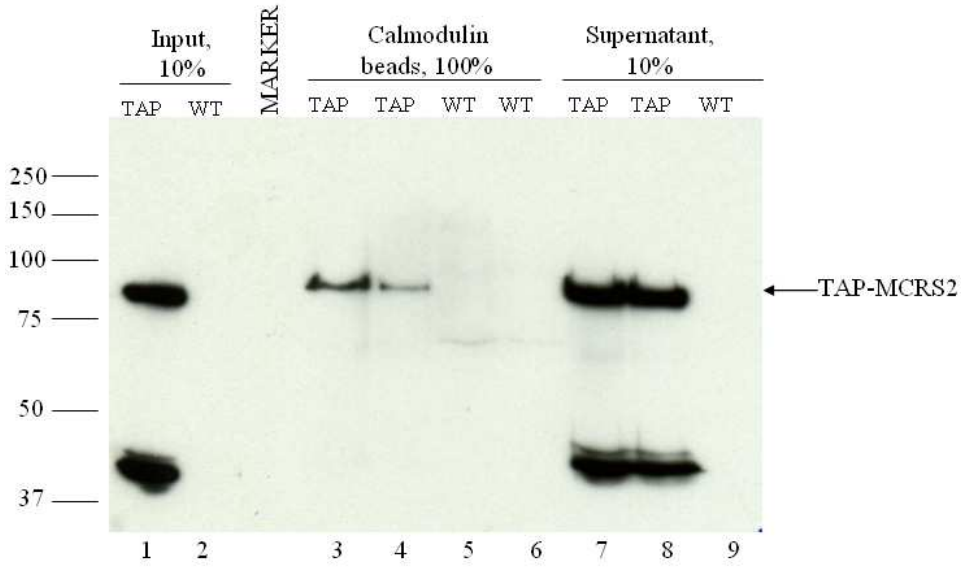


Figure 13. Verification of the binding of TAP-MCRS2 to calmodulin beads. Western blot is performed with PAP antibodies. 10% of starting material from both extracts (TAP-MCRS2 cell line and a wild type SL2 cells) was loaded as an input sample. Washed calmodulin beads were boiled in SDS loading buffer and 100% of material was loaded. To check how much unbound to calmodulin beads TAP-MCRS2 is left in the nuclear extract from TAP-MCRS2 cell line, 10% of the supernatant after incubation with calmodulin beads was loaded as well. Experiment is done in duplicate with two different batches of calmodulin beads to find the one that is better in binding TAP-MCRS2.

3.4.3. Biochemical purification of the NSL complex

Proteins, associated with MCRS2, were isolated from nuclear extracts from the cell line, stably expressing TAP tagged MCRS2. For each TAP purification, $2x10^9$ cells were required to prepare 1 ml of extract with approximately 6 μg/μl concentration of proteins. Nuclear extract from wild type SL2 cells was subject to the same procedure, and used as a mock control for TAP purification. Purified eluted material was run on a gel and silver stained (Fig. 14, lanes 3-4). Identification of proteins was done by

excision of specific to TAP MCRS2 purification bands (performed by Sven Fraterman, EMBL-Heidelberg), and by analysis of total complex elutions (Adrian Cohen, NCLMS, Netherlands).

Mass spectrometric sequence analysis of eluted bands revealed that MCRS2 (bait) associates with NSL1, NSL2, NSL3, WDS and MBDR2 (Table 2). None of these proteins were present in material obtained from untagged control mock purification from wild type SL2 cells. Interestingly, all of these proteins were purified together with MOF (Mendjan et al. 2006), and there were no new proteins found in TAP MCRS2 purification.

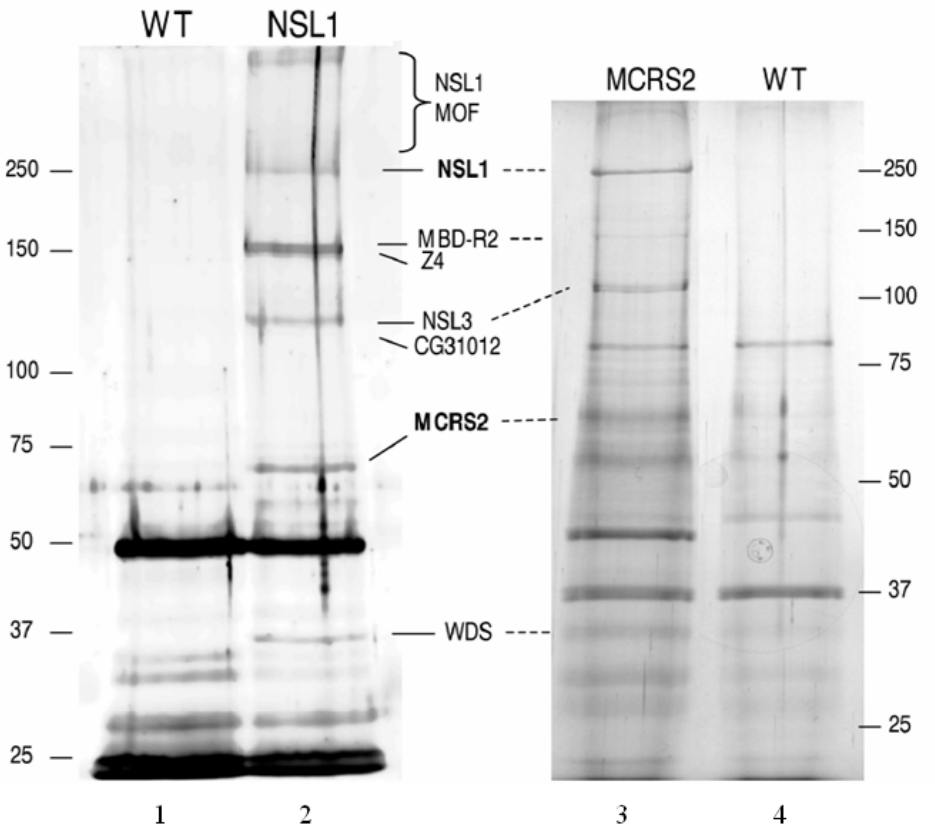


Figure 14. Silver stained gel of FLAG/HA purification of NSL1 (lanes 1-2) and TAP purification of MCRS2 (lanes 3-4); WT indicates corresponding mock purifications. 1.5 ml of nuclear extract from each cell line is used with the concentration of 6 µg/µl, and 50% of the purified eluted material was loaded on a gel; the rest of the material was left for subsequent mass spectrometry analysis.

At the same time, sequential FLAG/HA purification of nuclear extract from a stable cell line expressing TAP-FLAG-HA-NSL1 was performed in the lab by Philipp

Gebhardt. It also resulted in isolating the same set of proteins – NSL1 (bait), NSL2, NSL3, WDS, MCRS2 and MBDR2 – as well as MOF, Z4 and Chromator. Thus, the existence of the complex was confirmed by the purification performed of another tagged member of the complex. For the control, mock purification from nuclear extract of wild type SL2 cells was used (Fig. 14, lanes 1-2).

Table 3. Proteins identified in TAP-MCRS2 purification.

Name	Acession number	Mass [Da]	Mascot score (Number of peptides)	Number of experiments present (out of 6)	Domains
NSL1	gi|24647245	170587	48 (12)	4	coiled coil, PEHE
NSL2	gi|23172607	57054	202 (3)	3	two C/H-rich domains
NSL3	gi|17862340	114582	38 (12)	4	a/b hydrolase fold
MCRS2	gi|16767858	63840	127 (5)	6	forkhead-associated domain (FHA)
MBDR2	gi|45551883	120748	53 (8)	4	CHAP1, 2 3 Tudor, MBD1, ZnF, PhD finger
WDS	gi|6946677	39530	241 (5)*	4	seven WD40 repeats

*number taken from another experiment

3.4.4. NSL complex composition

Results of purifications as well as those of immunoprecipitation show the existence of a novel complex composed of the evolutionary conserved proteins: MOF, NSL1, NSL2, NSL3, MCRS2, MBDR2, WDS, Z4 and Chromator. MSL proteins are not present in these purifications, indicating that they are not a part of the same complex. The new complex has been named the NSL complex, and the next section describes its members.

NSL1 is a novel protein of 170 kDa, composed of 1570 aa. NSL1 has a putative coil-coiled domain and a PEHE domain at its C-terminus. In *Drosophila*, the only other protein with a PEHE domain is MSL1 (Marin 2003). Just like MSL1, NSL1's interaction with MOF might most probably be via the PEHE domain, as it was shown

to be true for human homologues: PEHE domain of hNSL1 interacted directly with hMOF in a GST pulldown (Mendjan et al. 2006).

NSL2 is a novel protein of 127 kDa, composed of 1200 aa, that is evolutionary conserved from flies to humans. It is ubiquitously present in mice and in *Xenopus laevis* embryos from stage IV onwards (Shim et al. 2000; Mata et al. 2003). The protein has no pronounced domain structure, with sequence analysis only revealing an evolutionary conserved region that is rich in cysteine and histidine motif (Taipale and Akhtar 2005).

NSL3 is a novel protein of 114 kDa, composed of 1066 aa. It contains an α/β-hydrolase domain (Taipale and Akhtar 2005), this is one of the largest and oldest structural domain superfamilies that share a common fold and catalytic triad (Nardini and Dijkstra 1999). In NSL3, only the β-hydrolase common fold is retained; the catalytic triad is not. NSL3 may not have enzymatic activity, but could possibly retain the ability to recognize substrates. Additional functional and structural studies would be necessary to determine more conclusive answers.

MBDR2 is a poorly characterized protein composed of 1081 aa (120 kDa). It contains DNA-binding, two Tudor and methyl binding (MBD) domains, as well as C2H2-type zinc-finger and a PHD finger (Taipale and Akhtar 2005). This suggests that MBDR2 might have a spectrum of functions. MBD domains bind to methylated DNA and are involved in transcriptional repression in mammals (Bird 2002). In *Drosophila*, the function of DNA methylation is not clear as, in general, DNA methylation happens much more seldom than in mammals (Lyko et al. 2000). Tudor domains have been shown to bind methylated arginine residues in Sm proteins involved in splicing, and methylated H3K79 in yeast (Huyen et al. 2004). Tudor domains share similarity to chromodomains, which also can bind methylated residues (Lachner et al. 2001).

WDS is the smallest protein in the complex, with only 361 amino acids (39 kDa). It is an evolutionary conserved protein, present from *Arabidopsis thaliana* to mammals. WDS belongs to the WD protein family and contains seven WD40 repeats (Hollmann et al. 2002). These repeats are modules involved in protein-protein interaction and are present in a variety of chromatin-associated complexes (Cao et al. 2002). The

mammalian ortholog, WDR5, binds specifically to dimethylated H3K4, and is also a constituent of H3K4-specific methyltransferase complexes (Wysocka et al. 2005). Methylated H3K4 in *Drosophila*, as well as in other organisms, is linked to transcriptional activation (Zhang and Reinberg 2001). Similar to MBDR2, the presence of WD40 repeats suggests possible functions of WDS in *Drosophila*, namely enhancing transcription from male X chromosome by bringing MOF to dosage compensated genes to induce local hyperacetylation and increase gene expression. WDS is ubiquitously expressed in *Drosophila* at all developmental stages and is essential for viability.

Z4 is a 105 kDa protein of 996 aa. Z4 has 7 zinc-fingers, similar to the insulator binding protein CTCF, which is a protein involved in organizing chromosomal domains (Eggert et al. 2004). Z4 is essential for *Drosophila* development and acts in a dose-dependent manner on the development of several tissues. It is involved in chromosome compaction and higher-order chromatin structure formation (Eggert et al. 2004). Z4 mutant flies loose their band/interband pattern on chromosomes, with the interband chromatin exhibiting an overall decompaction of chromosomal material. Z4 mutants also have dose-dependent defects on position effect variegation. Z4 can be immunoprecipitated with Chromator, suggesting that that both proteins might be responsible for the chromosome higher-order structure during interphase (Eggert et al. 2004).

Chromator/Chris consists of 926 aa and has a molecular weight of 101 kDa. It is essential, and is ubiquitously expressed in Drosophila (Rath et al. 2004). Like MOF and MSL3, it also contains a chromodomain. Chromator localizes on polytene chromosomes together with the interband-binding protein Z4 (Gortchakov et al. 2005). During mitosis, Chromator detaches from the chromosomes and aligns in a spindle-like structure. The C-terminal half of Chromator, lacking the chromodomain, is sufficient for both nuclear and spindle localization. Chromator is an essential protein; RNAi depletion of Chromator in SL2 cells induces abnormal microtubule spindle morphology and chromosome segregation defects (Rath et al. 2004).

3.5. NSL1 directly interacts with MCRS2 and MOF

In order to dissect the interactions between the NSL proteins, baculovirus expression system was used. Constructs with tagged (FLAG or HA), as well as untagged NSL1, MCRS2 and MOF were expressed in baculovirus system. Copurification of the proteins revealed that stable interaction of NSL1 can be detected with MCRS2 and MOF (Fig. 15, lanes 1-2), which resists 200 mM salt. It showed that interaction between NSL1-MCRS2 and NSL1-MOF is direct and present between these proteins also when there are no other NSL complex members present. Interestingly, presence of a tag can influence the result of copurifications. Thus, by reciprocal tagging of MCRS2, NSL1 and MOF, it was observed that N-terminal tagging of NSL1 decreases the interaction of this protein with both MOF and MCRS2. This difference most probably comes from the sterical obstructions that tags bring to the proteins, preventing protein interactions.

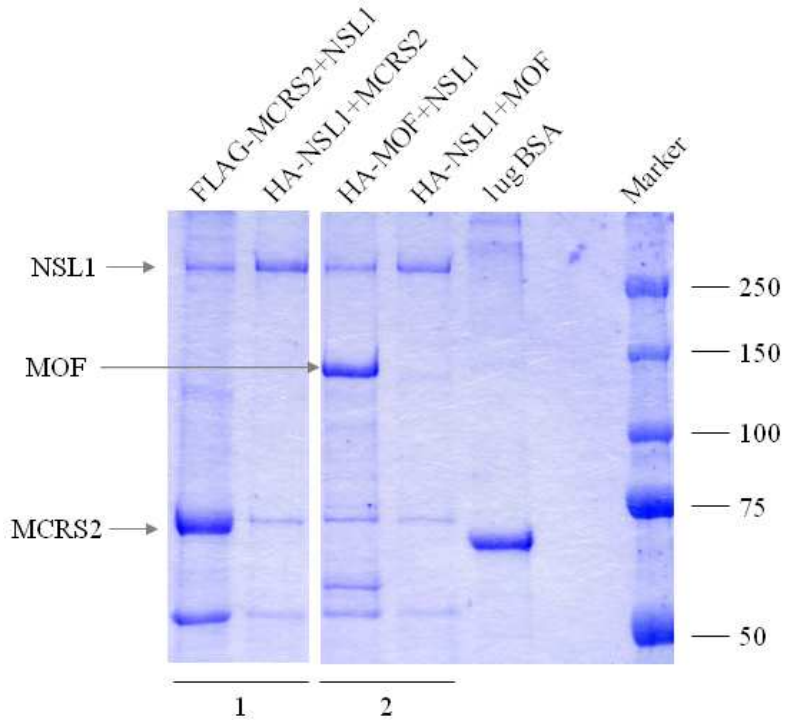


Figure 15. Reconstitution of NSLs interaction using baculovirus-expressed proteins. NSL1 interacts with (1) MCRS2 and (2) MOF upon incubation of protein extracts. After purification via corresponding tag, proteins were run on a gel and Coomassie stained. Experiment is performed by Herbert Holz.

3.6. NSLs bind to multiple sites on chromatin

Next, the *in vivo* localization of NSLs was investigated by immunostaining salivary gland polytene chromosomes of *Drosophila* male and female larvae with antibodies directed against MCRS2, MBDR2 and NSL1. As it is shown on the figure 16, these proteins broadly decorate all chromosomes, comprising autosomes and sex chromosomes. Merge images show an overlap of the proteins with each other on many binding sites.

