

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

Diplom-Biologe Philipp Gebhardt

born in Stuttgart, Germany

Date of the oral examination:

22 November 2007

**Purification and biochemical characterisation of
novel MOF-containing NSL complexes**

Referees:

Prof. Dr. Michael Sattler

Prof. Dr. Gabriele Petersen

SUMMARY

Every cell of a eukaryotic organism contains the whole genome in a membrane-bound nucleus where the DNA is compacted in highly organized chromatin fibers. The association of DNA with histone proteins allows for efficient packaging of the genetic material. Chromatin is a dynamic polymer that is packed and unwrapped in a highly regulated manner to match the multiple tasks that it encounters during the lifetime of a cell. Regulation of DNA accessibility is achieved by the concerted action of chromatin-associated proteins, such as chromatin remodeling enzymes, variant histone proteins and chromatin modifying enzymes. The latter class of enzymes brings about post-translational modifications of histones. Modifications, such as acetylation, can lead to changes in chromatin structure either directly or by recruiting other effector proteins.

A process that is linked to histone acetylation is dosage compensation in the fruit-fly *Drosophila melanogaster* which is studied as a model system for the regulation of chromatin. The multi-protein complex involved, harbours an enzyme, MOF (males absent on the first), with histone acetyltransferase activity directed towards histone H4 lysine 16. Other proteins that have previously been known to associate with it are the MSL (male specific lethal) proteins.

Recent biochemical purification of MOF containing complexes revealed that, in addition to the MSL proteins, a number of novel proteins co-purified with MOF in *Drosophila* and mammals (Mendjan et al., 2006). During my PhD I was therefore interested to study whether these novel proteins, which we named NSL proteins (non-specific lethal), exist in a complex similar to the MSL proteins in *Drosophila* and if so, what might be their function.

Interestingly, we found that one of the NSL proteins, NSL1 has a very similar domain architecture to MSL-1. A major part of this thesis has therefore been to study the NSL1 protein in detail. Using co-immunoprecipitation as well as *in vitro* interaction assays, I was able to show that NSL1 indeed interacts directly with MOF.

By applying an affinity purification strategy tagging the NSL1 protein, my PhD work has demonstrated that NSL1 co-purifies with other NSL proteins in a complex that is distinct from the MSL complex. I have also been able to show by immunofluorescence microscopy that two components of this complex, NSL1 and MCRS2, co-localise on hundreds of sites on all polytene chromosomes, suggesting that very likely these

proteins not only interact biochemically but may also function together *in vivo*. This part of my work has therefore provided novel insights into the existence of the NSL complex in *Drosophila* and has showed that MOF associates with two distinct sets of proteins, namely the MSLs and the NSLs.

In the second part of my PhD work I studied the overall contribution of MSL and NSL complexes in modulating histone acetylation using quantitative mass spectrometric analysis of endogenous histones from *Drosophila* cells that were RNAi-depleted of MOF, MSL-1 and NSL1. This work revealed that MOF contributes to a majority of histone H4 K16 acetylation in *Drosophila* cell lines and that MSL-1 plays an important role in modulating the activity of MOF *in vivo*. Surprisingly, we did not observe any major changes in histone acetylation levels upon NSL1 depletion *in vivo*.

Interestingly, NSL1 depleted cells displayed cell proliferation and segregation defects. The *in vivo* function of NSL1 remains elusive. Future work will therefore be required to elucidate the mechanism of action of these novel proteins in *Drosophila*. The foundations for it were layed by this work.

ZUSAMMENFASSUNG

Das Genom im Inneren jeder eukaryotischen Zelle wird von einer Kernmembran umgeben. Die darin eingeschlossene DNA ist stark komprimiert und in hochkomplexen Chromatinfasern organisiert. Die Assoziation der DNA mit Histon-Proteinen erlaubt eine effiziente Verpackung des genetischen Materials. Das Chromatin ist ein dynamisches Polymer, welches sich in regulierter Weise kondensiert und dekondensiert, um den vielfältigen zellulären Aufgaben zu genügen. Die Regulierung des Zugangs zur DNA wird durch die konzertierte Aktion Chromatin-assoziiierter Proteine, wie ‚Chromatin-remodeler‘-Enzymen, Histon-Varianten und Chromatin-modifizierender Enzyme erreicht. Die zuletztgenannte Enzymklasse bewerkstelligt die post-translationale Modifikation von Histonen. Solche Modifikationen, die Acetylierung als Beispiel, können entweder direkt oder durch die Rekrutierung anderer Effektor-Proteine zu strukturellen Veränderungen des Chromatins führen.

Mit der Histon-Acetylierung eng verknüpft, ist der Prozess der Dosiskompensation in der Fruchtfliege *Drosophila melanogaster*, welcher als Modellsystem für die Erforschung der Chromatinregulation dient. Der involvierte Multiproteinkomplex beinhaltet ein Enzym, MOF (males absent on the first), mit H4K16-spezifischer Histonacetyltransferase-Aktivität. Andere Proteine, die schon länger als MOF-assoziierte Proteine bekannt sind, umfassen die MSL (male specific lethal) Proteine. Die jüngste biochemische Aufreinigung MOF-enthaltender Komplexe, brachte zusätzlich zu den MSL Proteinen, eine Reihe an neuartigen, mit MOF interagierenden *Drosophila*- und Säuger-Proteinen zum Vorschein (Mendjan et al., 2006). In meiner Doktorarbeit war ich daher daran interessiert, ob diese neuartigen Proteine, die wir NSL (non-specific lethal) Proteine nannten, in einem Komplex (ähnlich der MSL-Proteine) vorliegen und was deren Funktion ist. Interessanterweise fanden wir, dass eines der NSL-Proteine, NSL1, eine MSL-1-ähnliche Domänenarchitektur aufweist. Ein Großteil dieser Doktorarbeit befasst sich daher mit der detaillierten Analyse des NSL1-Proteins. Unter Verwendung von Koimmunopräzipitations- und *in vitro* Interaktions-Experimenten, konnte ich zeigen, dass NSL1 eine direkte Interaktion mit MOF eingeht.

Durch die Strategie einer Affinitätsreinigung des NSL1-Proteins konnte ich in meiner Doktorarbeit zeigen, dass NSL1 mit anderen NSL-Proteinen in einem

Komplex interagiert, der verschieden vom MSL-Komplex ist. Ich konnte weiterhin durch Immunfluoreszenzmikroskopie demonstrieren, dass zwei Komponenten dieses Komplexes, NSL1 und MCRS2, an Hunderten von Stellen auf allen polytären Chromosomen kolokalisierten, was nahelegte, dass diese Proteine sehr wahrscheinlich nicht nur biochemisch interagieren sondern auch *in vivo* zusammenarbeiten. Dieser Teil meiner Arbeit hat dadurch neue Einblicke in die Existenz des NSL-Komplexes in *Drosophila* gewährt und konnte zeigen, dass MOF mit zwei unterschiedlichen Sätzen an Proteinen, den MSL- und den NSL-Proteinen, assoziiert vorliegt.

Im zweiten Teil meiner Doktorarbeit analysierte ich die Gesamtbeteiligung der MSL- und NSL-Komplexe an der Modulation der Histonacetylierung. Dies wurde durch eine quantitative massenspektrometrische Analyse von endogenen Histonen aus *Drosophila* Zelllinien, die mittels RNA-Interferenz von MOF, MSL-1 oder NSL1 depletiert wurden, erreicht. Diese Arbeit offenbarte, dass MOF an der Mehrheit der H4K16 –Acetylierung in *Drosophila* Zelllinien beteiligt ist und dass MSL-1 eine wichtige Rolle bei der Regulierung der MOF-Aktivität *in vivo* übernimmt. Überraschenderweise konnten wir keine wesentlichen Veränderungen des Histon-Acetylierungsniveaus nach NSL1 Depletion *in vivo* feststellen. Interessanterweise zeigten NSL1-depletierte Zellen Defekte in der Zellproliferation und –segregation. Die Funktion von NSL1 *in vivo* bleibt schwer zu fassen. Zukünftige Arbeiten werden daher notwendig sein, um das Wirkprinzip dieser neuen *Drosophila* Proteine zu erklären. Der Grundstein dafür wurde gelegt.