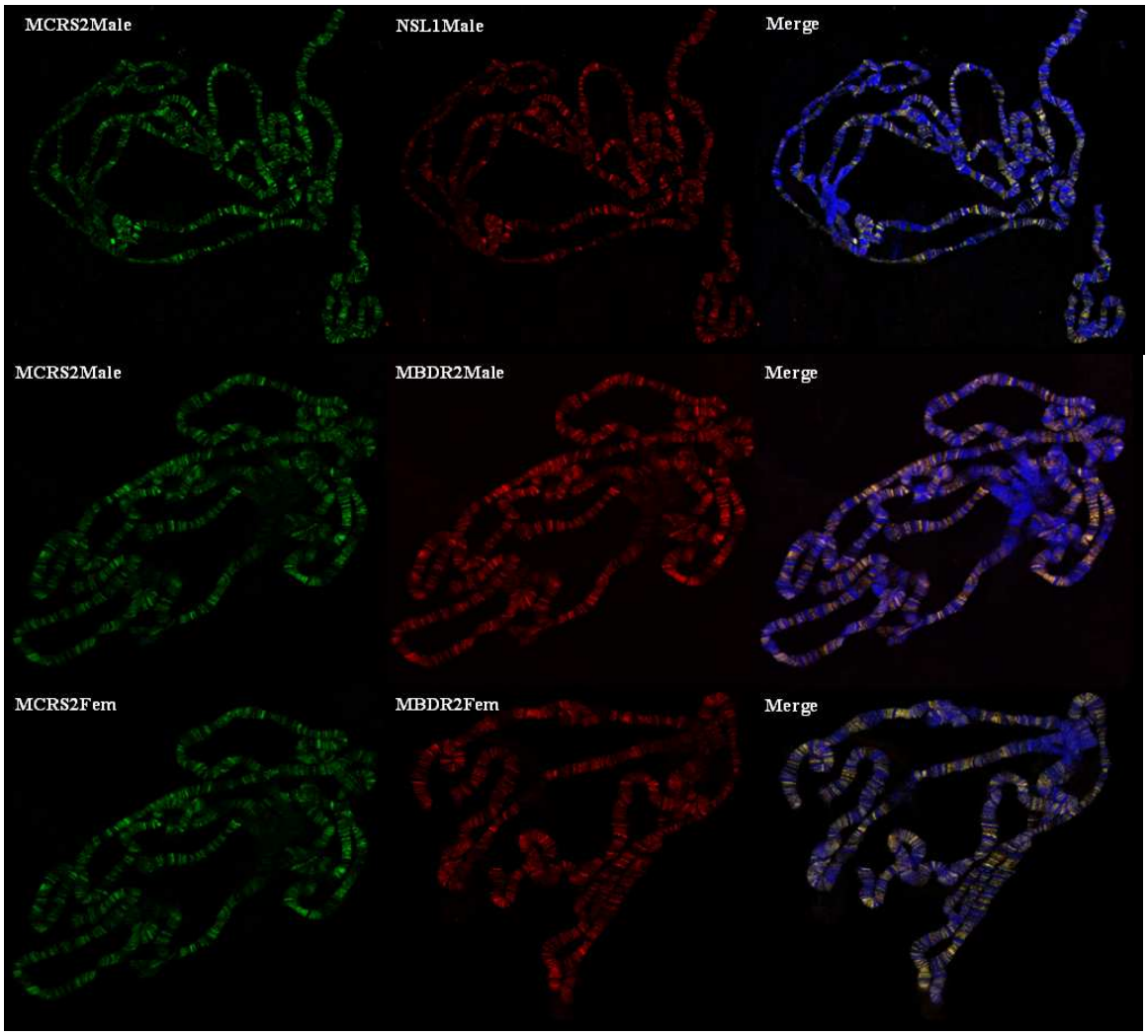


Figure 16. Colocalization of MCRS2, MBDR2 and NSL1 on wild type male and female *Drosophila* 3rd instar larvae on polytene chromosomes. Confocal microscopy is performed on polythene chromosomes of 3rd instar wild type *Drosophila* larvae immunostained with MCRS2 MBDR2 and NSL1. DNA is visualized by Hoechst.

The NSLs were found in the MOF TAP purification (Mendjan et al. 2006). In addition, immunoprecipitation experiments discussed above indicate interaction

between MOF and NSLs. NSL1 purification also resulted in identifying MOF in elutions. Genome-wide MOF profiling analysis carried in the lab (Kind et al. 2008) revealed many sites of MOF binding on autosomes. It was therefore interesting to see if NSL proteins coreside with MOF on same chromosomes. For this reason, immunostaining of *Drosophila* salivary gland polytene chromosomes with antibodies directed against MOF and MCRS2 was performed. As shown on the figure 14, it reveals many overlapping positions of both proteins on all chromosomes, with MOF being enriched on male X chromosome.

In contrast to MSL complex, which localizes to the male X chromosome, the NSL proteins are widely binding all chromosomes in males and females. Immunostaining with MSL1 protein (Fig. 17), a member of the MSL complex, shows a specific recognition of the male X chromosome, and MCRS2 staining on male X does not completely overlap with that of MSL1, representing their presence in different complexes.

Altogether, these data raise the possibility of an additional function of the MOF protein in *Drosophila*, independent of MSLs and, thus, dosage compensation, which it performs together with NSL complex.

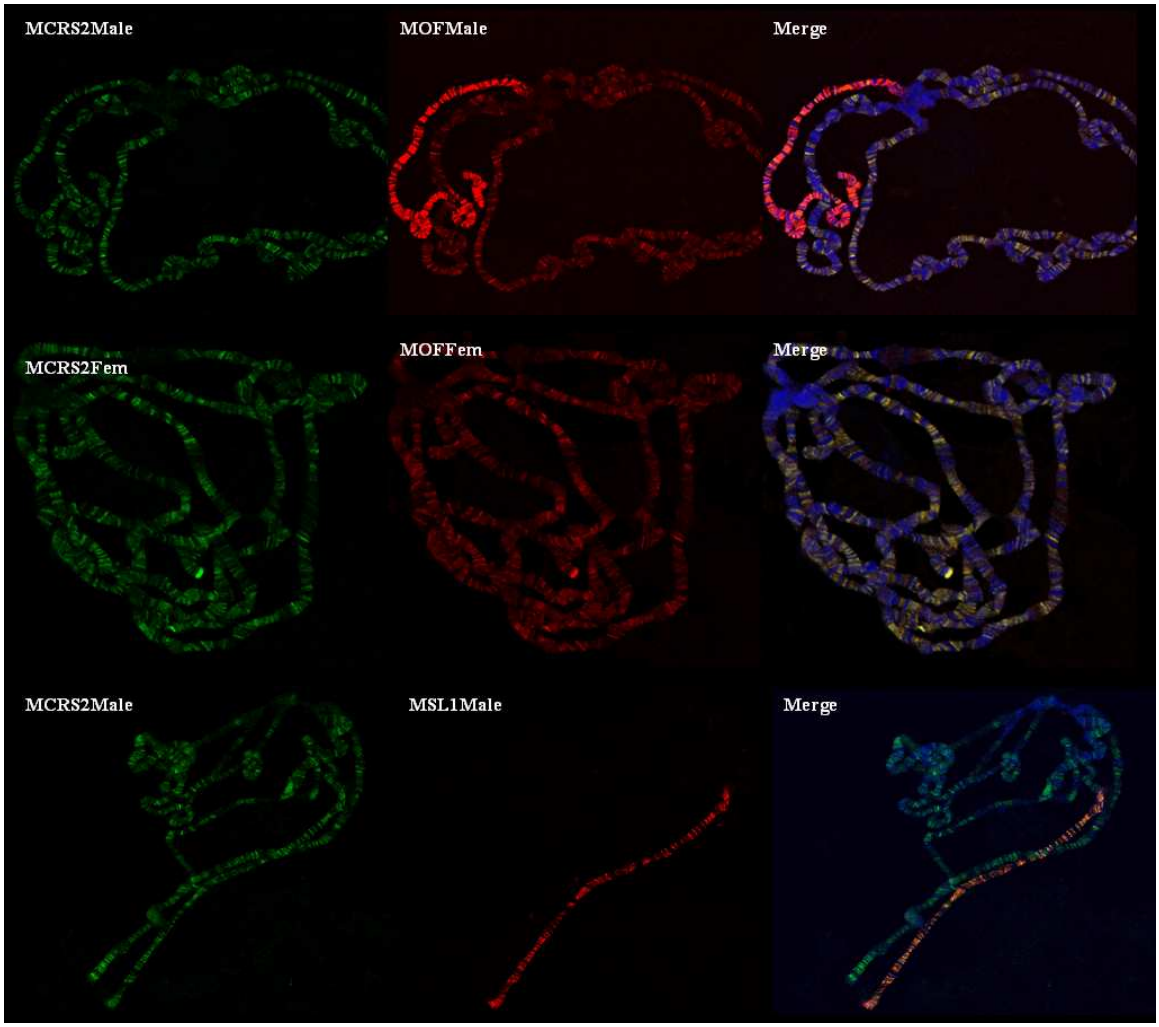


Figure 17. Colocalization of MCRS2 and MOF on *Drosophila* 3rd instar larvae polytene chromosomes in males and females. As a control of a binding specificity, coimmunostaining of MCRS2 and MSL1 is shown. Confocal microscopy is performed on polytene chromosomes of 3rd instar wild type *Drosophila* larvae immunostained with MCRS2, MOF and MSL1, and additionally stained with Hoechst to visualize DNA.

3.7. The NSL complex binds to promoters of MOF target genes

The immunostaining of polytene chromosomes presented above provides a global picture of NSL proteins binding on chromosomes. To gain a higher resolution of the binding sites of the complex on chromatin, a chromatin immunoprecipitation (ChIP) analysis was performed. Antibodies against MCRS2, MBDR2, NSL1 and MOF, in conjunction with their corresponding pre-immune sera serving to control the experiments, were used for ChIP experiments.

For this analysis, chromatin was prepared from 3rd instar larva glands from *Drosophila* wild type male and was used for the consistency of results obtained from the immunostainings. Besides, it was interesting to see the binding profile of the NSL complex in the differentiated tissue. Sheered formaldehyde cross-linked chromatin gave fragments ranging in size from several thousand bp to tens of bp, with a small peak at 200 bp (Fig. 18, A; for a detailed protocol of chromatin preparation and chromatin immunoprecipitation see Materials and Methods). To immunoaffinity isolate chromatin associated with particular components, sheared chromatin was incubated with antibodies of interest or with their corresponding pre-immune sera, immunocomplexes were recovered on protein A/G sepharose beads and DNA was then purified. Quantitative PCR on Gprk2, a gene discovered through whole genome analysis of sites bound by MOF (Kind et al. 2008), was used to determine the proportion of material that could be recovered by immunoaffinity purification (Fig. 18, B). Primers were designed to probe the occupancy of MOF, NSL1 and MCRS2 on the promoter region, body of the gene and 3' end of the coding sequence. The reason for such a combination of primers comes from the studies in the lab, where it was shown that MOF has a bimodal distribution on the genes where it is bound to. On X-linked genes MOF is present through out the gene with two peaks: on promoter and 3' end. On autosomal genes MOF binds only to promoters. It was interesting to compare profiles of NSLs binding with that of MOF. As a result of the experiment, NSL complex proteins were observed specifically bound to promoter of Gprk2 gene, with a high percentage of input recovery.

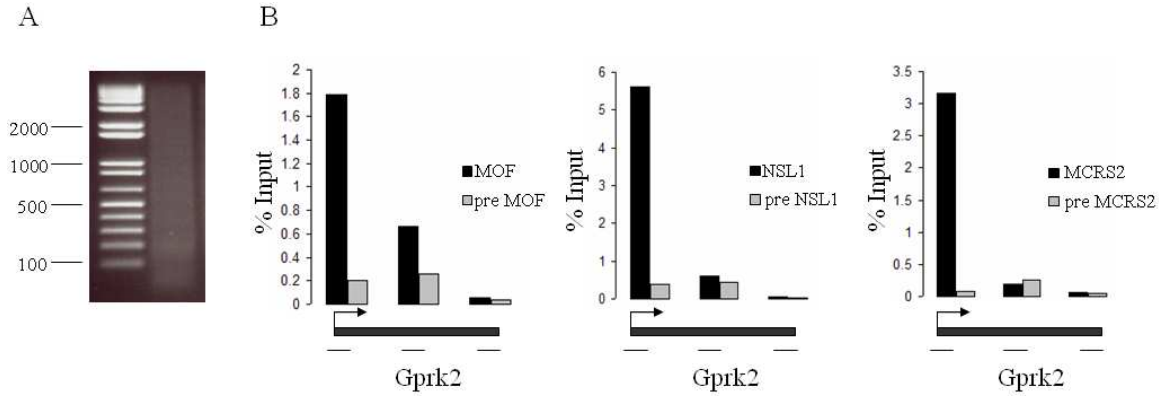


Figure 18. (A) Chromatin isolated from salivary glands. Formaldehyde crosslinked chromatin was sheared, the crosslinking reversed and the resulting material was treated with RNAse A and proteinase K prior to separation on a 0.8% agarose gel. (B) Chromatin immunoprecipitation performed on *Drosophila* wild type male larva glands using MOF, NSL1 and MCRS2 antibodies with respective preimmune sera. The quantity of purified DNA was determined by real time qPCR with three primer pairs spanning the promoter, middle and end of the Gprk2 gene. Results are presented as a percentage of input material used (primer sequences are listed in Materials and Methods).

In parallel, same set of experiments was performed with the chromatin from SL2 cell line, as it was interesting to see if there are differences in NSL binding profile in undifferentiated cells of embryonic origin. In conclusion, same profile was observed for the NSL binding between salivary glands and SL2 cells, and it was found out that the proteins were present on the promoter region of MOF-bound genes, as shown on the figure 19 on the example of Gprk2 gene. However, chromatin from salivary glands gave much higher levels of recovery, it was therefore decided that further ChIP experiments would be performed only with chromatin from larval salivary glands.

Having the material and established the conditions for ChIP, more genes were checked for the presence of the NSL complex. The location in relation to gene structure was determined using qPCR to scan the proximal promoter, gene body and 3' end of transcripts. Genes known to be bound by MOF were chosen as primary targets, on the X chromosome and on autosomes, according to ChIP-chip data produced in our lab (Kind et al. 2008). For negative controls, genes not bound by MOF genes were also characterized (Table 4).

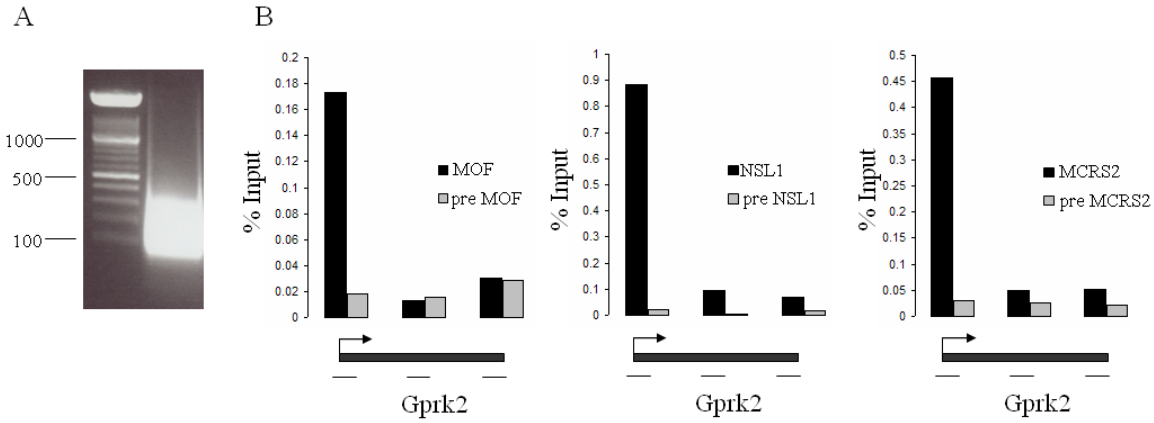


Figure 19. (A) Chromatin isolated from SL2 cells. Formaldehyde crosslinked chromatin was sheared, the crosslinking reversed and the resulting material was treated with RNAse A and proteinase K prior to separation on a 0.8% agarose gel. (B) Chromatin immunoprecipitation performed on male larva glands of *Drosophila* wild type using MOF, NSL1 and MCRS2 antibodies with respective preimmune sera. The quantity of purified DNA was determined by real time qPCR with three primer pairs spanning the promoter, middle and end of the Gprk2 gene.

Table 4. MOF-bound and not-bound genes looked at NSL complex localization.

Genes bound by MOF		**X chromosomal genes not bound by MOF**
X chromosomal	**Autosomal**	
CG6506	CG4245	CG6398
CG4406	CG9536	OdsH
CG32560	Sec5	Runt
Dspt6	HBS1	
Rb	NSL3	
roX2	Gprk2	

The results of qPCR performed on ChIP purified material are summarized in the Figure 20. The data presented is an average of five experiments that used five independently prepared chromatin samples. It shows that, within this restricted data set, MOF associates over the entire transcribed region of X chromosomal genes, but is present only at promoter regions of autosomal genes, as previously described (Kind et al. 2008). NSL1, MCRS2 and MBDR2 are present on the same genes as MOF; however, in contrast to the profile of MOF on the X chromosome, they recognize only promoter regions, independently of the chromosomal location of the gene.

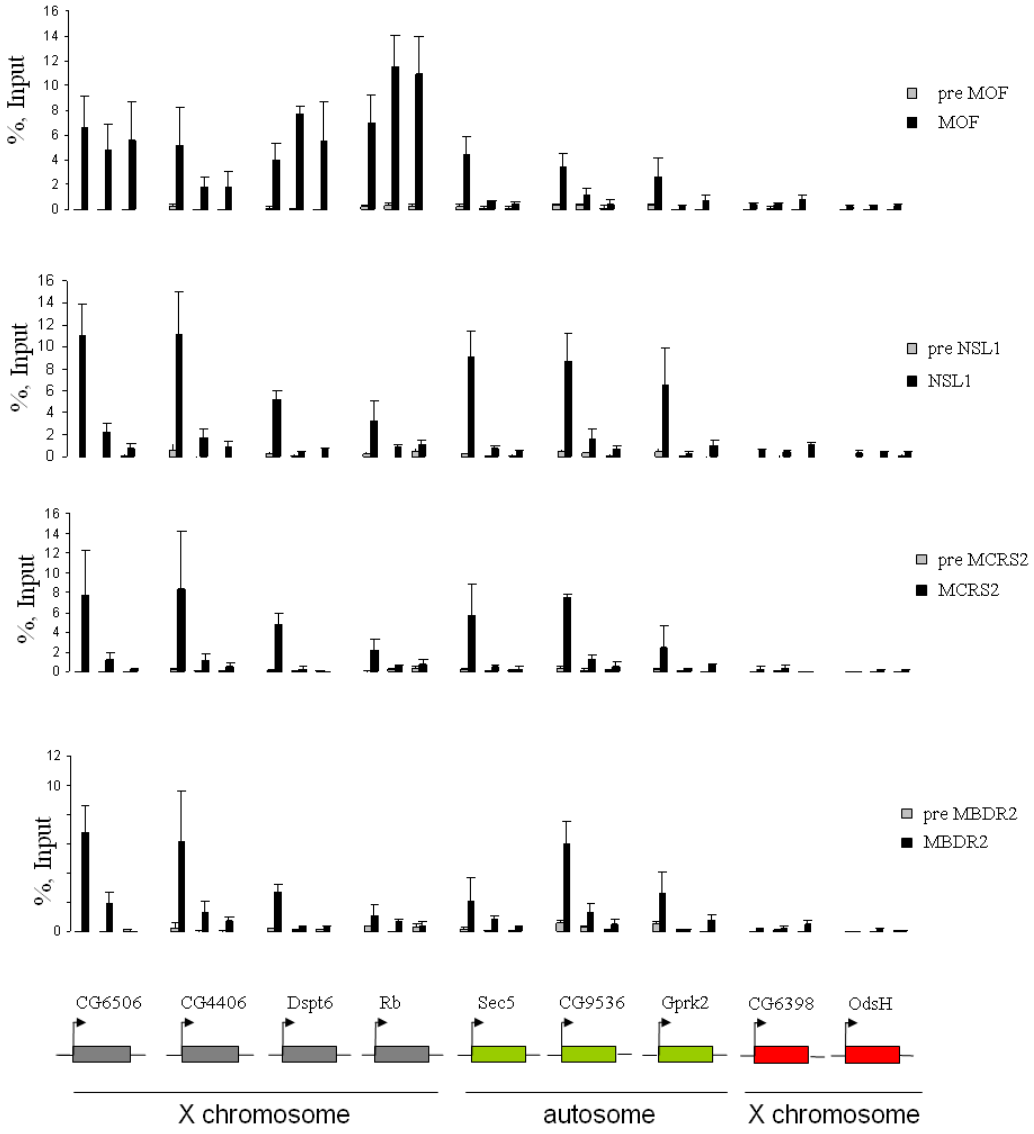


Figure 20. ChIP analysis from larval salivary glands using antibodies against MOF, NSL1, MCRS2 and MBDR2. Preimmune sera is used as a negative control for each antibody. Immunoprecipitated DNA was amplified by real-time PCR with primer sets indicated in the material and methods. Each gene is evaluated at promoter region, middle of the transcribed sequence and at the 3' end. Percent input is determined as the amount of immunoprecipitated DNA relative to input DNA. The data presented is an average result from five biologically independent experiments. The standard deviations of the replicate results are shown as error bars.

The ChIP analyses performed on these selected genes is limited, as firstly requiring prior knowledge of which genes are targeted by particular complex, and secondly, only a limited number of genes can be characterized by real-time quantitative-PCR. To see if the promoter binding is a general phenomenon, the comprehensive, genome-

wide determination of binding sites for the NSL complex will be obtained massively parallel sequencing of DNA enriched for association with NSL components (ChIP-seq). This work is in progress.

3.8. NSLs and transcription regulation

Results of ChIP with MCRS2, MBDR2 and NSL1 revealed that NSL complex members are localized on promoters of autosomal and sex chromosomes. In addition, the polytene chromosome staining suggests that there are plenty of sites on the genome where they bind on both X and autosomes. One of the most obvious questions that arise is whether the NSL complex is involved in the regulation of gene expression. To address this question, fly lines expressing RNAi were used. In collaboration with Sunil Jayaramaiah Raja, using RNAi mediated depletion of MCRS2, MBDR2 and NSL3 in salivary glands, it was shown that expression of many target genes is affected by depletion of the NSL components and that, as anticipated, expression is downregulated when the NSL complex is compromised (Fig. 21).

Several genes, based on the MOF presence, as well as chromosomal location, were used in this study. Two groups of genes were chosen on the X chromosome: MOF bound Ucp4a, Dspt6, Par6, RB, CG2967, CG6506, and CG4406; not bound by MOF CG14788 and CG12065. Similarly, two groups of autosomal genes were chosen: MOF bound Gprk2, CG9536, Sec5, PI3K92E, Eyg, NSL1, NSL2, NSL3 and MCRS2; not bound by MOF – MBDR2 and Nrv2. The efficiency of knock down was around 80% for MCRS2, MBDR2 and NSL3, as shown on the figure 20. It was found, that irrespectively of the gene location and the gender, expression of MOF bound genes is downregulated more than two fold upon RNAi knockdown of the NSL components comparing to the control. Interestingly, the experiment also shows that expression of NSL proteins themselves is effected by the knock down of each other.

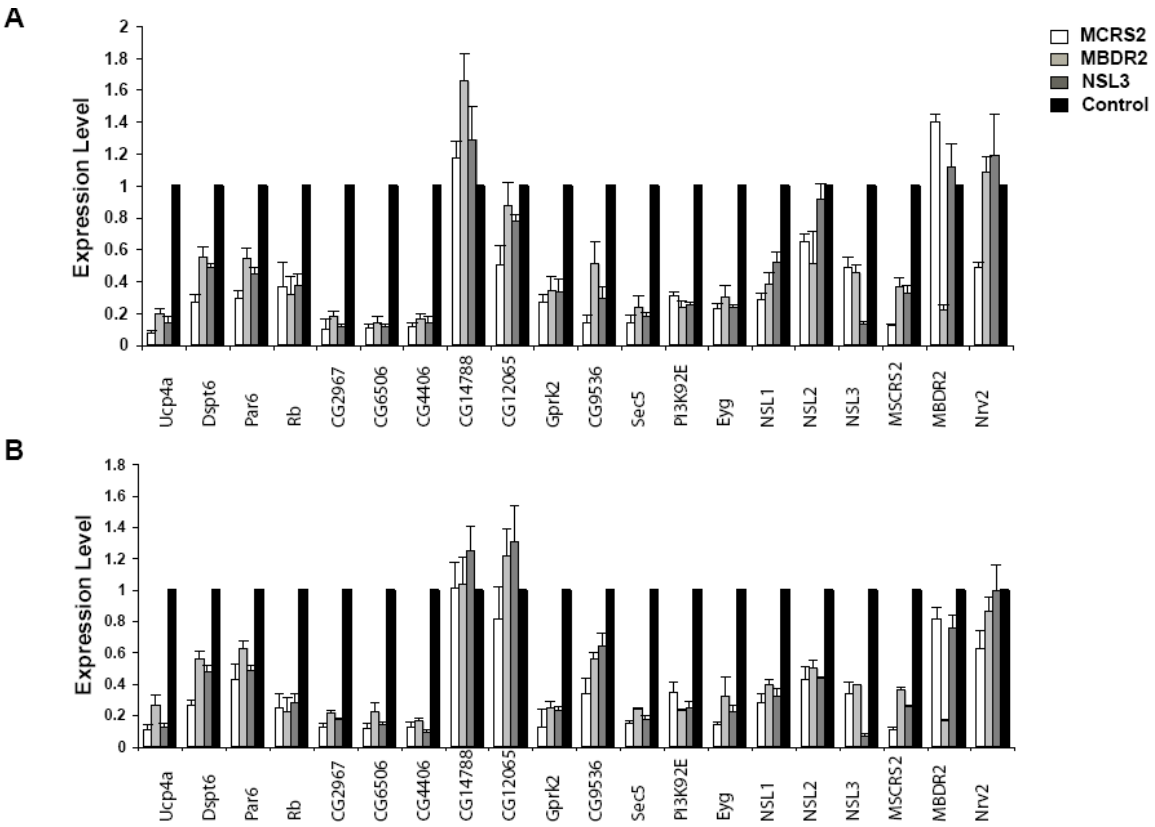


Figure 21. Expression analysis of X chromosomal (Ucp4a – CG12065) and autosomal (Gprk2 – NSL2) genes upon MCRS2, MBDR2 and NSL3 knock down in salivary glands, in male and female *Drosophila* 3rd instar larvae. CG14788 and CG12065 are autosomal genes not bound by MOF; MBDR2 and Nrv2 – X chromosomal not bound by MOF genes. Expression of many MOF target genes is downregulated more than two fold upon RNAi knockdown of the NSL components comparing to the control.

These results suggest that the NSL complex members act as transcriptional co-activators on MOF target genes. Upon depletion of MCRS2, MBDR2 or NSL3 components of the NSL complex, the expression levels of many X chromosomal as well as autosomal genes in both males and females are reduced. The complex is involved in regulation of a wide number of genes and its effect is likely to be more general than the MSL complex, whose activity is restricted to the male X-chromosome.

3.9. NSLs affect each other's stability

Decreased RNA levels of the NSLs upon MCRS2 knock down were then compared to the protein levels of the respected genes. For this, RNAi mediated knock-down of MCRS2 in SL2 cells was performed and the level of knock down was checked in Western blot by probing membranes with MCRS2 antibodies. MCRS2 knock down was compared to the EGFP knock down. As shown on the figure 22, the efficiency of the MCRS2 knock down is around 90%. Analysis of the cellular levels of MBDR2, Z4, Chromator, MOF and WDS by western blot determined that these components of the NSL complex are reduced when MCRS2 is depleted; cellular levels of tubulin are unaltered by RNAi treatment (Fig. 22). Additional analysis of other members of the NSL complex in the MCRS2 knock down on (data not shown) shows that among other effected proteins are NSL2 and NSL3.

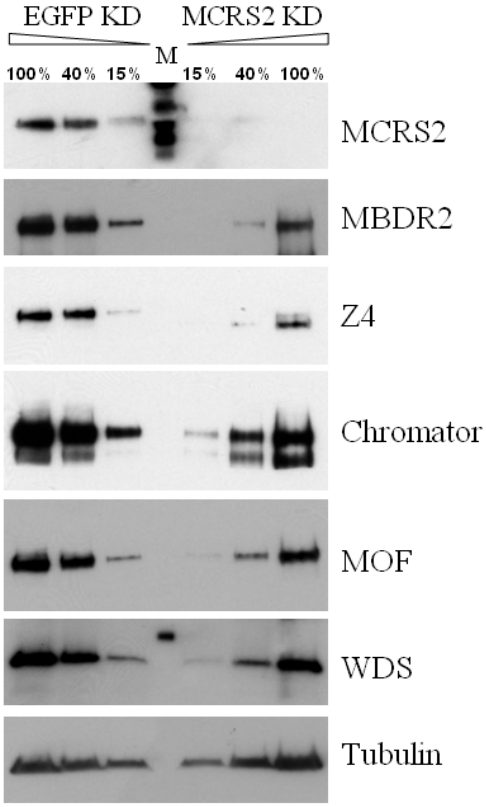


Figure 22. RNAi mediated depletion of MCRS2 reduces the intracellular level of other components of the NSL complex. SL2 cells are treated with EGFP or MCRS2 dsRNA with components of the NSL complex then analyzed by western blot. Tubulin is used as a control protein whose stability should not be affected by siRNA treatment.

To show that effect of MCRS2 knock down is specific to the NSLs, it was compared to the one of MSL1. As shown on figure 23, the efficiency of MCRS2 knock downs was around 80%, while the one of the MSL1 was almost complete. Probing same membranes with MBDR2 antibodies, it was shown that the effect of MCRS2 knock down is specific to the NSL complex, as indicated by decreased levels of the MBDR2 protein, and it does not influence protein levels of MSL1. MSL1 protein levels stay unchanged upon MCRS2 knock down; reciprocally, MCRS2 protein level is also not affected in the MSL1 knock down cells (Fig. 23).

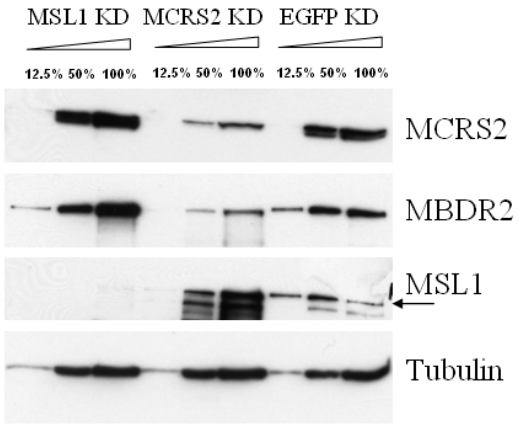


Figure 23. RNAi mediated reduction in MSL1 and MCRS2 does not influence each others protein levels, suggesting that MSL and NSL complexes operate independently of each other. Control cells are treated with EGFP dsRNA. Tubulin is used as a control protein, unaffected by siRNA treatment.

4. DISCUSSION

In *Drosophila*, dosage compensation is achieved by the MSL complex, composed of the MSL1, MSL2, MSL3, MOF and MLE proteins together with noncoding RNAs – *roX1* and *roX2*. It achieves dosage compensation by specifically binding to the male X chromosome and inducing global hyperacetylation of H4K16, which is associated with transcriptional activation. Recently, a connection between the dosage compensation complex and a new group of proteins was found through purification of MOF containing complexes, isolated from *Drosophila* embryos, a *Drosophila* cell line and a human cell line (Mendjan et al. 2006). Within the same study, isolation of MSL3 co-purified a similar set of interaction partners, although several proteins identified in the MOF purification were not found when MSL3 was used as bait. The interconnection between proteins classically believed to be involved only in dosage compensation with the range of nuclear and transcriptional processes was reinforced by the observed evolutionary conservation of complex components between fly and human, despite differing mechanisms of achieving dosage compensation in these organisms. Collectively, the functional connections between MOF and a large set of proteins suggest that it has functional roles in gene regulation beyond dosage compensation.

4.1. Purification of a novel complex, termed the NSL complex

Several proteins that purify with MOF, but not with MSL3, had not been previously characterized, those include MCRS2, NSL1, NSL2 and NSL3. Immunoaffinity purification of MCRS2 co-precipitated MOF, MBDR2, NSL2 and NSL3 from nuclear extracts, whereas no interaction of MCRS2 with MSL1 or MSL3 was detected. It suggested that there are different complexes where MOF is residing. Tandem affinity purification MCRS2 showed that, indeed, there is a novel complex consisting of several proteins, namely NSL1, NSL2, NSL3, MBDR2 and WDS, in addition to MCRS2. All of these proteins had been previously identified in purification with MOF. Surprisingly, no MSL constituents, apart from MOF, were co-isolated with MCRS2. Parallel purification of NSL1 revealed a complex with the same components as when MCRS2 was used as bait, although, in addition, MOF, Chromator and Z4 proteins were also co-purified. Chromator and Z4 were already known to co-purify

with MOF. Collectively, the data from complexes defined by purifications of MCRS2 and NSL1 show that a distinct complex exists. This novel assembly of proteins, the NSL complex, contains MOF, NSL1, NSL2, NSL3, MCRS2, MBDR2, WDS, Z4 and Chromator. The connection of MOF with another complex suggests that it has additional functions beyond dosage compensation.

Experiments to directly probe interactions between NSL complex members, using baculovirus-expression, showed that NSL1 copurified with MOF and MCRS2 in the absence of other complex members. This suggested that the interaction between these proteins is specific and direct. In addition, MCRS2 was found previously in our lab, by yeast two-hybrid technology, to interact with MOF, again indicating a direct interaction between MCRS2 and MOF.