SUMMARY.....	X
--------------	---

ZUSAMMENFASSUNG.....	XII
----------------------	-----

INTRODUCTION

1 Chromatin – definition and function	5
2 The structure of chromatin	6
2.1 The different flavours of chromatin: euchromatin and heterochromatin	8
3 Chromatin dynamics and the regulation of chromatin structure.....	9
3.1 Chromatin-remodeling	10
3.2 Chromatin modifications and their signaling potential.....	12
3.2.1 Post-translational modifications of histone tails	12
3.2.2 Post-translational modifications of histone cores.....	15
3.2.3 Translating the message – Effector proteins and their specialized domains.....	16
3.2.4 ‘Message disposal’ – Enzymes that erase the modifications.....	18
3.3 For a change – Histone variants.....	20
4 Focus on histone acetyltransferases (HATs)	22
4.1 Classification of HATs	23
4.1.1 Catalytic HAT-domains and their function	24
4.2 HATs – specific non-soloists.....	25
4.3 Biological relevance of HATs.....	25
4.3.1 Balancing acetylation – A delicate matter.....	26
4.3.2 Global versus gene-specific histone acetylation	27
5 Dosage Compensation – a model for the regulation of chromatin.....	28
5.1 Different Species – Different Mechanisms.....	29
5.2 A closer look at dosage compensation in <i>Drosophila</i>	30
5.2.1 The DCC – confined to the X.....	31
5.2.2 MOF and the 2x X	34
6 The MSL complex and beyond.....	35
6.1 <i>Drosophila</i> MSL-1 and NSL1 – similar players.....	36

AIMS AND OBJECTIVES

MATERIALS AND METHODS

7	Cloning of full-length NSL1 and NSL3 cDNAs	38
7.1	NSL1	38
7.2	NSL3	39
7.3	Sequence alignments	39
8	Generation of NSL1- and NSL3-specific antibodies	39
8.1	Expression and purification of NSL antigens	39
8.1.1	Protein expression.....	39
8.1.2	Affinity purification of GST-His ₆ -NSL fusion proteins	40
8.2	Antigens.....	41
8.3	NSL1 and NSL3 antibodies.....	41
9	Coimmunoprecipitation experiments (CoIPs)	42
10	Western Blot	42
11	Recombinant protein expression using the Baculovirus / SF9 system	43
11.1	Generation of NSL1 and NSL3 Baculoviruses	43
11.2	Baculovirus protein expression and purification from SF9 cells	43
11.3	Co-expression and purification	44
12	Purification of the NSL complex from <i>Drosophila melanogaster</i> Schneider cells	45
12.1	Affinity-tagging of the NSL1 protein.....	45
12.2	Generation of stable <i>Drosophila melanogaster</i> Schneider cell lines	46
12.3	Preparation of nuclear extracts.....	47
12.4	Tandem affinity purification (TAP)	48
12.5	FLAG/HA affinity purification	49
13	Analysis of the purified NSL complexes	50
13.1	Silver-staining of SDS polyacrylamide gels	50
13.2	Flamingo Staining of SDS polyacrylamide gels.....	51
13.3	Identification of interacting proteins by mass spectrometry.....	51
14	Biochemical characterization of NSL complexes	52
14.1	<i>In vitro</i> assembly of polynucleosomes	52
14.2	Histone acetyltransferase assays (HAT assays).....	52
14.3	Site-directed mutagenesis of histone H4 tail lysine residues	53
14.4	Preparation of wild-type and H4-mutant histone octamers.....	54
14.4.1	Expression of recombinant histone proteins	54
14.4.2	Inclusion Body Preparation.....	54
14.4.3	Histone unfolding.....	54
14.4.4	Histone purification by ion exchange chromatography.....	55
14.4.5	Histone refolding and reconstitution of histone octamers	55
15	Targeted quantitative mass spec analysis of histone lysine acetylation	56
16	RNA interference in <i>Drosophila</i> S2 and Kc cells	57

17	Quantitative RT-PCR to assess knockdown efficiency	58
18	Acid histone extraction of dsRNA-treated <i>Drosophila</i> S2 and Kc cells	59
19	Immunofluorescence staining of S2 and Kc cells	59
19.1	Fixation of cells	59
19.2	Immunostaining of cells.....	60
20	Immunofluorescence staining of polytene chromosome squashes	60
20.1	Preparation of chromosome squashes	60
20.2	Immunostaining of polytene chromosome squashes.....	60
21	Confocal microscopy	61
22	EGFP-tagging of NSL1 for transfection in <i>Drosophila</i> S2 cells	61
23	EGFP-tagging of <i>Drosophila</i> histone H2B and generation of a stable S2 cell line	62
24	Live Cell Imaging	62

RESULTS AND DISCUSSION

25	Getting started - cloning of the full-length NSL1 and NSL3 cDNAs	64
26	Generating the tools for detection – antibodies against NSL1 and NSL3	66
27	Intracellular localization of the <i>Drosophila</i> NSL1 protein	69
27.1	NSL1 is a nuclear protein.....	69
27.2	NSL1 localizes to chromosomes	73
28	Who interacts with whom – Coimmunoprecipitations reveal the dichotomy	75
29	The <i>Drosophila</i> NSL1 PEHE domain mediates MOF interaction	78
30	Affinity-purification and analysis of the <i>Drosophila</i> NSL complex	80
30.1	Analysis of eluted proteins by western blot.....	83
30.2	Identification of NSL1 interacting proteins by mass spectrometry	85
30.2.1	<i>Drosophila</i> MSL-1 and NSL1 – MOF's exclusive friends	88
30.3	Description of identified <i>Drosophila</i> NSL complex members	89
30.3.1	Confirmed members of the NSL complex	89
30.3.2	Newly identified members of the NSL complex	92
30.4	NSL1 co-localizes with NSL complex members on polytene chromosomes	93
31	Biochemical characterisation of the histone acetyltransferase activity in the NSL complex	95
31.1	MOF retains histone H4 specificity upon integration in the NSL complex.....	95
31.2	The NSL complex preferentially acetylates H4K16.....	96
31.2.1	Histone acetyltransferase assays with mutant polynucleosomes as substrate	97
31.2.2	Mass spectrometric analysis reveals the specificity of the NSL complex	99
32	Targeted quantitative analysis of lysine acetylation by MOF	101

32.1	Quantification of MOF-dependent H4K16 acetylation.....	105
33	Effect of NSL1 knockdown in S2 cells – the NSL1 RNAi phenotype	108
33.1	RNAi-mediated knockdown of NSL1 results in a retarded growth phenotype of <i>Drosophila</i> cell culture cells	109
33.1.1	Knockdown of NSL1 impacts on NSL3 protein stability	111
33.2	NSL1 knockdown cells exhibit apoptotic features	112
33.3	Live cell imaging	113
33.3.1	RNAi-mediated knockdown of NSL1 causes severe chromosome segregation defects	113
	FINAL DISCUSSION AND PERSPECTIVES.....	117
	REFERENCES.....	122
	ABBREVIATIONS.....	133
	ACKNOWLEDGEMENTS.....	135
	PUBLICATIONS.....	136

INTRODUCTION

Isn't it an overwhelming thought that all of the information that is needed to build a complex organism like a human being is contained in its genome? This genetic library holds all the instructions to form a whole body plan and to endow it with specific functions. It is not less incredible that this vast amount of genetic information is packed into every cell of an organism. From single-celled protists to multicellular organisms the genetic material, DNA, is contained in a membrane-bound nucleus inside the cell. This compartmentalization has a functional necessity but at the same time minimizes the space that is available to fit the genetic material in the cell. For illustration, human cells need to package approximately 2 meters of DNA into their nuclei that have an average diameter of about 5 to 10 μm . Depending on the cell type, the nuclear volume normally comprises 5 - 10 % of the cellular volume.

It is self-evident, that the genetic material needs a high degree of compaction while maintaining some functional organization in the nucleus. To achieve this goal, DNA in the nucleus is wrapped around histone proteins, which are evolutionarily conserved proteins sharing similar structural motifs. Being highly basic proteins, their positive charge allows them to bind tightly to the acidic DNA, thereby forming repeating units of so called nucleosomes that are sitting on the DNA like "beads-on-a-string" (Olins et al., 1974). Higher-order states of compaction are necessary to finally build up the chromosomes.