4.2. Similarities between the MSL and NSL complexes

An interesting similarity between the MSL and NSL complexes is that both of them have a component with a rare PEHE domain in their structure: these are the NSL1 and MSL1 proteins. They are the only two proteins in *Drosophila* that have this domain, and it is shown, that interaction of them with MOF occurs though this domain.

To address the functional significance of the MCRS2 protein, studies of the complex in conditions of depleted MCRS2 were carried out. It was found, that upon RNAi mediated knock down of MCRS2 in SL2 cells, protein levels of MCRS2, MBDR2, Z4, Chromator, NSL2 and NSL3 are decreased. These results suggest interdependency between protein components of the NSL complex, which is reflected in their stability. Depletion of MCRS2 severely affects the amount and localization of other NSL complex members, indicating that MCRS2 might be a central component of the NSL complex, perhaps through nucleating complex assembly.

Interestingly, MCRS2 exhibits a similar effect on the NSL complex, as MSL2 has on the MSL complex. MSL2 is a crucial component of the MSL complex. It is present only in males, as in females translation of *msl2* mRNA is inhibited (Kelley et al. 1997; Gebauer et al. 2003; Grskovic et al. 2003; Beckmann et al. 2005). Without this protein, the dosage compensation complex is not assembled. The mechanism of the

effect of MCRS2 on the NSL complex still needs to be discovered. It could be that MCRS2 is involved in controlling the expression of other NSLs, or it provides a platform for the complex assembly, or there are other mechanisms that we still do not know.

4.3. NSL complex colocalizes with MOF on chromatin

Staining of polytene chromosomes shows that MSL proteins bind multiple sites on the male X chromosome and colocalize with MOF. However, the MOF staining pattern is slightly different: while being enriches on the male X chromosome, MOF has a broad distribution over all chromosomes in both sexes (Kind et al, 2008). These data suggested that MOF has an additional function, apart from the dosage compensation. Immunostaining of polytene chromosomes with antibodies against NSL complex members indicated that NSLs bind to all chromosomes in a broad pattern both in males and female. Interestingly, the chromosomal staining patterns of the NSLs overlap at many sites with that of MOF on X chromosome, as well as on all autosomes both in males and females.

The staining of the chromosome squashes provided valuable information concerning the localization of the NSL proteins towards MOF. In order to gain resolution of NSL binding sites at the level of individual transcriptional units, chromatin immunoprecipitation experiments using antibodies against the NSL proteins and MOF were performed. To do so, a set of MOF bound genes, provided by the MOF profiling done in the lab (Kind et al. 2008), was used. ChIPs performed to determine the binding of NSL1, MCRS2, MBDR2 and MOF show an association of these proteins with promoters of MOF bound genes. In addition, MOF binds the 3’ end of the X chromosomal genes. Other data from the lab have shown that the binding of MOF at the 3’ end of the genes correlates with the MSL binding. NSL colocalize with MOF at promoter regions and is absent from the 3’ end, emphasizing that MOF function with NSL is independent of MSL.

Subsequent comprehensive, genome-wide determination of binding sites of the NSL complex will be obtained by massively parallel sequencing of immunoaffinity purified DNA using antibodies against NSL components (ChIP-seq). As NSLs are

present on genes where MOF is bound, and that MOF associated with a considerable number of genes, it is likely that a strong correlation, indicated by overlapping regions, will occur between MOF and NSL targets on chromatin.

Preliminary experiments from our lab indicate that binding of MOF binding is reduced upon RNAi mediated depletion of MCSR2. This result is very interesting as it shows that the NSL complex contributse towards targeting MOF to target gene promoters. Our lab has previously shown (Kind et al. 2008) that H4K16 acetylation is downregulated upon the reduction of MOF. It would be therefore interesting to test whether the level of H4K16ac mark on target genes is affected in NSLs knock down cells. Reduced level of this mark could be an explanation for the downregulation of expression of MOF bound genes, as H4K16ac is linked to an increased rate of transcription. These experiments are currently underway.

4.4. NSL complex is involved in transcription regulation

The fact that NSLs are localized on the promoters of many genes prompted an evaluation of the transcriptional changes of genes subject to regulation by NSL. To address this question, total RNA from the salivary glands of fly lines specifically expressing RNAi targeting NSL components in salivary glands were isolated and subject to reverse transcription with subsequent real-time quantitative analysis by qPCR using gene-specific primers. The depletion of NSL components results in a strong decrease in transcription of NSL-bound genes in males and females, on autosomes as well as on the X chromosome. The expression levels of non-bound genes are not reduced, indicating that this decrease is a direct effect of the absence of the NSLs. Altogether, these data strongly suggest that the NSL complex functions generally as an activator of transcription. The broad polytene chromosome staining of NSLs implies that many chromatin regions are occupied by the NSL complex, which is likely to be involved in regulation of a wide spectrum of genes. A better understanding of the proportion of genes regulated by the NSLs and their nature will be derived by genome-wide profiling of binding sites of the NSL complex members in comparison with expression data, derived either from expression arrays or by massively parallel sequencing. Comparison of the data from wild type and MOF and

NSLs mutants will facilitate the understanding of the involvement of these complexes in the global regulation of gene expression in *Drosophila*.

4.5. Mechanism of targeting to promoters

An important question to answer is how the NSL complex is targeted to responsive promoters. One alternative is that the complex is directly binding the chromatin. Among the NSLs there are several proteins, which can be potentially involved in targeting the complex. One of the proteins, which may be involved in targeting the NSL complex to promoters, is MBDR2. MBDR2 has several interesting domains, among which are Tudor and methyl binding (MBD) domains, (Taipale et al. 2005). Tudor domains share similarity to chromodomains, which also bind methylated residues (Lachner et al. 2001). In yeast, Tudor domains have been shown to bind methylated H3K79 (Huyen et al. 2004). MBD domains bind to methylated DNA and are involved in transcriptional repression in mammals (Bird 2002). In *Drosophila*, however, DNA methylation happens much more seldom than in mammals and the function of it is not very clear (Lyko et al. 2000). MBDR2 has however the potential to recognize modified chromatin and target the NSL complex to it.

Another protein with a potential of binding to chromatin is WDS. It belongs to the WD family and contains seven WD40 repeats (Hollmann et al. 2002). It is known, that these repeats are involved in protein-protein interaction and are present in many chromatin-associated complexes (Cao et al. 2002). The mammalian ortholog of this protein, WDR5, binds specifically to dimethylated H3K4 and also is a constituent of H3K4-specific methyltrasnferase complexes (Wysocka et al. 2005). Methylated H3K4 is associated to transcription activation (Zhang and Reinberg 2001). Therefore, WDS can be potentially one of the proteins in the NSL complex that specifically recognizes H3K4 methylated chromatin of promoters and brings the complex to regulate transcription of the target genes by, for example, bringing MOF to induce local hyperacetylation and increase gene expression.

It was proposed, that the MSL complex recognizes some degenerate sequences at 3' end of genes, which leads to the binding of it to the chromatin. It would be interesting

to know, if there is a similar kind of a sequence that the NSL complex is recognizing at the promoter regions, with a subsequent binding to it.

Another alternative is that the NSL complex can be brought to promoters not only by direct binding of its members with chromatin, but through an interaction with components of transcription machinery or with regulatory proteins present on promoters. These interactions are most probably transient, and reflecting an inherent instability, no such proteins have been purified with the NSL complex. An interesting issue is a possible interdependency between MOF and NSLs in binding to chromatin. From the MSLs studies it is known that, although MOF binding is compromised on the body of genes upon MSL1 knock down, the association of MOF to promoters is not affected in these conditions (Kind et al. 2008). This suggests that binding of MOF to promoters is independent of MSL. It would be interesting to investigate whether MOF binding to promoters is also independent of the NSL complex, or MOF requires an assistance of the NSL complex to associate with promoter regions. For this, ChIP with NSLs antibodies using chromatin derived from NSLs and MOF mutant flies or from MOF RNAi mediated knock down cells will provide an answer.

4.6. The model of the MSL and NSL complexes function

Dosage compensated genes are upregulated through the action of histone acetylase activity of MOF within the MSL complex. Consequently, genes associated with MSL have local hyperacetylation of H4K16, which promotes gene expression. Components of the MSL complex bind all over the coding sequence of compensated genes, peaking at 3' end. MSL1 and MSL2 recognize 3' regions of target genes probably through degenerate sequence elements (Dahlsveen et al. 2006; Gilfillan et al. 2006; Kind and Akhtar 2007). This leads to the recruitment of MSL3, which stabilizes the MSL1/MSL2/chromatin interaction. Binding of MSL3 to H3K36me3 might be an event that brings MOF to the body if the gene (Larschan et al. 2007). However, in contrast to other components of the MSL complex, MOF is also present on the proximal promoter region of genes, on all chromosomes and in both males and females. This binding is independent of the MSL complex (Kind et al. 2008). The profile of H4K16ac correlates with the binding of MOF on autosomes and on the X chromosome. On the X chromosomal genes, H4K16ac peaks at the 3' end of genes

(Kind et al. 2008). The work presented in this thesis shows that the NSL complex coresides with MOF specifically at the promoters of many genes (Fig. 22). Since the depletion of the NSL proteins resulted in the down regulation of target genes, the NSL complex might likely be implicated in the gene activation of the tested genes. An open question is whether this gene regulation is direct effect of the MOF acetylation, or is the result of the function of the NSL proteins. To answer this question, the binding of MOF in the NLS proteins knock down background and vice versa has been investigated. However, the efficiency of the knock down was not enough to provide a clear answer, and this work is still in progress.

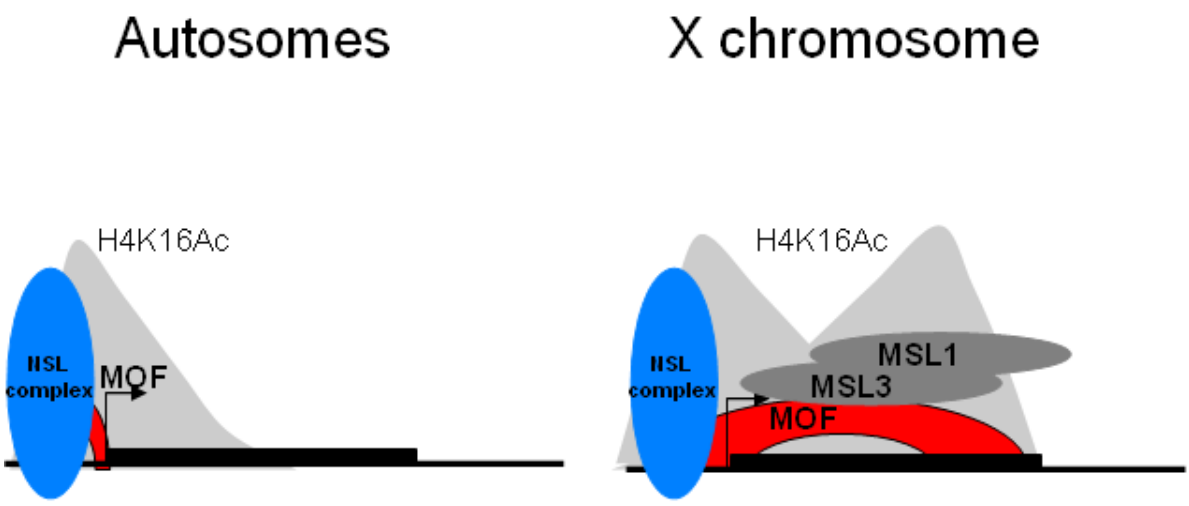


Figure 22. The differential distribution of the NSL and MSL complexes on genes. The NSL complex, in association with MOF, recognizes gene promoters and binds them. This induces H4K16 acetylation of proximal promoter regions, thereby increasing the transcriptional potential of the gene. On the male X chromosome, the NSL complex works in conjunction with the MSL complex, which then spreads itself and MOF over the body of the gene, thereby provoking general acetylation of male X chromatin. This correlates with an increase of transcription from the male X chromosome and results in dosage compensation.

4.7. Evolution conservation of the NSL proteins

Interestingly, the NSL proteins are conserved between two very evolutionary distant organisms: fly and human. First experiments showed that they might also reside in one complex together in human cells, as they segregate in the same fractions of a

nuclear extract after gradient centrifugation (Mendjan et al., 2006). This conservation indicates that NSLs are important for cell function. Although they have been identified in the dosage compensation studies, they are involved in general transcriptional regulation, in addition to X chromosomal genes dosage compensation. Correspondingly, in man, hMOF is linked with cancer. In normal human cells, about 60% of total histone H4 is monoacetylated, mostly at lysine 16, while this acetylation is frequently lost in cancer (Munks et al. 1991). hMOF is responsible for the specific acetylation of H4K16 and its depletion leads to global reduction of H4K16ac in HeLa cells (Taipale et al. 2005). In addition, hMOF depleted cells have an impaired DNA repair response following ionizing radiation (Taipale et al. 2005). hMOF is responsible not only for H4K16 acetylation, but is also able to acetylate the tumor suppressor protein p53, and this modification of p53 mediates the behavior of p53 in response to DNA damage (Sykes et al. 2006). It suggests that hMOF has a role in transcriptional regulation, cell proliferation, differentiation and the DNA repair response (Smith et al. 2005; Taipale et al. 2005). Knowledge that NSLs are associated with hMOF in mammalian cells brings new directions to the study of the NLS complex in mammalian system, as well as hMOF itself. It is possible that NSLs are also binding genes in mammalian genome and regulate their transcription together with hMOF.

In summary, the work described in this thesis defines a novel MOF containing NSL complex. This complex consists of a number of evolutionary conserved proteins. Its members colocalize on *Drosophila* polytene chromosomes. ChIP analysis reveals that the complex binds to promoters of MOF target genes, and this binding is functional, as it depletion of the complex members by RNAi mediated knock down leads to a decrease in the expression of X chromosomal and autosomal target genes. Thus, the NSL complex is a novel gene expression regulator, which, in *Drosophila*, provokes general transcriptional upregulation of a large number of genes.

5. MATERIALS AND METHODS

5.1. Biochemical methods

5.1.1. MCRS2 antibodies production

Polyclonal antibodies against MCRS2 protein were raised in rats and rabbits. For this purpose, N-terminally GST tagged MCRS2 (1-319 aa) protein was expressed in BL21 Rosetta (EMD Biosciences) using the pET41a vector system (EMD Biosciences). The protein formed inclusion bodies, which were collected and solubilized in 7 M guanidine HCl, 20 mM tris, pH 8.0, 10 mM DTT; unfolded in 7M urea, 20 mM sodium acetate pH 5.2, 0.2 M NaCl, 1 mM EDTA, 0.2 mM PMSF, 0.2 mM β-mercaptoethanol, 1 mM DTT; and refolded back by dialysis against urea free buffer: 50 mM Tris HCk pH 7.5, 150 mM NaCl, 1% Triton, 0.2 mM PMSF, 1 mM DTT. The protein was affinity purified on glutathione agarose. This material was then formulated with Titre-Max (Sigma) adjuvant and used to immunize 3 rats and 2 rabbits at three week intervals for a series of six injections.

Table 1. Usage of αMCRS2 antibodies in different applications.

Animal	Bleed	Western	IF	IP (µl)	ChIP
Rat1	final	1:1000	1:200*	4	-
Rat2	final	1:1000	-	4	-
Rat3	final	1:1000*	-	4*	-
Rabbit CDA	final	1:1000	1:200*	-	4*
Rabbit CCH	final	1:1000	1:200	-	4

IF – immunofluorescence;
IP – immunoprecipitation;
ChIP – chromatin immunoprecipitation;
* - best antibody to use for a given purpose.

5.1.2. Coimmunoprecipitation (CoIP)

For coimmunoprecipitation experiments (Fig. 5 in the Results part), nuclear extract (25 mg/ml) from wild-type *Drosophila* embryos was used. The protocol for the experiment is as follows.

1. Mix 100 µl extract with 600µl IP150 buffer for pre-cleaning with protein G beads (Sigma) for 30min at 4°C to remove unspecific resin-binding proteins.
2. Mix the supernatant with 4 µl of the respective antibody serum or preimmune serum for 1 hour, rotating at 4ºC.
3. Wash 4 times with 700 µl IP150-buffer each wash.
4. Resuspend the beads in 50 µl of 4xSDS-loading buffer, keep at 95°C for 5min.
5. Use 40µl of the supernatant on a SDS-PAGE for separation and subsequent western blot analysis with the corresponding antibodies.

<u>IP150 buffer</u>

HEMG150 (25 mM Hepes, 150 mM KCl, 0.2 mM EDTA, 12.5 mM $MgCl_2$, 10% (vol/vol) glycerol)
0.5% Tween-20
0.2 mg/ml BSA
0.2 mM PMSF
0.5 mM DTT
complete protease inhibitor (Roche).

5.1.3. Western blot

<u>SDS-PAGE gel, transfer</u>

1. Run the gel with constant current or voltage (usual running time under 50 mA constant is about 1.5-2.5 hrs).
2. Transfer for 1.5 hrs at 120 constant voltage with a ice block, stirring. Bigger proteins might take longer to transfer.
3. Immerse membrane in blocking buffer for one hour or overnight, 4ºC.

Reagents for gel

30% bis/acrylamide mix (i.e., 29.2% acrylamide and 0.8% N,N'-methylene-bis-acrylamide)

1.5 M Tris, pH 8.8

1.0 M Tris, pH 6.8

10% sodium dodecyl sulfate (SDS)

10% ammonium pesrsulfate (APS)

TEMED (N,N,N',N'-tetramethylethylenediamine)

5X running buffer:

30.28 g Tris (FW 121.1)

144.13 g glycine

10 g SDS (or 10 ml 10% SDS)

ddH2O to 2 L

4X Protein Loading Buffer

10 ml 1 M Tris pH 6.8;

20 ml 20% SDS;

20 ml glycerol;

0.2 g bromophenol blue;

1.234 g dithiothreitol (DTT; FW 154.2);

5ml beta-mercaptoethanol.

Store in 0.5 ml aliquots at -20°C for 6 months.

Incubation with primary and secondary antibodies

1. Incubate membrane with the primary antibody in blocking buffer for at least an hour at room temperature or overnight at 4ºC.
2. Wash three times 5 min each with washing buffer (PBS containing 0.1% Tween20).
3. Incubate the membrane with HRP-coupled secondary antibody at a dilution of 1:20.000 for 45-60 min at room temperature.
4. Wash three times 5 min each with washing buffer (PBS containing 0.1% Tween20).

5. Treat the membrane with ECL solutions (Amersham) and expose to X-ray film.

5.1.4. Coomassie staining protocol

The gel must be fixed prior to staining by a non-modifying, precipitation procedure such as the ethanol (or methanol)-acetic acid method. If the protein is not fixed in the gel as a separate step from the staining, the protein will be washed away and results will be compromised.

1. Soak the gel in the gel-fixing solution for 1hr. The purpose of this step is to washing the SDS-containing gel buffers out of the gel.
2. Cover the gel with the gel-washing solution, and continue to fix the proteins in the gel by incubating overnight at room temperature with gentle agitation. The gel should be covered during this process to avoid contamination and to prevent the evaporation of the solution.
3. Cover the gel with the Coomassie stain. Stain the gel at room temperature for 3 to 4 hr with gentle agitation.
4. Cover the gel with the destain solution and allow the gel to destain with gentle agitation. Change the destain solution several times. Continue the destaining until the protein bands are seen without background staining of the gel.
5. Equilibrate the gel in the storage solution for at least 1 hr. The gel should return to its original dimensions during this process.
6. Store the gel in the storage solution as needed. It might be convenient to carefully transfer the gel to a heat-sealable bag for longer-term storage.

Reagents

Gel-fixing solution: Add 500 ml of USP-grade 95% (v/v) ethanol to 300 ml of water. Add 100 ml of acetic acid and adjust the total volume to 1000 ml with water. The final concentrations are 50% (v/v) ethanol in water with 10% (v/v) acetic acid.

Gel-washing solution: Add 500 ml of methanol to 300 ml of water. Add 100ml of acetic acid and adjust the total volume to 1000 ml with water. The final concentrations are 50% (v/v) methanol in water with 10% (v/v) acetic acid.

Stain: Dissolve 0.4 g of Coomassie blue R350 in 200 ml of 40% (v/v) methanol in water with stirring as needed. Filter the solution to remove any insoluble material. Add 200 ml of 20% (v/v) acetic acid in water. The final concentration is 0.1% (w/v) Coomassie blue R350, 20% (v/v) methanol, and 10% (v/v) acetic acid.

Destain: Add 500 ml of HPLC- grade methanol to 300 ml of water. Add 100 ml of acetic acid and, after mixing, adjust the total volume to 1000 ml with water. The final concentrations are 50% (v/v) methanol in water with 10% (v/v) acetic acid.

Storage solution: Add 25 ml of acetic acid to 400 ml of water. After mixing, adjust the final volume to 500 ml with water. The final concentration of acetic acid is 5% (v/v).

5.1.5. Silver staining

Gel-separated proteins are most commonly detected and quantitated by dye binding, utilizing the property of some dyes to bind to proteins non-specifically, making the proteins optically detectable and quantifiable. One of the most commonly used procedures is silver staining.

Protocol of silver staining:

1. Fix gel with destaining solution (45 methanol : 5 acetic acid : 45 water) on a shaking table for 20-30 mins.
2. Rinse with water (20-60 mins, or overnight). Change water several times to remove acid completely and to avoid background.
3. Sensitize gel for 3 mins with 0.02% sodium thiosulfate (prepare fresh: 0.1 g sodium thiosulfate in 500 ml water).
4. Discard solution and rinse the gel with two changes of water, 1 min each.

5. Incubate gel in chilled 0.1% silvernitrate, (prepare fresh: 0.2 g silvernitrate in 200 ml water) in the cold room f0r 20-40 mins without shaking, covered with alufoil.
6. Discard/collect solution and rinse the gel with two changes of water, 1 min each.
7. Develop the gel with 0.04% formaldehyde, 2.5% natrium carbonate (prepare fresh: 12.5 g natrium carbonate, 150 µl 37% formaldehyde, 500 ml water). Replace developing solution when it turns yellow.
8. Quench developing solution when sufficient staining is obtained by discarding the solution and adding 1% acetic acid.
9. Store silver stained gel at cold, or dry.

5.1.6. Destaining silver gels

Silver particles bind to the proteins and thus inhibit their hydrolysis. All silver stained gel bands cut out for the purpose of in-gel digestion should be destained. It is done prior to performing the in-gel digest, but after excision of the bands from the gel.

1. Add 5 ml of 50% sodium thiosulfate and 7.5% potasium hexacyanoferrate per 150 ml water.
2. Use table shaker and incubate gel until no band is visible anymore.
3. Rinse well with water, of possible, overnight. Change water several times.

5.1.7 Flamingo staining

Since not all proteins can be visualized by silver staining, other staining reagents can be used. One of them is Flamingo staining (BioRad). It is a novel dye that was developed from a class of dyes that are minimally fluorescent at low pH in the absence of protein, but acquire strong fluorescence in the presence of denatured protein. Flamingo staining is fully compatible with peptide mass fingerprinting by MALDI-MS.

Protocol of the Flamingo staining:

1. Fix the gel o/n with 40% ethanol and 10% acetic acid
2. Incubated in 1:10 diluted Flamingo staining (BioRad) solution for 3 to 5 hours.
3. Incubate for 10min in a 0.1% Tween-20 solution.
4. Scan on a PharosFX scanner (BioRad).

5.1.8. SL-2 nuclear extracts

Nuclear extract from *Drosophila* SL-2 cells was used for TAP purification of MCRS2.

1. Harvest $0.3x10^9$ cells.
2. Pool cells, in falcon tubes and pellet by centrifugation 2000rpm, 10 min, at cold.
3. Rinse cells, in cold PBS (5-10 ml), respin 2000 rpm for 10 min.
4. Dounce 50 times on ice using the 1.5 ml pestle (type B).
5. Centrifuge 10 min at 4500 rpm (SS34 or eppendorf centrifuge).
6. Discard the supernatant.
7. Wash pellet in 5 ml of buffer B (in corex 15 ml tubes), spin at 4500 rpm for 10 min.
8. Resuspend in 1 ml of buffer B.
9. Create a sucrose gradient layer by carefully laying the resuspended nuclei on 1 ml of better B+0.8 M sucrose in corex tubes.
10. Spin in HB4 swing out rotor 10 min, 4000 rpm.
11. Resuspend the nuclei pellet in 800µl of buffer B 150 (i.e. 150 mM KCl).
12. Precipitate with 4 M ammonium sulphate (pH8.0), 82µl/ml of resuspension, rotate for 1 hour at cold.
13. Ultracentrifuge 1 hour using Ti55, 16200 rpm or SW40, 26000 rpm.
14. Take the supernatant and precipitate with 1 volume of ammonium sulphate (pH8.0) for 30 min, at cold.
15. Spin in corex tubes or eppendorf, 12000 rpm, 30 min.
16. Resuspend pellet in 40 µl of buffer C.

Buffer B: 15 mM Hepes, pH 7.6; x mM KCl; 5 mM $MgCl_2$, 0.1 mM EDTA; 0.5 mM

EGTA; 1 mM PMSF. B10=10 mM KCl, B150=150 mM KCl.

Buffer C: 25 mM Hepes, pH 7.6; 50 mM NaCl; 5 mM $MgCl_2$, 0.1 mM EDTA; 1 mM DTT; 1 mM PMSF, PI cocktail.

5.1.9. Nuclear *Drosophila* embryos extract

Nuclear extract from *Drosophila* embryo (0–12 hr collections) was used for coimmunoprecipitation experiments.

1. Suspend embryos in 1 ml NU-I buffer.
2. Homogenize in a 60 ml glass homogenizer with a motor-driven Teflon pestle (up to 30 g at the time). One slow stroke at 2000 rpm followed by 5 strokes at 1500 rpm.
3. Pass homogenate through a single layer of miracloth (Calbiochem) supported by a funnel over a GSA sentrifugation beaker. Use a new piece of miracloth for every 30 g embryos. Wash with 2 ml NU-I buffer per g embryos. Directly add to the beaker further 3 ml NU-I buffer per g embryos (total 6 ml of NU-I buffer per g embryo for the spin).
4. Spin in GSA rotor (HB4 for smaller preps) for 15 mins at 8 K.
5. Pour off cytoplasm, watch the nuclei pellet. Wipe side of the tubes with tissue to remove lipid.
6. Resuspend pellet in 3 ml of NU-I buffer per g embryos, leaving behind the much tighter yellow yolk pellet (use dounce with a loose pestle to fully suspend). This step is optional.
7. Spin again 15 mins at 8 K in a fresh beaker.
8. Pour off the supernatant, wipe sides again, if necessary.
9. Nuclei extraction: resuspend nuclei (again avoid yolk pellet) in 1 ml NU-II buffer per g embryo using a dounce with loose pestle. Measure the volume of the resuspended nuclei.
10. Place into ultracentrifuge tube(s) and add 1/10 volume of 4 M ammonium sulfate to tubes and mix vigorously. The solution will become very viscous.
12. Rotate the tubes in the coldroom for about 20 mins.
13. Spin in a cooled untracentrifuge for 1 hr: Ti60 or Ti70 rotor at 35 krpm.

14. Remove supernatant with a 10 ml pipet by plunging the pipet tip well below the upper lipid layer and removing steadily. Leave behind the bulk of the lipid which interferes with subsequent ammonium sulfate precipitation.
15. Measure volume of supernatant and place into a beaker on ice.
16. Add to the supernatant 0.3 g per ml of finely ground ammonium sulfate in small aliquots over a 5 min period while stirring. Leave stirring for further 10 mins.
17. Spin in precooled Sorvall at 15 krpm for 30 min (SS34 rotor or equivalent)
18. Pour off the supernatant, drain well and dry the sides of the tub with kimwipe wrapped around a spatula.
19. Resuspend pellet by adding 10 µl of HEMG40 per g embryo, then mixing into a paste with a pipet tip. Add 180 µl HEMG40 per g embryo and resuspend fully by pipetting up and down through a 5 ml pipet. Dialyse against 3x1 liter of HEMG40 until the conductivity is equal to HEMG 100 (HEMG + 100 mM KCl).
20. Spin out the precipitated protein for 5 min in an eppendorf centrifuge at 10 krpm in an HB4 (SS34) rotor.
21. Snapfreeze supernatant in liquid nitrogen and store at -80^0C. Expect 4-5 mg nuclear protein per g of dechorionated embryo. Protein concentrations in Bradford assay are usually between 15-20 mg/ml.

NU-I buffer: 15 mM Hepes-KOH, pH 7.6; 10 mM KCl; 5 mM $MgCl_2$; 0.1 mM EDTA, pH 8.0; 0.5 mM EGTA; pH8.0; 350 mM sucrose.

NU-II buffer: 15 mM Hepes-KOH, pH 7.6; 110 mM KCl; 5 mM $MgCl_2$; 0.1 mM EDTA, pH 8.0.

HEMGx: 25 mM Hepes-KOH, pH 7.6; x mM KCL; 12.5 mM $MgCl_2$; 0.1 mM EDTA, pH 8.0; 10% glycerol. HEMG 40 = HEMG + 40 mM KCl

5.1.10. TAP-tagging of the MCRS2 protein

The full-length open reading frame of MCRS2 protein was subcloned into the multiple-cloning site of the pBSactshort-N-TAP vector (a gift from Elisa Izzaurralde). This pBluescript-derived vector has an N-terminal TAP tag that can be fused to the

protein of interest. Expression of a tagged protein is driven from a shortened Actin5C-promoter with efficient termination of transcription conferred by a 3' BgH1 terminator sequence. The Actin5C-promoter drives low-level expression of TAP tagged MCRS2. This results in complex formation with TAP-tagged MCRS2 under conditions similar to those of endogenous MCRS2 expression, to preclude forcing the formation of aberrant complexes through over-expressing MCRS2. The MCRS2 coding sequence was subcloned from pFastBac-FlagCG1135 by digestion with *Eco*RI (5') and *Spe*I (3'). Both ends were blunted with Klenow polymerase (NEB). The acceptor vector, pBSactshort-N-TAP, was prepared by restriction digest with *BamH*I, followed by blunting. Analytical digests were performed with *Eco*RI and *BamH*I to confirm the presence and orientation of the insert. The resulting construct was named pBSactshort-N-TAP-MCRS2.

5.1.11. Generation of stable *Drosophila* SL2 cell line

A stable *Drosophila* TAP-MCRS2 producing SL2 cell line was established by cotransfecting the pBSactshort-N-TAP-MCRS2 expression plasmid and the pUC-NEO resistance vector. Transfection was facilitated by Effectene (Qiagen), used according to the manufacturer's instructions. Transfection with only the expression vector was performed in parallel. Following transfection, cells were incubated for 24 hours, after which the medium was exchanged. Selection with geneticin (G418, Invitrogen) was initiated 48 hours after transfection, with a range of antibiotic concentrations, between 0.8 and 1.2 mg/ml of geneticin, used. Selection was monitored by the complete death of mock transfected cells and colony formation of stably expressing cells in cells transfected with the selection vector. Heterogeneous populations of transfected cells were then used to derive cell lines.

5.1.12. Tandem affinity purification (TAP)

The TAP purification protocol (Rigaut, 1999) was adapted for Drosophila embryo nuclear extracts (Sascha Mendjan).

1. Spin down extracts immediately after thawing at max speed for 15min.

2. Dilute nuclear extract is in IgGBB150 to about 5mg/ml protein concentration, and spun down at rpm max for 15min.

3. IgG beads (Roche) should be tested by boiling in 1xSDS LB for IgG release. If they release IgG (fat coomasie band at 50kDa) crosslink the beads with dimethyl-pimelidate. Beads are equilibrated in IgGBB150 before binding.

4. Diluted extract is bound to IgG beads at 4°C for 1h-1h30min max.

5. Save supernatant (for binding control) and wash beads 3x with IgGBB150 and 3x with IgGBB200 (identical to IgGBB150 but with 200 mM KCl). The last 2 washes should be 5-10' each and at room temperature RT.

6. Resuspend beads in TEV cleavage buffer CB150. Wash once in CB150 without TEV at RT, before you add the CB150+TEV.

7. Cleave at 18°C for 2h rotating slowly in an appropriate tube (2/3-3/4 full).

8. Take off the supernatant, spin it down at max speed for 5' 4°C, and save the beads (freeze). Add 3µl of 1M $CaCl_2$ per 1ml of cleavage supernatant.

9. Dilute cleavage supernatant in calmodulin binding buffer CalBB150 in a 1:3 ratio. Equilibrate calmodulin beads in CalBB150 before binding. Bind for 2h at 4°C.

10. Spin down beads and save the supernatant (binding control). Wash beads (each wash 5-10min) 2xCalBB150 at 4°C, 2xCal150 at RT, and 2xCalBB200 at RT. Final wash is in CalBB150 (with Tris pH7.6 instead of Hepes: this final change to the Tris buffer system is because of the subsequent PAGE run that is also Tris based).

11. To avoid presence of non-specifically bound proteins, as well as to result in purifying a complex for further biochemical analysis, elute proteins with CalEl150 for 15'-30' at 4°C rotating/shaking. If the non-specific binding is low, proteins can be directly eluted by boiling in 1xSDS (with beta-mercaptoethanol, no DTT).