1 Chromatin – definition and function

The term chromatin - DNA decorated with nucleosomes – was coined by Walther Flemming when he made an observation in 1882. He recognized that the substance in the nucleus, which was called the 'nuclear-scaffold' at that time, could be stained by a dye. He phrased his conclusions as follows: "The scaffold owes its capability of refraction, the way how it behaves, and in particular its colorability to a substance which, with regard to its latter attribute, I have termed Chromatin. I'll retain the name Chromatin as long as Chemistry has decided

about it, and I empirically refer to it as that substance in the cell's nucleus which takes up the dye upon staining the nucleus”.

The term chromatin still stands today and we have learned a lot about its structure and many of its constituents. It became clear that it is not only the DNA and the associating histones that form chromatin but also other proteins (e.g. chromatin-bound enzymes, transcription factors, scaffold proteins) and RNA molecules are part of this polymer.

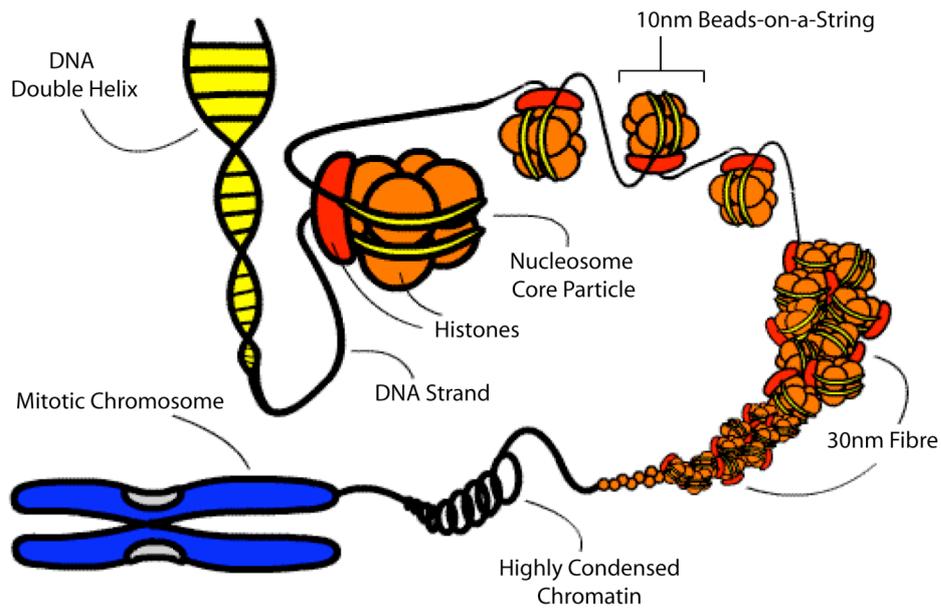
There are several essential functions that chromatin has to fulfil: First of all, it compacts the DNA into a small volume. Secondly, it regulates the flow of information that is stored in it, e.g. it makes DNA sequences accessible for transcription, it allows the DNA to be replicated and to be repaired when needed. Another very important feature of chromatin is that it has to ensure proper and damage-safe transmission of the genetic data to the next cell generation.

2 The structure of chromatin

The basic unit of chromatin is the nucleosome which was identified by Roger Kornberg in 1974 (Kornberg, 1974). The nucleosome consists of a histone octamer that is composed of the five canonical histones, H1, H2A, H2B, H3 and H4 (see Figure 2-1 B). A tetramer of histones H3 and H4 and two dimers of histones H2A and H2B are wrapped by approximately 1.7 turns (~147 base pairs) of DNA (Davey et al., 2002).

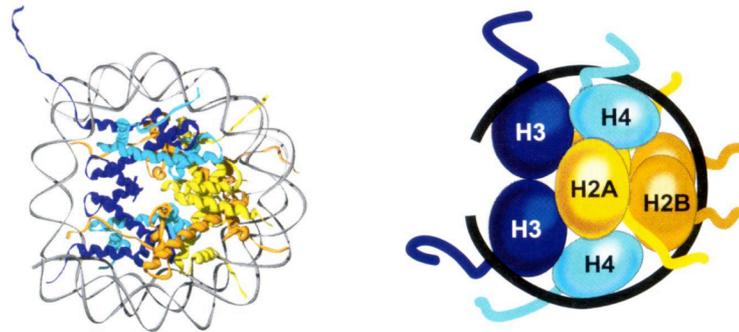
The physiological form of euchromatin that contains actively transcribed genes is thought to be organised in the highly dynamic 10nm fibre (beads-on-a-string), which consists of nucleosomes that are distributed on the DNA. In 1989, Stillman et al. (Smith et al., 1989) found, that a purified human protein, Chromatin assembly factor 1 (CAF-1), was able to wrap newly-synthesized DNA around histones. The generated ‘beads-on-a-string’ structure allows for transcription when transversed by RNA polymerase molecules.

A From nucleosome to higher order chromatin structure



Adapted from original by Kyle Dunn

B Nucleosome structure



C Higher order structural models of the 30nm fibre

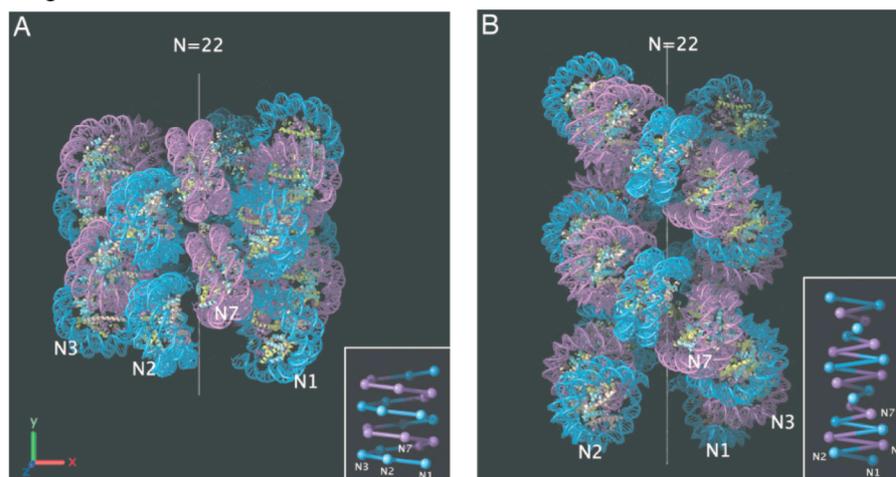


Figure 2-1: From nucleosome to higher order chromatin structure

(A) From nucleosome to higher order chromatin structure. Different organizational states of chromatin are depicted (B) Nucleosome structure (C) Higher order structural models of the 30nm fibre. Left: Interdigitated one-start helix model (Robinson et al.) , Right: Two-start helical crossed linker model (adapted from Schalch et al.)

By involving the linker histone H1, the 'beads-on-a-string' structure can compact further and form the 30nm fibre. However, the exact structure of this EM-visible filament was debated for a long time. Different models were applied to explain the real configuration (Figure 2-1 C). In 2004, experimental evidence was presented which favoured the helical two-start model rather than the solenoid model (Dorigo et al., 2004). However, accurate electron microscope based measurements of *in vitro* reconstituted 30nm fibres more recently suggested a compact one-start helical configuration where the nucleosomes are organized in an interdigitating manner (Robinson et al., 2006).

Using high-resolution approaches will help to finally resolve this conflicting issue in the coming years and elucidate the detailed structure of the 30-nm chromatin fiber. The 30nm fibre itself can be folded into higher-order structures, ultimately yielding highly compacted chromatids and metaphase chromosomes which are transcriptionally silent. This compaction is dynamic since the different levels of compaction vary throughout the stages of the cell cycle.

2.1 The different flavours of chromatin: euchromatin and heterochromatin

A long standing view on chromatin implied that there are two forms of chromatin that are present in non-dividing cells: heterochromatin and euchromatin. This was first recognized by E. Heitz (Heitz, 1928) when he detected differences in the chromosomal staining behaviour in moss species. Since then, more characteristic features have been attributed to these two fractions of chromatin. Briefly, heterochromatin was thought to correspond to the highly compacted fraction of chromatin that is enriched in repetitive sequences and is not transcribed (Craig, 2005), whereas euchromatin is open, actively transcribed chromatin. However, it turned out that this distinction was rather too simplistic. There are in fact heterochromatic regions where transcription can occur and vice versa not all of the genes in euchromatin are transcribed (Dimitri et al., 2005; Gilbert et al., 2004).