IgGBB150 buffer: (25mM Hepes pH7.6, 150mM KCl, 5mM MgCl2, 0.5 mM EDTA, 20% glycerol and 0.5mM DTT, 0.2%Tween20, 0.4mM PMSF, Complete protease inhibitor (Roche).

IgGEl150 buffer: (20mM Tris pH7.6, 150mM KCl, 5mM MgCl2, 0.5 mM EDTA, 20% glycerol and 0.4mM PMSF, 200ng/ml FLAG peptide, 1/100 elution volume RNAsin (Promega) if RNA is co-purified.

CB150 buffer: (20mM Hepes pH 7.6, 150mM KCl, 0.5mM EDTA and 0.5mM DTT, 0.1% Tween20, 0.4mM PMSF, 10microg/ml TEV and 1/100 volume of RNasin (Promega) if RNA should be co-purified.

CalBB150 buffer: (20mM Hepes/10mMTris pH7.6, 150mM KCl, 2mM Mg-Acetate, 1mM Imidazol, 3mM CaCl2, 20% glycerol and 10mM beta-mercaptoethanol, 0.2%Tween20, 0.4mM PMSF, Complete protease inhibitor)

CalEl150 buffer: (10mM Tris pH7.6, 150mM KCl, 2mM Mg-Acetate, 1mM Imidazol, 3mM EGTA, 20% glycerol and 10mM beta-mercaptoethanol, 0.4mM PMSF, RNasin).

5.1.13. Identification of purified proteins by mass spectrometry

Both excision of specific to TAP-tagged MCRS2 purification bands (performed by Sven Fraterman, EMBL-Heidelberg) and analysis of total complex elutions have been done (Adrian Cohen, NCLMS, Netherlands).

5.1.14. Chromatin from *Drosophila* SL-2 cells

1. Fix 40 ml almost dense cell culture ($4x10^6$ cells / ml) with 4 ml of fixation mix (7.1 ml Paro fix solution (50mM HEPES, 1mM EDTA, 0.5mM EGTA, 100mM NaCl) + 2.9 ml FA (37% stock)).
2. Incubate 8 mins at RT.
3. Stop with 2 ml 2,5M glycine, 10 min on ice, spin 5 mins cold.
4. Wash pellet in 40 ml Paro Pinse 1 (10mM tris pH8, 10mM EDTA, 0,5mM EGTA, 0,25% triton X 100), incubate 5 mins on ice, spin.

5. Wash pellet in 40 ml Paro Pinse 2 (10mM tris pH8, 1mM EDTA, 0,5mMEGTA, 0,2mM NaCla), incubate 5 mins on ice, spin.
6. Resuspend in Ripa buffer in the appropriate for sonication volume (400 µl).
7. Sonicate 3 times 15sec.
8. Spin at high speed for 10 mins, cold.
9. Use supernatant (can store at -80). Take 25-100 ug/ul of the chromatin for an IP.

5.1.15. Chromatin from *Drosophila* third instar larvae salivary glands

Chromatin from *Drosophila* embryos was prepared according the protocol developed by Orlando et al (Orlando et al, 1997).

1. Dissect male III instar larvae salivary glands (10 pairs per ChIP) 2. Fix for 15 min at room temperature in 1mL of fixing solution (50mM HEPES pH 7.6, 100mM NaCl, 0.1mM EDTA pH8, 0.5mM EGTA pH8, 2% formaldehyde), on a wheel.
3. Centrifuged at 2000rpm for 1min.
4. Wash once in PBS-0,01%Triton X100-0,125M glycine.
5. Wash for 10 min in 1mL of buffer A (0,25% Triton X100, 10mM EDTA pH8, 0.5mM EGTA pH8, 10mM Tris pH8).
6. Wash for 10 minutes in Buffer B (200mM NaCl, 10mM Tris pH8, 10mM EDTA pH8, 0,5mM EGTA pH8). Can freeze in N2 and keep for weeks.
7. Resuspend the glands in 500 ml of sonication buffer (10mM Tris pH8, 1mM EDTA pH8, 0.5mM EGTA pH8). Transfer in the specific tubes for sonication.
8. Sonicate 3 times 10s using a Branson Sonifier 250, power 2, with a microtip.
9. Sonicate 8 min (8 times pulse 30s, paused for 30s, high) using a Bioruptor (Cosmo Bio).
10. Adjust the samples to 0.5% sarcosyl (add 50µl for 1mL of extract).
11. Incubate on a wheel for 10 min at room temperature.
12. Centrifuged at 14000rpm for 10 min.
13. Dialyze at 4°C ON against 5% glycerol, 10mM Tris pH8, 1mM EDTA pH8, 0.5mM EGTA pH8.
14. Perform chromatin immunoprecipitation using 500µl of chromatin, according to Orlando et al, 1997.

5.1.16. Chromatin immunoprecipitaion, ChIP

(adapted from J. Muller lab)

1. Thaw an aliquot of chromatin and use the appropriate amount for the ChIP
2. Adjust the volume to 500µl with ice-cold dialysis buffer (4% glycerol, 10mM TrisHCL pH 8, 1mM EDTA, 0.5mM EGTA)
3. Adjust to RIPA conditions by addition of TritonX100, sodium deoxycholate, SDS and NaCl.
4. Add 40µl of the 50% (v/v) ProteinA Sepharose (PAS) suspension (100mg Protein A Sepharose CL4B, Sigma, equilibrated in 1ml RIPA buffer for 30-60min; this will swell the beads with RIPA up to 500 µl volume. Spin down 30 sec and take up to 1 ml new RIPA buffer). RIPA Buffer: 140mM NaCl / 10mM Tris-HCl pH8,0 / 1mM EDTA / 1% TritonX100 / 0,1% SDS / 0,1% sodium deoxycholate, 1mM PMSF (on ice); add PMSF immediately before use).
5. Incubate the chromatin with the PAS for 1h at 4°C with gentle mixing, then spin down for 30sec at max speed. This acts as a preclearing step to reduce non-specific binding to protein A sepharose.
6. Remove the chromatin to a new tube and add 2-5µg of the appropriate antibody and control (preimmune serum).
7. Incubate overnight at 4°C with gentle mixing. Purify immunocomplexes by adding 40µl 50% PAS suspension (100mg Protein A Sepharose CL4B, Sigma, equilibrated in 1ml RIPA buffer for 30-60min. Spin down and take up in new 1ml RIPA buffer) and incubate for 3h at 4°C with gentle mixing. Use filtered tips during the ChIP.
8. Wash complexes 5 times with 1ml RIPA for 10 min each, once in LiCl buffer (250mM LiCl / 10mM Tris-HCl pH 8,0 / 1mM EDTA / 0,5% NP-40 / 0,5% sodium deoxycholate; prepare it well in advance so that NP-40 has time to get dissolved properly) and twice in TE (10mM Tris-HCl pH 8,0 / 1mM EDTA). Carry out all steps at 4°C using 1ml wash buffer and always spin at max speed for 30 sec to pellet PAS before removing the supernatant.
9. Add DNase-free RNase A (in approx. 50 µl TE buffer) to the PAS complexes up to 50µg RNase A/ml, and incubate for 30min at 37°C.
10. Adjust the samples to 0,5% SDS, 0,5mg/ml Proteinase K (premix SDS and Proteinase K in at least 40 µl buffer TE) and incubate overnight at 37°C, followed by 6h at 65°C to reverse the cross-link.

11. Phenol/chloroform extract the sample with adding 1 Vol phenol, 20 sec vortex, 1 Vol chloroform, 20 sec vortex, 2 min max speed centrifugation. Take the upper phase and add 2 Vol chloroform, 20sec vortex, 2min centrifuge. OR, if you use Phase Lock Gel Heavy® tubes, spin down gel in tubes (maximum speed 30 seconds), put liquid from Proteinase K into tube, add one volume V of phenol (take it from the lower phase), stir vigorously (around two minutes), add one volume V of chloroform, and stir vigorously again. Centrifuge at 16000g (max speed in microfuge) 5 min. Add then two volumes V of chloroform, stir vigorously, and again spin down at top speed five min. Take phase that is on top of gel.
Precipitate the upper phase by adding 1µl 20mg/ml glycogen (as carrier), 1/10 volume (V', the new volume you got after extraction) 3M NaOAc pH 5.2, add 2,5 V' ethanol. Put at –80°C for 30min before centrifuging at 4°C for 20min at full speed. Wash pellet in 1ml 70% ethanol, turn tube, centrifuge at 4ºC 3 min max speed, discard supernatant, air-dry and resuspend in 500µl H2O. Store at –20°C.

As an option to step 11, one can use MinElute columns (Quagen) for to purify DNA after RNase A, proteinase K treatments and reverse crosslinking.

Table 5. Primers used for qPCR in chromatin immunoprecipitation.

Name	Forward primer (5'-3')	Reverse primer (5'-3')	Position relative to a TSS
CG4406-5'	ACAGCTGGCGAGGATCAG	TCGATACTCGAGGCGTTG	+60
CG4406-mid	CCAACTCCTGGCTGGTTATC	GGCAGCAATGTGCTCATCTA	+570
CG4406-3'	TTGAAGGGCTTTTTGGTCAC	TTGAAGGGCTTTTTGGTCAC	+1280
CG6506-5'	AGGGCCCGATAAGTAAACAA	GCCCCAGTGCTCTGTTTG	+60
CG6506-mid	CAGCTGGTCCCACTGGAG	ATTCCTGGCCAGCACCTT	+850
CG6506-3'	GCAATGGAAATGGCAATG	TGAAGTTATCCCCGCAATTT	+1270
Rb-5'	AAAAATCATCAGCACGGAAA	ATTGCTGGCCGAGTTCTG	+300
Rb-mid	TGCCCGCCAAGTATTTCT	CGCTGGCATGTTCAGGTA	+2340
Rb-3'	CTCCAAAAGCCTCGTGCT	CCATGGATCCAATGACCA	+3840
dSPT6-5'	CGCTCACAAACTCTTCGTTT	ACACCTACCTCCGATTCCTC	+110
dSPT6-mid	CGAGGCGATAGTTGTACCAG	CATAGGGACTGCTGTTGGAC	+2663
dSPT6-3'	TACAATGTCACTGGGACGTG	CGAGGACATACCCCGATTAT	+6454
Sec5-5'	GCCAAGATTTCACCACTGAC	ATGCGGAAAAACTGATCAAA	-60
Sec5-mid	ACTCCCATTGGCGATAAACT	TGGTGTGCTGATCAAATGTC	+1350
Sec5-3'	TGAGACTGCCAAGTGAGTGA	CAGCGCTTCCATGAAGTAGT	+2780
CG9536-5'	AGACCACCCGGTTCCAGT	CACCGATCGCTTCTCCTG	-30
CG9536-mid	CGGAGAGCTTCACGTTCG	CCCGCAAACAGCAATTGTA	+2780
CG9536-3'	CCAGCTGCCCATCACAAC	CACCTTGACCCGGAACAT	+2040
OdsH-5'	CAGTGTCAGCAAAAGCATTG	GATGAACCATGGGGATGTT	+10
OdsH-mid	TCTGGGGCAGAATGATTGTA	CGCTATACGACCCTCCATTA	+20770
OdsH-3'	GTTGAACCGGAGTACGTGA	GAGGGTCTTATTCTGCATCG	+22750
Gprk2-5'	CTTGTTTTGCGAGCCTTTTC	CAGAACACACACACGCACAC	+220
Gprk2-mid	GTCGCTTCTTGGATGTCGAG	CTGCGAGTTGTTGCTGTTGT	+48540
Gprk2-3'	TTGCCCATTGGGTATGCT	TTTGCAAAAGCGCACTCC	+54540
roX1-prom	GTGTATTTTGCAATTGGA	CGCATTCATGCAGTTCCC	+48540

5.2. Cytological methods

5.2.1. Immunofluorescence on SL-2 cells

SL-2 cells are grown on cover slips at a density of 7 x10^6 cells ml-1. After 1 wash in PBS, cells are fixed in 3.7% Formaldehyde in PBS–10 min RT and blocked in solution 1 (5% BSA, 0.1%Tween, 0.1% Triton in 4X SSC) for 1 hour at RT or o/n at 4°. Cells are incubated with primary antibody diluted (anti-MSL-1, anti-MSL-2, anti-MSL-3 and anti-MOF: 1/500; anti-MLE: 1/250) in solution 1 for 1 hour at RT. After washing 3X 10 min in 0.1% Tween, 0.1% Triton, 4X SSC, cells were incubated in secondary antibody and Hoechst, diluted in solution 1, for 1 hour RT. After 3X 10 min washes in 0.1% Tween, 0.1% Triton, 4X SSC coverslips were mounted on a slide with a drop of Fluoromont-G.

5.2.2. Immunostaining of polytene chromosome.

Preparation of 3rd instar larvae

1. Add a large drop of live baker's yeast on top of the dried medium.
2. Let the flies lay eggs just to the point where larvae will hatch under uncrowded conditions (<100 larvae/bottle).
3. Grow larvae at 18^0C.
4. For salivary gland preparations use 3rd instar larvae that are still crawling and have not started to pupate, yet.

Chromosome squashes

1. Dissect two pairs of salivary glands in PBS.
2. Fix glands (3.7% Paraformaldehyde in H20 for 10 min) in poly-lysine treated slide.
3. Cover glands with a SIGMA-cote treated cover slip.
4. Tap the coverslip with a pencil until cells are broken up. Hold the coverslip and spread extensively the chromosomes. Remove excess fixative by pressing slides (coverslip down) onto blotting paper.
5. After freezing slides in liquid nitrogen flick off coverslip with a razorblade.

6. Wash slides two times for 15 min. in PBS slowly shaking the rack.

7. Proceed with the immunostaining or keep the slides (up to one week) in 100% Methanol or Ethanol (for EGFP staining).

Immunostaining

1. Stored slides are washed 2-x 15 min. in PBS. Block for 1 hour in blocking solution at room temperature.

2. Add 20 µl to each slide of affinity purified primary antibodies (i.e. rabbit polyclonal antibodies; dilutions 1:50 to 1:500 in blocking solution need to be adjusted for each individual primary antibody). Cover with coverslip and incubate for 1h at room temperature in a humid chamber.

3. Rinse in PBS

4. Wash 15 min in PBS, 300mM NaCl, 0.2% NP40, 0.2% Tween20-80;

15 min in PBS, 400mM NaCl, 0.2% NP40, 0.2% Tween20-80;

(If background problems persist, NaCl conc. can be raised to 500mM)

5. Shake rack thoroughly during washing procedure.

6. Rinse in PBS

7. Add 20 µl diluted secondary antibody (fluorescent labeled like Cy3- Anti-Rabbit IgG (Fc) (Dianova), or Anti-Rabbit IgG (Fc) HRP Conjugate, Promega Kat. Nr.: W4011, 1:100 dilution) + Hoechst (1:3000 dilution) in blocking solution. Cover with coverslip and incubate for 40 min. at rt. in humid chamber.

8. Rinse in PBS.

9. Wash 15 min in PBS, 300mM NaCl, 0.2% NP40, 0.2% Tween20-80;

3-15 min in PBS, 400mM NaCl, 0.2% NP40, 0.2% Tween20-80;

10. Shake rack thoroughly during washing procedure.

Rinse in PBS.

11. Mount the chromosomes in 10µl Fluoromont G.

5.2.3. Confocal microscopy

For cells and polytene chromosomes, images were captured with an AxioCamHR CCD camera on a Leica SP2 FCS spectral filterless confocal microscope (Leica

Microsystems) using 63x PlanApochromat NA 1.32 oil immersion objective and the Leica Confocal Software V2.61.

5.2.4 Double-stranded RNA interference

RNA interference was performed essentially as described before (Clemens et al. 2000) with the few modifications. S2 cells were grown at 25°C in Schneider's Drosophila medium (Gibco) supplemented with 10% foetal bovine serum and a mix of 100 U/ml penicillin and 100 mg/ml streptomycin (Invitrogen). Gene-specific dsRNAs were amplified by PCR from corresponding cDNAs using T7-tailed oligonucleotides. Products were about 300 (for MCRS2, MBDR2 and EGFP) or 600 (for MOF, MSL1) nucleotides long. The resulting PCR products were transcribed using the T7 RiboMAX Express Large Scale RNA Production System (Promega). A total of $6x10^6$ S2 cells were incubated with 45µg dsRNA per $1x10^6$ cells, additional 45µg dsRNA were added on day 2 and harvested after 5 days.

Table 6. Primers used for RNAi mediated knock down.