The plasticity of chromatin is becoming even more apparent by highlighting that heterochromatin can be further subdivided in constitutive and facultative heterochromatin.

Facultative heterochromatin, also regarded as the silenced chromatin state, is developmentally regulated and can vary between individuals or tissues. It is assembled when needed to silence genes and it can encompass chromosome regions, complete chromosomes or even whole genomes (Dimitri et al., 2005). A prominent example of facultative heterochromatin is the random inactivation of one of the X chromosomes (Xi) in cells of female mammals (Avner et al., 2001; Lyon, 1961). This happens early in embryogenesis and the resulting chromosome-wide silenced heterochromatic structure is called the Barr body.

Interestingly, it has been observed that neocentromere formation, which is the *de novo* centromere formation from ectopic locations, is often accompanied by a switch from a euchromatic to a heterochromatic state (Amor et al., 2002).

In contrast, constitutive heterochromatin is permanent in all cells and remains condensed. It is enriched in highly repetitive DNA sequences and in transposable element-related sequences. It was previously thought to represent a waste dump for 'junk' DNA with little functional significance (Ohno, 1972). However, this view was revised with the assignment of functional domains of gene expression within constitutive heterochromatin (Reinhart et al., 2002; Saffery et al., 2003; Volpe et al., 2002). Constitutive heterochromatin is involved in centromere formation and can be found at telomeres (Craig, 2005).

3 Chromatin dynamics and the regulation of chromatin structure

Generally speaking, chromatin structure and the degree of compaction are directly linked to the accessibility of the underlying DNA sequences and to their potential to be replicated, transcribed or repaired. It becomes clear, that there must be mechanisms existing that are able to control chromatin structure in order to ensure timely expression of genes, DNA repair, replication and other processes. It is the interplay of chromatin-remodelling complexes, chromatin-modifying enzymes and histone variants that act in the cell to organize chromatin structure. The different components will be described in the next sections.

3.1 Chromatin-remodeling

The DNA inside the nucleosome is normally not accessible for DNA-binding factors. Yet, nucleosomes are stable and exhibit limited mobility on their own. It was shown that artificially immobilized nucleosomes are able to prevent transcription by RNA polymerase (Gottesfeld et al., 2002). Rendering them dynamic requires so called chromatin-remodelling complexes (Saha et al., 2006). In general, the remodelers use energy from the hydrolysis of ATP to transiently disrupt the histone–DNA interactions and rearrange nucleosomes. Specialized remodelling complexes have evolved to provide access to the underlying DNA and allow for transcription, repair, chromosome condensation and other chromatin-related processes.

The core of all chromatin remodelers contains a helicase-like ATPase subunit which belongs to the SWI2/SNF2 superfamily (Smith et al., 2005). The yeast SWI/SNF complex was the first identified chromatin remodeling complex. The eponymous subunits, Swi2 and Snf2, were initially identified in *S. cerevisiae* as positive transcriptional regulators of HO and SUC2 genes.

Swi, Swi2 and Swi3 proteins are involved in regulation of the HO endonuclease gene, which in turn is required for mating type switching - hence the name SWItch (Peterson et al., 1992; Stern et al., 1984). The invertase encoding SUC2 gene, is belonging to the sucrose-non-fermenter genes SNF2, SNF5 and SNF6 - giving rise to the name SNF (Neigeborn et al., 1984). Later research revealed that SWI2 and SNF2 essentially represented the same gene and that the gene products were working together in a complex to positively regulate transcription (Peterson et al., 1994; Peterson et al., 1992).

The catalytic ATPase domain that is inherent to all known chromatin remodelers can be used to subgroup these enzymes into five different classes: SWI/SNF, ISWI, NURD/Mi-2/CHD, INO80 and SWR1 (Saha et al., 2006).

Normally, the ATPase subunits are incorporated in large multi-protein complexes that differ in the composition of their various core members. For example, remodelers of the ISWI class share an identical ATPase subunit, ISWI. The biological specialization is achieved by association of unique proteins to form complexes like ACF (ATP-utilizing chromatin assembly and remodeling factor), NURF (nucleosome-remodeling factor), and CHRAC (chromatin accessibility

complex) (Vignali et al., 2000). This reflects the *in vivo* specialization of different complexes and some main tasks can be assigned to the different classes:

SWI/SNF remodelers are transcriptional regulators, which organize the nucleosome positioning to enhance the binding of transcription-factors (Martens et al., 2003). In contrast, ISWI complexes adopted functions in chromatin assembly after DNA replication, maintenance of chromosome structure and in – mostly repressive – transcriptional regulation (Corona et al., 2004; Haushalter et al., 2003; Langst et al., 2001). However, the chromatin context also influences the specific tasks that are accomplished by these complexes (Saha et al., 2006).

It remains the question how chromatin remodeling is achieved on the molecular scale? There are multiple ways proposed how chromatin remodelers manage to mobilize nucleosomes and change chromatin structure. Accumulated data is mostly based on *in vitro* experiments and is still under debate. However, some general principles can be described.

The movement of a nucleosome translationally along the DNA is called ‘sliding’ and results in the exposure of a formerly occluded DNA region. Nucleosome-sliding was proposed to be a mechanism used by some of the SWI/SNF and ISWI (e.g. ACF and CHRAC) remodelling complexes. *In vitro* experiments on nucleosome arrays showed that their action resulted in contrasting nucleosome-sliding properties. ISWI remodelers were able to phase disordered nucleosome arrays (Ito et al., 1997; Varga-Weisz et al., 1997) whereas SWI/SNF remodelers did not show the ability to evenly space nucleosomes (Flaus et al., 2003; Owen-Hughes et al., 1996).

The NURF complex, possessing the ISWI ATPase as a catalytic subunit, was shown to redistribute nucleosomes *in vitro* by transiently decreasing the activation energy and possibly ‘looping’ the DNA around the nucleosome particle (Hamiche et al., 1999).

Yet, another remodelling mechanism was uncovered by studying the *Drosophila* HSP70 gene (Schwartz et al., 2005). It turned out that a transcription elongation-coupled histone replacement was taking place at sites of active transcription. In this case, histone H3 was exchanged with the histone variant H3.3 to create variant nucleosomes.

3.2 Chromatin modifications and their signaling potential

Chromatin-remodeling enzymes as one class of major players that are able to regulate the dynamic nature of chromatin were already described. However, chromatin-remodeling complexes work in concert with chromatin modifying enzymes, which are represented in the second class (Narlikar et al., 2002). It contains those enzymes that add or remove covalent modifications on the freely accessible N-terminal histone tails.

3.2.1 Post-translational modifications of histone tails

Already in the 1960s, Vincent Allfrey (Allfrey, 1966) noted that histones from different eukaryotes were decorated with acetyl-, phospho- and methyl-groups. Although the physiological consequences of these modifications could not be fully understood at that time, successive research was building on his observations that certain modifications were associated with transcriptionally active chromatin sources.

Different classes of posttranslational modifications on histone tail residues have been identified up to date (Figure 3-1). Two categories can be distinguished. The first one is composed of small chemical groups that are covalently attached to specific amino acid residues. This includes the acetylation of lysines (K), methylation of lysines and arginines (K, R), serine/threonine phosphorylation (S, T) and ADP ribosylation of glutamic acid (E). The second category comprises the histone modification with relatively large polypeptides. Here, lysines can either be ubiquitylated or sumoylated by attachment of ubiquitin and sumo respectively (Kouzarides, 2007). Their functions are thought to be implicated in the aforementioned processes.

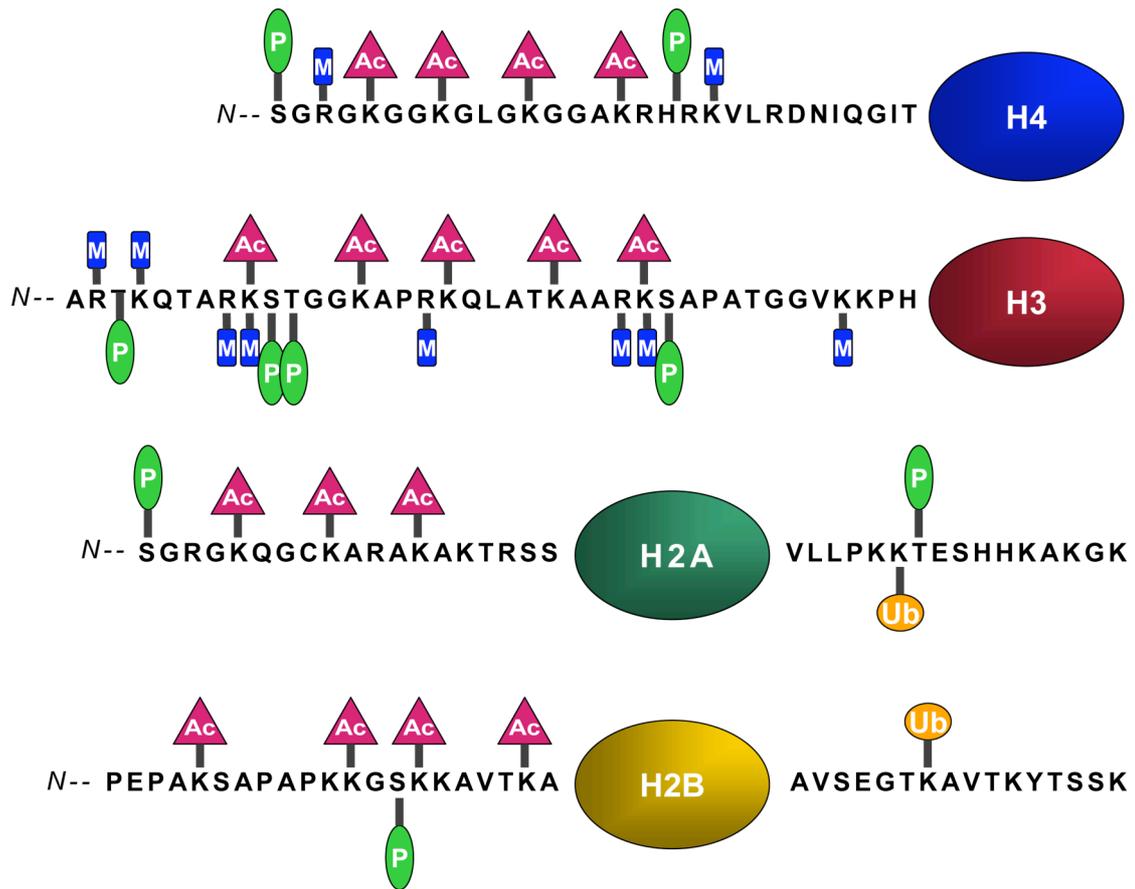


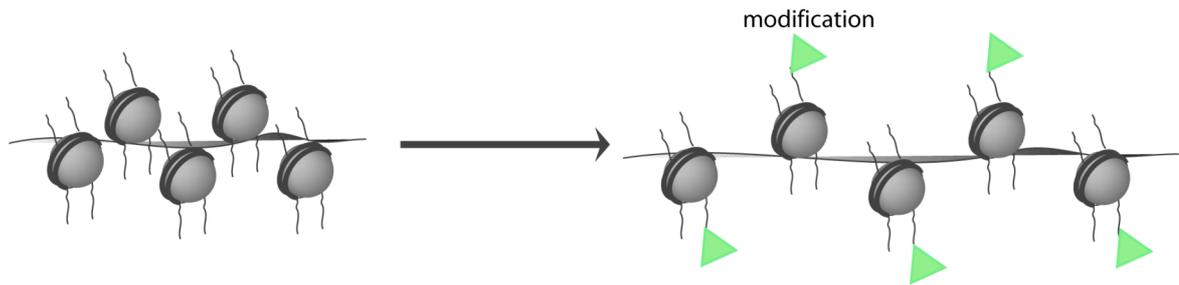
Figure 3-1: Post-translational modifications of histone tails from human cells

The four core histones and the identified post-translational are depicted.

Symbols: acetylation (Ac, magenta), methylation (M, blue), phosphorylation (P, green), ubiquitylation (Ub, orange)

The variety of modifications on the histone tails has been thought to constitute a coded message that results in a defined output – comparable with the triplet code that is used universally at the DNA level. The hypothesis was formulated as the so-called ‘histone code’ (Jenuwein et al., 2001; Strahl et al., 2000; Turner, 1993; Turner, 2000) and proposed that the different histone modifications result in distinct outcomes. The encoding of the message is thought to involve sequential or combinatorial action of the modifications. Two modes of code translation are conceivable (see Figure 3-2): either the modifications act directly on the interplay between DNA and histone proteins (cis-effects), thereby altering higher-order chromatin structure, or they specify interactions with downstream partners, such as chromatin-associated proteins (trans-effects). The latter can be further subdivided in the recruitment of positive-acting factors or the inhibition of binding of a negative-acting factor.

cis-effect (e.g. acetylation or phosphorylation)



trans-effect : specific binding of chromatin-associated factor to a modification mark (e.g. HP1 binding to methylated H3K9)



Figure 3-2: Changes to chromatin structure by *cis*- and *trans*-effects

An example for a *cis*-effect is acetylation which reduces the interactions between DNA and nucleosomes, thereby creating active chromatin domains and facilitating transcription. The ‘charge theory’ (Grunstein, 1997) is explaining the underlying mechanism for the mentioned chromatin relaxation: negatively charged acetylgroups attached to the positively charged amino groups of lysine residues in the tails of histones H3 and H4 are neutralizing the tail:DNA interaction. As an example, acetylation of H4K16 was shown to disrupt chromatin compaction and to influence protein-histone interactions (Shogren-Knaak et al., 2006). This of course might be a means to establish transcriptionally active euchromatic regions.

Applying the same principle of a *cis*-effect, the negative charge of phosphorylated serine residues is thought to create ‘charge patches’ and can be seen to result in condensed, non-permissive chromatin (Dou et al., 2000; Nowak et al., 2004).

Interestingly, most of the histone modifications identified to date could be classified as either activating or repressive marks.

A covalent modification with a variety of possible consequences is methylation. Methylation has been shown to result in both, activating and repressing readouts (Zhang et al., 2001). This is not only determined by the residue to be methylated but also by the methylation state, as lysines can be mono-, di- or trimethylated.

The picture gets even more complex by taking into account that mono- and dimethylation can occur also on arginines (Bauer et al., 2002; Chen et al., 1999; Strahl et al., 2001; Wang et al., 2001).

Some modifications seem to have orthologous localization in different genomes (e.g. *S. pombe*, *A. thaliana*, mammals) which in turn can be associated with distinct chromatin states. An example is the trimethylation of H3K4 which can be associated with activated promoter regions (Bernstein et al., 2005; Santos-Rosa et al., 2002). On the other hand, silenced heterochromatin can be correlated with methylation on histone H3 residues lysine 9 and lysine 27 (Litt et al., 2001; Ringrose et al., 2004).

Still, it is well possible that a considerable variation of histone modification patterns in different organisms exists and therefore the proposed histone code (Jenuwein et al., 2001; Strahl et al., 2000; Turner, 1993; Turner, 2000) - in contrast to the genetic code - is not likely to be universal. The highly dynamic properties of the chromatin modifications may function rather in a combinatorial way than acting singularly. This might be the basic message of the histone code hypothesis.

An alternative model to the histone code hypothesis, the signaling network model, compares the modification of histones with the well-studied process of signal transduction (Schreiber et al., 2002). Cytoplasmic signal transduction is mostly characterized by serine/threonine phosphorylation events in the cytoplasm. Histone post-translational modifications would similarly represent a nuclear signal transduction pathway that is DNA-associated. In this scenario, the modifications could create docking-sites for different effector-proteins that bind through various modification-recognition modules. The effect of this would be the formation of a signalling network with 'backup' properties (bistability) due to the redundancy of single modifications that could be bound by multiple effector proteins.

3.2.2 Post-translational modifications of histone cores

Apart from their tails, histones do not show any sequence similarity. Nevertheless, they assemble into the characteristic 'histone fold' and exhibit a common domain structure (Arents et al., 1991). For about 40 years, identification of modified residues focused mostly on the histone tails. Owing to the advent of mass spectrometry in this field, it has become evident that posttranslational histone

modifications can also be found on the globular core domains. The location of the modifications on the histone cores can be used for classification and function prediction (Mersfelder et al., 2006):

Outer face modifications could modulate histone–protein interactions thereby influencing higher order chromatin structure. They are involved in DNA repair, heterochromatic gene silencing and transcription (Mersfelder et al., 2006).