Name	Forward primer (5'-3')	Reverse primer (5'-3')
T7-MOF	TTAATACGACTCACTATAGGGAGA ATGTCTGAAGCGGAGCTGGAACAG	TTAATACGACTCACTATAGGGAGACGA AGTCGTCAATGTTGGAACCAC TG
T7-MCRS2	TTAATACGACTCACTATAGGGAGAGCG TCTAGAGACCAGGAAGGTGAAGCGCAGA	TTAATACGACTCACTATAGGGAGACGC GAATTCCCTCCGAGTTCGACAACCAGACA
T7-MBDR2	TTAATACGACTCACTATAGGGAGAGGCC GTCCTCTTCGCCATACTCGCAGATGCA	TTAATACGACTCACTATAGGGAGA GCTCTCAGCGTCGTCCCATTTTGTCAGAT
T7-MSL1	TTAATACGACTCACTATAGGGAGA ATGTCTGAAGCGGAGCTGGAACAG	TTAATACGACTCACTATAGGGAGA CGAAGTCGTCAATGTTGGAACCACTGCC
T7-EGFP	TAATACGACTCACTATAGGGAGG ATGGTGAGCAAGG	TAATACGACTCACTATAGG GAGGATCGCGCTTCTCG

REFERENCES

Adler, D. A., E. I. Rugarli, et al. (1997). "Evidence of evolutionary up-regulation of the single active X chromosome in mammals based on Clc4 expression levels in Mus spretus and Mus musculus." Proc Natl Acad Sci U S A **94**(17): 9244-8.

Alekseyenko, A. A., E. Larschan, et al. (2006). "High-resolution ChIP-chip analysis reveals that the Drosophila MSL complex selectively identifies active genes on the male X chromosome." Genes Dev **20**(7): 848-57.

Aoyagi, S., G. Narlikar, et al. (2002). "Nucleosome remodeling by the human SWI/SNF complex requires transient global disruption of histone-DNA interactions." Mol Cell Biol **22**(11): 3653-62.

Avner, P. and E. Heard (2001). "X-chromosome inactivation: counting, choice and initiation." Nat Rev Genet **2**(1): 59-67.

Avvakumov, N. and J. Cote (2007). "The MYST family of histone acetyltransferases and their intimate links to cancer." Oncogene **26**(37): 5395-407.

Baek, S. H. and M. G. Rosenfeld (2004). "Nuclear receptor coregulators: their modification codes and regulatory mechanism by translocation." Biochem Biophys Res Commun **319**(3): 707-14.

Baker, B. S., M. Gorman, et al. (1994). "Dosage compensation in Drosophila." Annu Rev Genet **28**: 491-521.

Bannister, A. J., R. Schneider, et al. (2002). "Histone methylation: dynamic or static?" Cell **109**(7): 801-6.

Bao, Y. and X. Shen (2007). "SnapShot: chromatin remodeling complexes." Cell **129**(3): 632.

Beckmann, K., M. Grskovic, et al. (2005). "A dual inhibitory mechanism restricts msl-2 mRNA translation for dosage compensation in Drosophila." Cell **122**(4): 529-40.

Bird, A. (2002). "DNA methylation patterns and epigenetic memory." Genes Dev **16**(1): 6-21.

Bisoni, L., L. Batlle-Morera, et al. (2005). "Female-specific hyperacetylation of histone H4 in the chicken Z chromosome." Chromosome Res **13**(2): 205-14.

Bochar, D. A., L. Wang, et al. (2000). "BRCA1 is associated with a human SWI/SNF-related complex: linking chromatin remodeling to breast cancer." Cell **102**(2): 257-65.

Borrow, J., V. P. Stanton, Jr., et al. (1996). "The translocation t(8;16)(p11;p13) of acute myeloid leukaemia fuses a putative acetyltransferase to the CREB-binding protein." Nat Genet **14**(1): 33-41.

Brownell, J. E., J. Zhou, et al. (1996). "Tetrahymena histone acetyltransferase A: a homolog to yeast Gcn5p linking histone acetylation to gene activation." Cell **84**(6): 843-51.

Buscaino, A., T. Kocher, et al. (2003). "MOF-regulated acetylation of MSL-3 in the Drosophila dosage compensation complex." Mol Cell **11**(5): 1265-77.

Cao, R., L. Wang, et al. (2002). "Role of histone H3 lysine 27 methylation in Polycomb-group silencing." Science **298**(5595): 1039-43.

Cao, R. and Y. Zhang (2004). "The functions of E(Z)/EZH2-mediated methylation of lysine 27 in histone H3." Curr Opin Genet Dev **14**(2): 155-64.

Chai, B., J. Huang, et al. (2005). "Distinct roles for the RSC and Swi/Snf ATP-dependent chromatin remodelers in DNA double-strand break repair." Genes Dev **19**(14): 1656-61.

Chang, C. R., C. S. Wu, et al. (2005). "Targeting of cohesin by transcriptionally silent chromatin." Genes Dev **19**(24): 3031-42.

Chu, D. S., H. E. Dawes, et al. (2002). "A molecular link between gene-specific and chromosome-wide transcriptional repression." Genes Dev **16**(7): 796-805.

Chuang, P. T., D. G. Albertson, et al. (1994). "DPY-27:a chromosome condensation protein homolog that regulates C. elegans dosage compensation through association with the X chromosome." Cell **79**(3): 459-74.

Chuang, P. T., J. D. Lieb, et al. (1996). "Sex-specific assembly of a dosage compensation complex on the nematode X chromosome." Science **274**(5293): 1736-9.

Clemens, J. C., C. A. Worby, et al. (2000). "Use of double-stranded RNA interference in Drosophila cell lines to dissect signal transduction pathways." Proc Natl Acad Sci U S A **97**(12): 6499-503.

Cline, T. W. and B. J. Meyer (1996). "Vive la difference: males vs females in flies vs worms." Annu Rev Genet **30**: 637-702.

Copps, K., R. Richman, et al. (1998). "Complex formation by the Drosophila MSL proteins: role of the MSL2 RING finger in protein complex assembly." EMBO J **17**(18): 5409-17.

Corona, D. F., A. Eberharter, et al. (2000). "Two histone fold proteins, CHRAC-14 and CHRAC-16, are developmentally regulated subunits of chromatin accessibility complex (CHRAC)." EMBO J **19**(12): 3049-59.

Corona, D. F. and J. W. Tamkun (2004). "Multiple roles for ISWI in transcription, chromosome organization and DNA replication." Biochim Biophys Acta **1677**(1-3): 113-9.

Crews, D. (2003). "Sex determination: where environment and genetics meet." Evol Dev **5**(1): 50-5.

Czermin, B., R. Melfi, et al. (2002). "Drosophila enhancer of Zeste/ESC complexes have a histone H3 methyltransferase activity that marks chromosomal Polycomb sites." Cell **111**(2): 185-96.

Davis, T. L. and B. J. Meyer (1997). "SDC-3 coordinates the assembly of a dosage compensation complex on the nematode X chromosome." Development **124**(5): 1019-31.

Dawes, H. E., D. S. Berlin, et al. (1999). "Dosage compensation proteins targeted to X chromosomes by a determinant of hermaphrodite fate." Science **284**(5421): 1800-4.

Decristofaro, M. F., B. L. Betz, et al. (2001). "Characterization of SWI/SNF protein expression in human breast cancer cell lines and other malignancies." J Cell Physiol **186**(1): 136-45.

Delmas, V., D. G. Stokes, et al. (1993). "A mammalian DNA-binding protein that contains a chromodomain and an SNF2/SWI2-like helicase domain." Proc Natl Acad Sci U S A **90**(6): 2414-8.

Deng, X. and V. H. Meller (2006). "roX RNAs are required for increased expression of X-linked genes in Drosophila melanogaster males." Genetics **174**(4): 1859-66.

Dou, Y. and M. A. Gorovsky (2000). "Phosphorylation of linker histone H1 regulates gene expression in vivo by creating a charge patch." Mol Cell **6**(2): 225-31.

Dou, Y. and M. A. Gorovsky (2002). "Regulation of transcription by H1 phosphorylation in Tetrahymena is position independent and requires clustered sites." Proc Natl Acad Sci U S A **99**(9): 6142-6.

Dror, V. and F. Winston (2004). "The Swi/Snf chromatin remodeling complex is required for ribosomal DNA and telomeric silencing in Saccharomyces cerevisiae." Mol Cell Biol **24**(18): 8227-35.

Du, J., I. Nasir, et al. (1998). "Sth1p, a Saccharomyces cerevisiae Snf2p/Swi2p homolog, is an essential ATPase in RSC and differs from Snf/Swi in its interactions with histones and chromatin-associated proteins." Genetics **150**(3): 987-1005.

Durocher, D., S. J. Smerdon, et al. (2000). "The FHA domain in DNA repair and checkpoint signaling." Cold Spring Harb Symp Quant Biol **65**: 423-31.

Eggert, H., A. Gortchakov, et al. (2004). "Identification of the Drosophila interband-specific protein Z4 as a DNA-binding zinc-finger protein determining chromosomal structure." J Cell Sci **117**(Pt 18): 4253-64.

Ellegren, H., L. Hultin-Rosenberg, et al. (2007). "Faced with inequality: chicken do not have a general dosage compensation of sex-linked genes." BMC Biol **5**: 40.

Ercan, S., P. G. Giresi, et al. (2007). "X chromosome repression by localization of the C. elegans dosage compensation machinery to sites of transcription initiation." Nat Genet **39**(3): 403-8.

Fan, Y., T. Nikitina, et al. (2005). "Histone H1 depletion in mammals alters global chromatin structure but causes specific changes in gene regulation." Cell **123**(7): 1199-212.

Felsenfeld, G. and M. Groudine (2003). "Controlling the double helix." Nature **421**(6921): 448-53.

Fischle, W., B. S. Tseng, et al. (2005). "Regulation of HP1-chromatin binding by histone H3 methylation and phosphorylation." Nature **438**(7071): 1116-22.

Fischle, W., Y. Wang, et al. (2003). "Molecular basis for the discrimination of repressive methyl-lysine marks in histone H3 by Polycomb and HP1 chromodomains." Genes Dev **17**(15): 1870-81.

Franke, A. and B. S. Baker (1999). "The rox1 and rox2 RNAs are essential components of the compensasome, which mediates dosage compensation in Drosophila." Mol Cell **4**(1): 117-22.

Fyodorov, D. V. and J. T. Kadonaga (2002). "Dynamics of ATP-dependent chromatin assembly by ACF." Nature **418**(6900): 897-900.

Gangaraju, V. K. and B. Bartholomew (2007). "Mechanisms of ATP dependent chromatin remodeling." Mutat Res **618**(1-2): 3-17.

Gardner, R. G., Z. W. Nelson, et al. (2005). "Ubp10/Dot4p regulates the persistence of ubiquitinated histone H2B: distinct roles in telomeric silencing and general chromatin." Mol Cell Biol **25**(14): 6123-39.

Gebauer, F., M. Grskovic, et al. (2003). "Drosophila sex-lethal inhibits the stable association of the 40S ribosomal subunit with msl-2 mRNA." Mol Cell **11**(5): 1397-404.

Georgel, P. T., T. Tsukiyama, et al. (1997). "Role of histone tails in nucleosome remodeling by Drosophila NURF." EMBO J **16**(15): 4717-26.

Ghosh, S., R. N. Chatterjee, et al. (1989). "The LSP1-alpha gene of Drosophila melanogaster exhibits dosage compensation when it is relocated to a different site on the X chromosome." EMBO J **8**(4): 1191-6.

Gilfillan, G. D., T. Straub, et al. (2006). "Chromosome-wide gene-specific targeting of the Drosophila dosage compensation complex." Genes Dev **20**(7): 858-70.
Gortchakov, A. A., H. Eggert, et al. (2005). "Chriz, a chromodomain protein specific for the interbands of Drosophila melanogaster polytene chromosomes." Chromosoma **114**(1): 54-66.
Grant, P. A., D. E. Sterner, et al. (1998). "The SAGA unfolds: convergence of transcription regulators in chromatin-modifying complexes." Trends Cell Biol **8**(5): 193-7.
Graves, J. A. (2006). "Sex chromosome specialization and degeneration in mammals." Cell **124**(5): 901-14.
Grskovic, M., M. W. Hentze, et al. (2003). "A co-repressor assembly nucleated by Sex-lethal in the 3'UTR mediates translational control of Drosophila msl-2 mRNA." EMBO J **22**(20): 5571-81.
Grune, T., J. Brzeski, et al. (2003). "Crystal structure and functional analysis of a nucleosome recognition module of the remodeling factor ISWI." Mol Cell **12**(2): 449-60.
Hagstrom, K. A., V. F. Holmes, et al. (2002). "C. elegans condensin promotes mitotic chromosome architecture, centromere organization, and sister chromatid segregation during mitosis and meiosis." Genes Dev **16**(6): 729-42.
Hassan, A. H., P. Prochasson, et al. (2002). "Function and selectivity of bromodomains in anchoring chromatin-modifying complexes to promoter nucleosomes." Cell **111**(3): 369-79.
Heard, E. (2004). "Recent advances in X-chromosome inactivation." Curr Opin Cell Biol **16**(3): 247-55.
Henry, K. W., A. Wyce, et al. (2003). "Transcriptional activation via sequential histone H2B ubiquitylation and deubiquitylation, mediated by SAGA-associated Ubp8." Genes Dev **17**(21): 2648-63.
Hilfiker, A., D. Hilfiker-Kleiner, et al. (1997). "mof, a putative acetyl transferase gene related to the Tip60 and MOZ human genes and to the SAS genes of yeast, is required for dosage compensation in Drosophila." EMBO J **16**(8): 2054-60.
Hirota, T., J. J. Lipp, et al. (2005). "Histone H3 serine 10 phosphorylation by Aurora B causes HP1 dissociation from heterochromatin." Nature **438**(7071): 1176-80.
Hollmann, M., E. Simmerl, et al. (2002). "The essential Drosophila melanogaster gene wds (will die slowly) codes for a WD-repeat protein with seven repeats." Mol Genet Genomics **268**(4): 425-33.
Huang, J., J. M. Hsu, et al. (2004). "The RSC nucleosome-remodeling complex is required for Cohesin's association with chromosome arms." Mol Cell **13**(5): 739-50.
Huang, J. and B. C. Laurent (2004). "A Role for the RSC chromatin remodeler in regulating cohesion of sister chromatid arms." Cell Cycle **3**(8): 973-5.
Huang, J., B. Liang, et al. (2005). "ATP-dependent chromatin-remodeling complexes in DNA double-strand break repair: remodeling, pairing and (re)pairing." Cell Cycle **4**(12): 1713-5.
Huyen, Y., O. Zgheib, et al. (2004). "Methylated lysine 79 of histone H3 targets 53BP1 to DNA double-strand breaks." Nature **432**(7015): 406-11.
Inayoshi, Y., K. Miyake, et al. (2006). "Mammalian chromatin remodeling complex SWI/SNF is essential for enhanced expression of the albumin gene during liver development." J Biochem **139**(2): 177-88.