Modifications on the histone lateral surface appear to mediate histone–DNA interactions and act in processes like chromatin assembly, transcription and DNA damage repair (Cosgrove et al., 2004).

The residues that make up the histone–histone interfaces are also subject to modification. These core modifications are important for the regulation of nucleosome stability and could play a role in creating access to the DNA by loosening intranucleosomal interactions (Zhang et al., 2003).

3.2.3 Translating the message – Effector proteins and their specialized domains

The marks, deposited by the chromatin modifying enzymes can be subsequently interpreted by regulators of transcription, repair, replication, condensation and by other chromatin-associated factors.

The earlier mentioned trans-effects come about by the enzymatic modification of a histone tail residue (e.g. methylation of H3K9) which is followed by the association of chromatin-binding proteins (e.g. HP1, hetero-chromatin protein 1) with these sites. This results in downstream effects on chromatin structure.

The proteins that are able to recognize the post-translational modification pattern on the histone tails contain specialized protein domains that enable them to bind with high affinity and to generate a response according to the modification signal.

Figure 3-3 summarizes the enzymes that establish and remove small chemical modification marks and depicts some examples of protein domains that are involved in the recognition of the modifications presented.

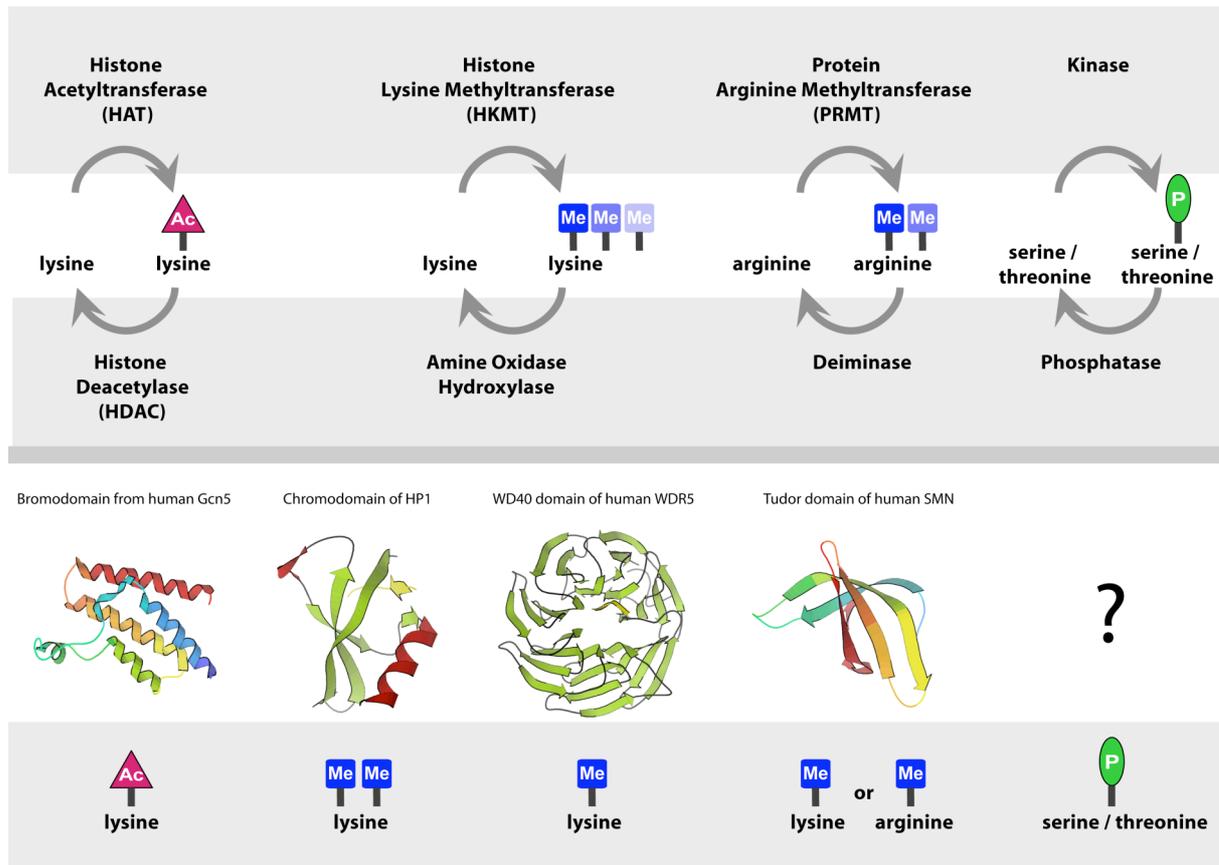


Figure 3-3: Different classes of chromatin-associated proteins regulate the modification, recognition and the turnover of post-translational histone modifications

Upper part: Covalent histone modifications and the histone-modifying enzymes involved (not depicted: ubiquitylation, sumoylation, ADP-ribosylation)

Lower part: Chromatin-binding domains from selected chromatin-associated proteins and the modifications they bind to.

An example for a modification-binding protein domain is the 40-50 amino acid-long chromodomain, a sequence motif that was originally identified in the *Drosophila* chromatin proteins, Polycomb (Pc) and Heterochromatin protein 1 (HP1). Functions that were assigned to the chromodomain include DNA-binding (Bouazoune et al., 2002) and the interaction with methylated lysine residues (Bernstein et al., 2006; Flanagan et al., 2005; Min et al., 2003). The HP1 protein was shown to bind via its chromodomain to methylated lysine 9 of histone H3 (Lachner et al., 2001) whereas Polycomb's (Pc) chromodomain binds H3K27 (Cao et al., 2002). The MOF chromodomain was initially shown to bind to roX2 RNA (Akhtar et al., 2000b). However, the domain was later classified as a chromo-barrel domain which is structurally distinct from the canonical chromodomains. This will be discussed in more detail in chapter 5.2.2.

A specialized domain involved in the binding of acetylated lysine residues is the bromo domain (Dhalluin et al., 1999; Zeng et al., 2002). Interestingly, bromodomain-containing proteins can be found amongst the HAT enzymes, such as Gcn5 and CBP/p300. They are often integral parts of large chromatin-associated complexes and contribute to their chromatin-binding capabilities, the Swi/Snf complex being such an example (Hassan et al., 2002). Binding-specificity to acetylated histones can be also found in other bromodomain-containing proteins, like Taf1 and Bdf1, that are part of the TFIID complex (Matangkasombut et al., 2003).

3.2.4 'Message disposal' – Enzymes that erase the modifications

The dynamic nature of histone modifications demands not only proteins that set a covalent modification but also mechanisms to remove the mark. Enzymes that are able to erase the message therefore counteract the activities of histone acetyltransferases (HATs), kinases, methyltransferases (MTs) and the other chromatin modifying enzymes. They are called histone deacetylases (HDACs), phosphatases and demethylases respectively (see Figure 3-3).

3.2.4.1 Histone deacetylases (HDACs)

Deacetylase enzymes can be grouped in two major families: the histone deacetylases (HDACs) and the sirtuins (*silent information regulator* like), which are NAD-dependent deacetylases (Grozingier et al., 2002).

The HDACs can be further subdivided into two groups according to their catalytic mechanism. These enzymes antagonize the action of acetyltransferases (not only HATs) by removing the acetyl groups from their substrate lysines.

The HDACs are mostly found to be part of large multisubunit complexes. Via these complexes the enzymes are targeted to genes where they lead to transcriptional repression. An example for this is the class I histone deacetylase Rpd3. When incorporated in a large complex including the HDAC Sin3 it can be targeted to regulatory DNA sequences (Kadosh et al., 1997; Kurdistani et al., 2003). However, Rpd3 can also be part of a small complex that is targeted to open-reading frames

and is thought to result in suppression of internal RNA polymerase II initiation (Carrozza et al., 2005).

The identification of modifying and opposing enzymes led to the widely-accepted model that, e.g. in the case of acetylation, transcriptional activators can recruit HATs to upstream activating sequences (UAS) and transcriptionally activate by local histone acetylation whereas URS-bound (upstream repressive sequences) repressors recruit HDACs for deacetylation of histones. An example for transcriptional repression associated with histone deacetylation comes from yeast, where the URS1 element binds the Ume6 repressor. This in turn recruits the Sin3-Rpd3 histone deacetylase complex (Kadosh et al., 1997).

This targeted system conveys the possibility to switch between transcriptional ON / OFF states by using sites of reversible acetylation. In addition, the interplay between HATs and HDACs also maintains the steady-state levels of global acetylation (Katan-Khaykovich et al., 2002).

3.2.4.2 Histone demethylases

The second class of enzymes that opposes chromatin-modifying activities are the demethylases. Although it was not possible for a long time to identify demethylases, research from recent years uncovered the existence of such enzymes. Like their counterparts, the methyltransferases, they have to be classified as enzymes acting on methylated arginine or on methylated lysine residues (Figure 3-3).

LSD1 (lysine-specific demethylase) was identified as an amine oxidase that could reverse mono- and dimethylated H3K4 (Shi et al., 2004).

Recently, it was shown that the JmjC domain-containing histone demethylase 1 (JHDM1) demethylates mono- and dimethylated H3K36 via an oxidative mechanism (Tsukada et al., 2006). Additional JmjC domain-containing proteins have been linked to demethylase activities. For example, it was demonstrated that the JMJD2A protein was able to reverse trimethylated H3K9 and H3K36 to a dimethylated product but was unable to further demethylate these lysine residues (Whetstine et al., 2006).

Demethylation of arginine follows another strategy. The PADI 4 (peptidyl arginine deiminase 4) protein was shown to be a deiminase, which converts arginine to citrulline thereby antagonizing arginine methylation. Furthermore, the enzyme is capable of directly deiminating mono-methylated, but not di-methylated, arginine (Cuthbert et al., 2004).

3.2.4.3 Phosphatases

Phosphatases, the enzymes that reverse phosphorylation marks, are existent but characterized to a much lesser extent (Holbert et al., 2005).

One example for the action of phosphatases in the context of histone-associated phosphorylation comes from studies on the modification of serine 10 of histone H3. It was shown that the H3S10-specific kinase Aurora-B was inhibited directly by the protein phosphatase 1 (PP1) which is involved in the control of histone phosphorylation (Hsu et al., 2000; Murnion et al., 2001).

3.3 For a change – Histone variants

The general view that octamers are made up by combination of the four canonical histones was massively extended when histone variants came into play.

The diversification of histones into variants offers further potential for chromatin differentiation and epigenetic regulation. These histone variants are used for replication-dependent and replication-independent replacement of canonical S-phase histones. While the basic histone fold is highly conserved in all histones the differences between canonical and variant histones range from a few amino acid substitutions to the addition of accessory domains, like the macro-domain in macroH2A (Kamakaka et al., 2005). The incorporation of histone variants in different regions of chromatin changes the nucleosome composition and can also establish specific domains of chromatin folding (Horn et al., 2002). Variants are known for all the histones, except H4. The reasons for the absence of H4 sequence variants are not known.

Up to date there is a growing list of histone variants and associated processes emerging. The following Table 3-1 gives an overview of some of the known histone variants in different species and their likely functions.

Histone variant	Species	Chromatin effect	Function
H10	Mouse	Chromatin condensation	Transcription repression
H5	Chicken	Chromatin condensation	Transcription repression
SpH1	Sea urchin	Chromatin condensation	Chromatin packaging
H1t	Mouse	Open chromatin	Histone exchange, recombination?
MacroH2A	Vertebrate	Condensed chromatin	X chromosome inactivation
H2ABbd	Vertebrate	Open chromatin	Transcription activation
H2A.X	Ubiquitous	Condensed chromatin	DNA repair/recombination/transcription repression
H2A.Z	Ubiquitous	Open/closed chromatin	Transcription activation/repression, chromosome segregation
SpH2B	Sea urchin	Chromatin condensation	Chromatin packaging
CenH3	Ubiquitous		Kinetochore formation/function
H3.3	Ubiquitous	Open chromatin	Transcription

Table 3-1: Histone variants and their implicated functions - adapted from (Kamakaka et al., 2005) – references can be found therein.

For example, the histone H3 variant CENP-A (centromere protein A) was found to be incorporated at centromeres where it replaces histone H3 and is involved in the formation of functional kinetochores (Ahmad et al., 2001; Earnshaw et al., 1985; Smith, 2002). It is possible that CENP-A is required for the maintenance of high compaction in the largely heterochromatic centromere region (Smith, 2002).

Some more examples include the H2A variant, H2A.X, which is involved in the repair of DNA double-strand breaks (Rogakou et al., 1998) and H2A.Z has a role in transcriptional regulation (Santisteban et al., 2000; Smith, 2002).

Yet another H2A variant, macroH2A, which is thought to act in transcriptional repression, localizes to the inactive X chromosome of female mammals (Costanzi et al., 1998). Its C-terminal macro-domain was suggested to be enzymatically involved in this transcriptional repression by epigenetically marking the inactive X chromosome (Ladurner, 2003).

The previously mentioned H3/H3.3 exchange at sites of active transcription in the HSP70 gene (Schwartz et al., 2005) is another example of how histone variants can impact on chromatin-associated processes.

Interestingly, the deposition of histone variants involves some of the known ATP-hydrolysis driven chromatin remodeling complexes which can act as exchangers.

In addition, this 'histone eviction' mechanism confer capabilities for the removal of modification marks from chromatin regions by simply replacing previously modified histones with unmodified histones or histone variants that bring about other features (Schwartz et al., 2005).

A mechanism for selective histone exchange which involves the concerted action of two distinct chromatin-remodeling enzymes was revealed by studies on the *Drosophila* Tip60 complex (Kusch et al., 2004). This complex contains both the histone acetyltransferase Tip60 and the ATPase Domino/p400 incorporated in one multiprotein complex. It was shown to catalyze the replacement of phosphorylated histone H2Av/H2B dimers with non-phosphorylated H2Av/H2B dimers at sites of double-stranded breaks.

4 Focus on histone acetyltransferases (HATs)

Histone acetylation is accomplished by a class of enzymes known as histone acetyltransferases (HATs). Chemically, they are catalysts for the transfer of an acetyl group from acetyl-CoA to the lysine ϵ -amino groups on the N-terminal tails of histones.

20 years after Vincent Allfrey initially observed the correlation between elevated levels of histone acetylation and increased levels of gene expression, an important part of the acetylation puzzle emerged from studies on *Tetrahymena*. Brownell and colleagues (Brownell et al., 1996) cloned and sequenced the *Tetrahymena* enzyme p55, a protein shown to be a histone acetyltransferase. Importantly, they could prove that the corresponding yeast homolog Gcn5p, a transcriptional coactivator, possessed HAT activity. This was the first nuclear histone acetyltransferase to be identified. Interestingly, the enzyme responsible for the removal of acetyl groups, Rpd3 - a histone deacetylase, was found nearly simultaneously (Taunton et al., 1996).

4.1 Classification of HATs

With the discovery of additional HATs (some of them already known as coactivators), a classification into distinct groups and families could be carried out. First of all, two groups of HATs can be distinguished according to their locality of action.

B-type HATs are localized to the cytoplasm and acetylate free core histones before they are imported in the nucleus and deposited in the chromatin structure. B-type enzymes primarily acetylate histones H3 and H4. In *Drosophila* and human cells it was shown that cytosolic B-type HATs preferentially diacetylate H4 at positions K5 and K12 (Sobel et al., 1995). This seems to be evolutionary conserved as the analogous residues (K4 and K11) in *Tetrahymena* carry the same acetylation pattern. Likewise, the HAT responsible for this acetylation seems to be conserved: Hat1p is the enzyme that catalyzes this reaction (Kleff et al., 1995). Newly synthesized histone H3 is also acetylated by B-type HATs in many organisms but this seems to be less conserved (Kuo et al., 1996; Sobel et al., 1995).

The second group, the **A-type HATs**, is represented by the nuclear histone acetyltransferases that acetylate histones in a chromatin context. Although HATs can modify lysine residues on all four core histones, they exhibit preferences for their substrates. Therefore, they can be further subdivided into three major families according to their catalytic domains.

The GNAT (Gcn5-related acetyltransferase) family is a large group of HAT enzymes (Gcn5, PCAF, Elp3, Hat1, Hpa2 and Nut1) (Brownell et al., 1996; Neuwald et al., 1997; Yang, 2004; Yang et al., 1996) that possess mainly histone H3 specificity.