Ito, T., M. Bulger, et al. (1997). "ACF, an ISWI-containing and ATP-utilizing chromatin assembly and remodeling factor." Cell **90**(1): 145-55.
Itoh, Y., E. Melamed, et al. (2007). "Dosage compensation is less effective in birds than in mammals." J Biol **6**(1): 2.
Jenuwein, T. and C. D. Allis (2001). "Translating the histone code." Science **293**(5532): 1074-80.
Joshi, A. A. and K. Struhl (2005). "Eaf3 chromodomain interaction with methylated H3-K36 links histone deacetylation to Pol II elongation." Mol Cell **20**(6): 971-8.
Julien, E. and W. Herr (2004). "A switch in mitotic histone H4 lysine 20 methylation status is linked to M phase defects upon loss of HCF-1." Mol Cell **14**(6): 713-25.
Kageyama, Y., G. Mengus, et al. (2001). "Association and spreading of the Drosophila dosage compensation complex from a discrete roX1 chromatin entry site." EMBO J **20**(9): 2236-45.
Karachentsev, D., K. Sarma, et al. (2005). "PR-Set7-dependent methylation of histone H4 Lys 20 functions in repression of gene expression and is essential for mitosis." Genes Dev **19**(4): 431-5.
Kelley, R. L., V. H. Meller, et al. (1999). "Epigenetic spreading of the Drosophila dosage compensation complex from roX RNA genes into flanking chromatin." Cell **98**(4): 513-22.
Kelley, R. L., I. Solovyeva, et al. (1995). "Expression of msl-2 causes assembly of dosage compensation regulators on the X chromosomes and female lethality in Drosophila." Cell **81**(6): 867-77.
Kelley, R. L., J. Wang, et al. (1997). "Sex lethal controls dosage compensation in Drosophila by a non-splicing mechanism." Nature **387**(6629): 195-9.
Kim, J., S. B. Hake, et al. (2005). "The human homolog of yeast BRE1 functions as a transcriptional coactivator through direct activator interactions." Mol Cell **20**(5): 759-70.
Kind, J. and A. Akhtar (2007). "Cotranscriptional recruitment of the dosage compensation complex to X-linked target genes." Genes Dev **21**(16): 2030-40.
Kind, J., J. M. Vaquerizas, et al. (2008). "Genome-wide analysis reveals MOF as a key regulator of dosage compensation and gene expression in Drosophila." Cell **133**(5): 813-28.
Lachner, M., D. O'Carroll, et al. (2001). "Methylation of histone H3 lysine 9 creates a binding site for HP1 proteins." Nature **410**(6824): 116-20.
Langst, G., E. J. Bonte, et al. (1999). "Nucleosome movement by CHRAC and ISWI without disruption or trans-displacement of the histone octamer." Cell **97**(7): 843-52.
Larschan, E., A. A. Alekseyenko, et al. (2007). "MSL complex is attracted to genes marked by H3K36 trimethylation using a sequence-independent mechanism." Mol Cell **28**(1): 121-33.
Lee, D. Y., C. Teyssier, et al. (2005). "Role of protein methylation in regulation of transcription." Endocr Rev **26**(2): 147-70.
Legube, G., S. K. McWeeney, et al. (2006). "X-chromosome-wide profiling of MSL-1 distribution and dosage compensation in Drosophila." Genes Dev **20**(7): 871-83.
Li, F., D. A. Parry, et al. (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**(20): 8913-24.
Lieb, J. D., M. R. Albrecht, et al. (1998). "MIX-1: an essential component of the C. elegans mitotic machinery executes X chromosome dosage compensation." Cell **92**(2): 265-77.
Lucchesi, J. C. (1998). "Dosage compensation in flies and worms: the ups and downs of X-chromosome regulation." Curr Opin Genet Dev **8**(2): 179-84.
Luger, K., A. W. Mader, et al. (1997). "Crystal structure of the nucleosome core particle at 2.8 A resolution." Nature **389**(6648): 251-60.
Lyko, F., B. H. Ramsahoye, et al. (2000). "DNA methylation in Drosophila melanogaster." Nature **408**(6812): 538-40.
Lyman, L. M., K. Copps, et al. (1997). "Drosophila male-specific lethal-2 protein: structure/function analysis and dependence on MSL-1 for chromosome association." Genetics **147**(4): 1743-53.
Macdonald, N., J. P. Welburn, et al. (2005). "Molecular basis for the recognition of phosphorylated and phosphoacetylated histone h3 by 14-3-3." Mol Cell **20**(2): 199-211.
Marin, I. (2003). "Evolution of chromatin-remodeling complexes: comparative genomics reveals the ancient origin of "novel" compensasome genes." J Mol Evol **56**(5): 527-39.
Martin, C. and Y. Zhang (2005). "The diverse functions of histone lysine methylation." Nat Rev Mol Cell Biol **6**(11): 838-49.
Mata, X., S. Taourit, et al. (2003). "Putative FLJ20436 gene characterisation in goat. Observed ubiquitous expression in goat and transgenic mice allowed to restrict the location of an hypothesised insulator element." Gene **321**: 137-44.
Melamed, E. and A. P. Arnold (2007). "Regional differences in dosage compensation on the chicken Z chromosome." Genome Biol **8**(9): R202.
Meller, V. H., P. R. Gordadze, et al. (2000). "Ordered assembly of roX RNAs into MSL complexes on the dosage-compensated X chromosome in Drosophila." Curr Biol **10**(3): 136-43.
Meller, V. H. and B. P. Rattner (2002). "The roX genes encode redundant male-specific lethal transcripts required for targeting of the MSL complex." EMBO J **21**(5): 1084-91.
Meller, V. H., K. H. Wu, et al. (1997). "roX1 RNA paints the X chromosome of male Drosophila and is regulated by the dosage compensation system." Cell **88**(4): 445-57.
Mendjan, S., M. Taipale, et al. (2006). "Nuclear pore components are involved in the transcriptional regulation of dosage compensation in Drosophila." Mol Cell **21**(6): 811-23.
Meyer, B. J. (2005). "X-Chromosome dosage compensation." WormBook: 1-14.
Meyer, B. J. and L. P. Casson (1986). "Caenorhabditis elegans compensates for the difference in X chromosome dosage between the sexes by regulating transcript levels." Cell **47**(6): 871-81.
Miller, D., J. Summers, et al. (2004). "Environmental versus genetic sex determination: a possible factor in dinosaur extinction?" Fertil Steril **81**(4): 954-64.
Min, J., Y. Zhang, et al. (2003). "Structural basis for specific binding of Polycomb chromodomain to histone H3 methylated at Lys 27." Genes Dev **17**(15): 1823-8.

Mishra, P., M. Socolich, et al. (2007). "Dynamic scaffolding in a G protein-coupled signaling system." Cell **131**(1): 80-92.
Mizuguchi, G., T. Tsukiyama, et al. (1997). "Role of nucleosome remodeling factor NURF in transcriptional activation of chromatin." Mol Cell **1**(1): 141-50.
Mohrmann, L. and C. P. Verrijzer (2005). "Composition and functional specificity of SWI2/SNF2 class chromatin remodeling complexes." Biochim Biophys Acta **1681**(2-3): 59-73.
Morales, V., C. Regnard, et al. (2005). "The MRG domain mediates the functional integration of MSL3 into the dosage compensation complex." Mol Cell Biol **25**(14): 5947-54.
Morales, V., T. Straub, et al. (2004). "Functional integration of the histone acetyltransferase MOF into the dosage compensation complex." EMBO J **23**(11): 2258-68.
Morrison, A. J., J. Highland, et al. (2004). "INO80 and gamma-H2AX interaction links ATP-dependent chromatin remodeling to DNA damage repair." Cell **119**(6): 767-75.
Muller, J., C. M. Hart, et al. (2002). "Histone methyltransferase activity of a Drosophila Polycomb group repressor complex." Cell **111**(2): 197-208.
Munks, R. J., J. Moore, et al. (1991). "Histone H4 acetylation in Drosophila. Frequency of acetylation at different sites defined by immunolabelling with site-specific antibodies." FEBS Lett **284**(2): 245-8.
Nardini, M. and B. W. Dijkstra (1999). "Alpha/beta hydrolase fold enzymes: the family keeps growing." Curr Opin Struct Biol **9**(6): 732-7.
Narlikar, G. J., H. Y. Fan, et al. (2002). "Cooperation between complexes that regulate chromatin structure and transcription." Cell **108**(4): 475-87.
Nguyen, D. K. and C. M. Disteche (2006). "Dosage compensation of the active X chromosome in mammals." Nat Genet **38**(1): 47-53.
Nielsen, P. R., D. Nietlispach, et al. (2002). "Structure of the HP1 chromodomain bound to histone H3 methylated at lysine 9." Nature **416**(6876): 103-7.
Noma, K., C. D. Allis, et al. (2001). "Transitions in distinct histone H3 methylation patterns at the heterochromatin domain boundaries." Science **293**(5532): 1150-5.
Nowak, S. J. and V. G. Corces (2004). "Phosphorylation of histone H3: a balancing act between chromosome condensation and transcriptional activation." Trends Genet **20**(4): 214-20.
Nusbaum, C. and B. J. Meyer (1989). "The Caenorhabditis elegans gene sdc-2 controls sex determination and dosage compensation in XX animals." Genetics **122**(3): 579-93.
Oh, H., Y. Park, et al. (2003). "Local spreading of MSL complexes from roX genes on the Drosophila X chromosome." Genes Dev **17**(11): 1334-9.
Okada, Y., Q. Feng, et al. (2005). "hDOT1L links histone methylation to leukemogenesis." Cell **121**(2): 167-78.
Pannuti, A. and J. C. Lucchesi (2000). "Recycling to remodel: evolution of dosage-compensation complexes." Curr Opin Genet Dev **10**(6): 644-50.
Park, S. W., M. I. Kuroda, et al. (2008). "Regulation of histone H4 Lys16 acetylation by predicted alternative secondary structures in roX noncoding RNAs." Mol Cell Biol **28**(16): 4952-62.
Park, Y., G. Mengus, et al. (2003). "Sequence-specific targeting of Drosophila roX genes by the MSL dosage compensation complex." Mol Cell **11**(4): 977-86.

Payer, B. and J. T. Lee (2008). "X chromosome dosage compensation: how mammals keep the balance." Annu Rev Genet **42**: 733-72.

Plath, K., J. Fang, et al. (2003). "Role of histone H3 lysine 27 methylation in X inactivation." Science **300**(5616): 131-5.

Puig, O., F. Caspary, et al. (2001). "The tandem affinity purification (TAP) method: a general procedure of protein complex purification." Methods **24**(3): 218-29.

Rath, U., D. Wang, et al. (2004). "Chromator, a novel and essential chromodomain protein interacts directly with the putative spindle matrix protein skeletor." J Cell Biochem **93**(5): 1033-47.

Raymond, C. S., J. R. Kettlewell, et al. (1999). "Expression of Dmrt1 in the genital ridge of mouse and chicken embryos suggests a role in vertebrate sexual development." Dev Biol **215**(2): 208-20.

Rea, S., F. Eisenhaber, et al. (2000). "Regulation of chromatin structure by site-specific histone H3 methyltransferases." Nature **406**(6796): 593-9.

Reid, G., R. Gallais, et al. (2009). "Marking time: the dynamic role of chromatin and covalent modification in transcription." Int J Biochem Cell Biol **41**(1): 155-63.

Reifsnyder, C., J. Lowell, et al. (1996). "Yeast SAS silencing genes and human genes associated with AML and HIV-1 Tat interactions are homologous with acetyltransferases." Nat Genet **14**(1): 42-9.

Richter, L., J. R. Bone, et al. (1996). "RNA-dependent association of the Drosophila maleless protein with the male X chromosome." Genes Cells **1**(3): 325-36.

Roth, S. Y., J. M. Denu, et al. (2001). "Histone acetyltransferases." Annu Rev Biochem **70**: 81-120.

Salma, N., H. Xiao, et al. (2004). "Temporal recruitment of transcription factors and SWI/SNF chromatin-remodeling enzymes during adipogenic induction of the peroxisome proliferator-activated receptor gamma nuclear hormone receptor." Mol Cell Biol **24**(11): 4651-63.

Sanders, S. L., M. Portoso, et al. (2004). "Methylation of histone H4 lysine 20 controls recruitment of Crb2 to sites of DNA damage." Cell **119**(5): 603-14.

Santos-Rosa, H., R. Schneider, et al. (2002). "Active genes are tri-methylated at K4 of histone H3." Nature **419**(6905): 407-11.

Scott, M. J., L. L. Pan, et al. (2000). "MSL1 plays a central role in assembly of the MSL complex, essential for dosage compensation in Drosophila." EMBO J **19**(1): 144-55.

Shen, X., G. Mizuguchi, et al. (2000). "A chromatin remodelling complex involved in transcription and DNA processing." Nature **406**(6795): 541-4.

Shim, S., C. S. Yoon, et al. (2000). "A novel gene family with a developmentally regulated expression in Xenopus laevis." Biochem Biophys Res Commun **267**(2): 558-64.

Sif, S., A. J. Saurin, et al. (2001). "Purification and characterization of mSin3A-containing Brg1 and hBrm chromatin remodeling complexes." Genes Dev **15**(5): 603-18.

Silva, J., W. Mak, et al. (2003). "Establishment of histone h3 methylation on the inactive X chromosome requires transient recruitment of Eed-Enx1 polycomb group complexes." Dev Cell **4**(4): 481-95.

Sims, R. J., 3rd, C. F. Chen, et al. (2005). "Human but not yeast CHD1 binds directly and selectively to histone H3 methylated at lysine 4 via its tandem chromodomains." J Biol Chem **280**(51): 41789-92.

Smith, E. R., C. Cayrou, et al. (2005). "A human protein complex homologous to the Drosophila MSL complex is responsible for the majority of histone H4 acetylation at lysine 16." Mol Cell Biol **25**(21): 9175-88.
Smith, E. R., A. Eisen, et al. (1998). "ESA1 is a histone acetyltransferase that is essential for growth in yeast." Proc Natl Acad Sci U S A **95**(7): 3561-5.
Song, H., Y. Li, et al. (2004). "Human MCRS2, a cell-cycle-dependent protein, associates with LPTS/PinX1 and reduces the telomere length." Biochem Biophys Res Commun **316**(4): 1116-23.
Sterner, D. E. and S. L. Berger (2000). "Acetylation of histones and transcription-related factors." Microbiol Mol Biol Rev **64**(2): 435-59.
Stokes, D. G. and R. P. Perry (1995). "DNA-binding and chromatin localization properties of CHD1." Mol Cell Biol **15**(5): 2745-53.
Stokes, D. G., K. D. Tartof, et al. (1996). "CHD1 is concentrated in interbands and puffed regions of Drosophila polytene chromosomes." Proc Natl Acad Sci U S A **93**(14): 7137-42.
Stuckenholz, C., V. H. Meller, et al. (2003). "Functional redundancy within roX1, a noncoding RNA involved in dosage compensation in Drosophila melanogaster." Genetics **164**(3): 1003-14.
Sykes, S. M., H. S. Mellert, et al. (2006). "Acetylation of the p53 DNA-binding domain regulates apoptosis induction." Mol Cell **24**(6): 841-51.
Taipale, M. and A. Akhtar (2005). "Chromatin mechanisms in Drosophila dosage compensation." Prog Mol Subcell Biol **38**: 123-49.
Taipale, M., S. Rea, et al. (2005). "hMOF histone acetyltransferase is required for histone H4 lysine 16 acetylation in mammalian cells." Mol Cell Biol **25**(15): 6798-810.
Teranishi, M., Y. Shimada, et al. (2001). "Transcripts of the MHM region on the chicken Z chromosome accumulate as non-coding RNA in the nucleus of female cells adjacent to the DMRT1 locus." Chromosome Res **9**(2): 147-65.
Tsukiyama, T., C. Daniel, et al. (1995). "ISWI, a member of the SWI2/SNF2 ATPase family, encodes the 140 kDa subunit of the nucleosome remodeling factor." Cell **83**(6): 1021-6.
Tsukiyama, T. and C. Wu (1997). "Chromatin remodeling and transcription." Curr Opin Genet Dev **7**(2): 182-91.
Turner, B. M., A. J. Birley, et al. (1992). "Histone H4 isoforms acetylated at specific lysine residues define individual chromosomes and chromatin domains in Drosophila polytene nuclei." Cell **69**(2): 375-84.
van Attikum, H., O. Fritsch, et al. (2004). "Recruitment of the INO80 complex by H2A phosphorylation links ATP-dependent chromatin remodeling with DNA double-strand break repair." Cell **119**(6): 777-88.
Varga-Weisz, P. D., M. Wilm, et al. (1997). "Chromatin-remodelling factor CHRAC contains the ATPases ISWI and topoisomerase II." Nature **388**(6642): 598-602.
Vradii, D., S. Wagner, et al. (2006). "Brg1, the ATPase subunit of the SWI/SNF chromatin remodeling complex, is required for myeloid differentiation to granulocytes." J Cell Physiol **206**(1): 112-8.
Wang, W. (2003). "The SWI/SNF family of ATP-dependent chromatin remodelers: similar mechanisms for diverse functions." Curr Top Microbiol Immunol **274**: 143-69.
Wang, Y., J. Wysocka, et al. (2004). "Human PAD4 regulates histone arginine methylation levels via demethylimination." Science **306**(5694): 279-83.

Woolfe, A. and G. Elgar (2008). "Organization of conserved elements near key developmental regulators in vertebrate genomes." Adv Genet **61**: 307-38.

Wu, J. L., Y. S. Lin, et al. (2009). "MCRS2 represses the transactivation activities of Nrf1." BMC Cell Biol **10**: 9.

Wysocka, J., T. Swigut, et al. (2005). "WDR5 associates with histone H3 methylated at K4 and is essential for H3 K4 methylation and vertebrate development." Cell **121**(6): 859-72.

Young, D. W., J. Pratap, et al. (2005). "SWI/SNF chromatin remodeling complex is obligatory for BMP2-induced, Runx2-dependent skeletal gene expression that controls osteoblast differentiation." J Cell Biochem **94**(4): 720-30.

Zhang, Y. and D. Reinberg (2001). "Transcription regulation by histone methylation: interplay between different covalent modifications of the core histone tails." Genes Dev **15**(18): 2343-60.

Zhou, S., Y. Yang, et al. (1995). "Male-specific lethal 2, a dosage compensation gene of Drosophila, undergoes sex-specific regulation and encodes a protein with a RING finger and a metallothionein-like cysteine cluster." EMBO J **14**(12): 2884-95.