A second HAT family, named after the founding members MOZ , YBF2 , SAS2 and IIP60), is the MYST family (Borrow et al., 1996; Takechi et al., 1999; Utley et al., 2003; Yamamoto et al., 1997; Yang, 2004). They usually show more specificity for histone H4. For example, the MYST family HAT *Drosophila* MOF is specific for H4K16 acetylation (Akhtar et al., 2000a; Hilfiker et al., 1997; Smith et al., 2000). For recent reviews on MYST HATs and their involvement in diseases consult (Avvakumov et al., 2007; Rea et al., 2007).

These two families of HATs are the predominant ones and structural studies have been performed on their catalytic domains (Clements et al., 1999; Lin et al., 1999; Rojas et al., 1999; Trievel et al., 1999).

A third family is composed of a divergent set of proteins that possess intrinsic HAT activity. CBP/p300 (Bannister et al., 1996; Ogryzko et al., 1996) and Taf1 (Mizzen et al., 1996) proteins are representative for that group. Nevertheless, the structure of their catalytic domains has not yet been solved.

4.1.1 Catalytic HAT-domains and their function

Histone acetyltransferases show sequence and structural homology to bacterial aminoglycoside N-acetyltransferases (Wolf et al., 1998). This feature and in particular the sharing of an invariant motif for acetyl-CoA recognition points to a common evolutionary origin of these and potentially related enzymes (Dutnall et al., 1998; Neuwald et al., 1997; Wolf et al., 1998).

The catalytic core region is comprised of two substructures. The first one is an antiparallel β -sheet composed of three β -strands. This connects via a loop to the second part, an α -helix followed by another β -strand, which pairs with the β -sheet (Tan, 2001). The acetyl-CoA and the histone tail substrate are envisaged to be accommodated between these domains.

Despite the structural homology of histone acetyltransferases, different catalytic mechanisms seem to be employed by them. The most prominent one is the formation of a ternary complex that includes the HAT, acetyl-CoA and the histone substrate. Using a conserved residue in the catalytic domain of the HAT, a nucleophilic attack of the substrate is directed towards the acetyl-CoA. This has been shown to be the case for Gcn5, PCAF and the MYST family member Esa1 (Lau et al., 2000; Tanner et al., 2000).

In contrast, p300 HAT seems to catalyze the reaction in a ping-pong kinetic pathway, where the formation of a covalent intermediate between acetyl-CoA and a conserved cysteine residue precedes the transfer of the acetyl group to the target lysine (Thompson et al., 2001).

4.2 HATs – specific non-soloists

HATs are most often incorporated into large multi-protein complexes that have a modular structure (Kimura et al., 2005). Reflecting the diversity of the HAT enzymes themselves, these complexes can also vary in their protein composition to contribute to the unique features of each complex. The substrate specificity of such HAT complexes is brought about by the subunit composition and the domains that are present in addition to the catalytic domain. For example, some subunits have domains that cooperate to recruit the HAT to the appropriate location in the genome; these include bromodomains, chromodomains, WD40 repeats, Tudor domains and PHD fingers.

A profound effect on specificity can be seen upon the choice of substrate. The fact that HATs often show decreased specificity *in vitro* can be influenced experimentally by providing the most natural substrate. Histone tail peptides, free octamers and nucleosomes as substrates generally show increasing HAT specificity in this order.

The need for associating factors in order to create optimal HAT activity and specificity is nicely illustrated in the publication of Morales et al (Morales et al., 2004) where they demonstrated that either MSL-1 or MSL-3 alone were not sufficient to achieve maximal HAT activity of the MOF enzyme but the full trimer was able to do so. Another example was contributed by work on yeast Gcn5. It was shown that recombinant Gcn5 protein alone was not able to acetylate nucleosomal histones efficiently, whereas the Gcn5-containing ADA and SAGA complexes could do it (Grant et al., 1997). Ada2 and Ada3 were identified as the shared components to modulate catalytic activity and specificity in these two complexes (Balasubramanian et al., 2002).

4.3 Biological relevance of HATs

The congruence of previously identified transcriptional co-activators with their later identification as histone acetyltransferases (such as Gcn5 and p300) put forth the idea of them being responsible for the local regulation of specific genes.

However, the biological functions of histone acetyltransferases expand beyond gene-specific regulation as consistent levels of acetylation can be detected

genome-wide in cells. The distinction between localized and broad-range acetylation can be achieved in different ways.

4.3.1 Balancing acetylation – A delicate matter

The global acetylation pattern observed is most likely the result of balancing acetylation and deacetylation activities in order to prevent full acetylation as well as fatal hypoacetylation (Vogelauer et al., 2000). The fast turnover of acetyl groups ensures the reversal to the basal state after the signal is removed (Katan-Khaykovich et al., 2002). All of this serves the function of an adaptable and temporal fine-tuning of gene expression.

An important mechanism to not only gain enzymatic specificity but at the same time inhibit chromosome-wide silencing by promiscuous chromatin-associated factors was suggested by van Leeuwen and Gottschling (van Leeuwen et al., 2002). This *trans*-effect, illustrated in Figure 4-1, would act through exclusion of these negative-acting factors by substantial pre-modification of target residues.

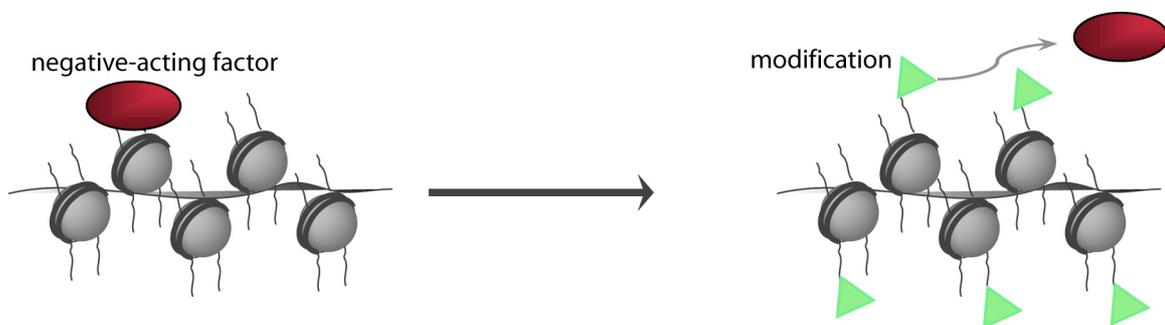


Figure 4-1: trans-effect involving negative-acting factor

Regulation in this manner was proposed for the interplay between acetylation - resulting in an 'open' chromatin configuration - and telomeric heterochromatin formation in yeast. The silencing information regulator Sir3 can interact with the H4 tail (residues 16-29) only when H4K16 is deacetylated (Kurdistani et al., 2003). *In vivo* observations confirmed the various experimental data on this, as lysine 16 of histone H4 is generally found to be in a deacetylated state in Sir3 binding regions (Braunstein et al., 1996).

Looking at modifications other than acetylation, an example for such an inhibitory mechanism was delivered by the observation that two adjacent post-translational modifications, phosphorylation of H3S10 and methylation of H3K9, can have a combinatorial readout. The interaction of HP1 protein with methylated H3K9 is abrogated upon H3S10 phosphorylation (Fischle et al., 2005; Hirota et al., 2005), which is a mark on mitotic chromosomes (similar to Figure 4-1). This finding is consistent with the experimental inhibition of mitotic Aurora B kinase resulting in retention of HP1 on mitotic chromosomes (Fischle et al., 2005; Hirota et al., 2005).

4.3.2 Global versus gene-specific histone acetylation

The following section will elaborate further on genome-wide versus gene-specific histone acetylation. A difference between global and local acetylation is characterized by the idea that the first one is mediated by non-specific interaction of histone acetyltransferases with chromatin whereas the latter one is regulated by sequence-specific DNA-binding proteins. Keeping in mind the genome-wide background levels of acetylation, gene-specific acetylation can be seen as local perturbations of acetylation states. This is mostly the result of interplay between HATs and HDACs. Both classes of enzymes can be targeted to specific sites in the genome.

Some HAT enzymes belonging to the MYST family have been linked to regulation of chromosome-wide gene expression (Kimura et al., 2005). Yeast Sas2 and Sas3 enzymes have been implicated in long-range gene repression dependent on chromosomal location (Ehrenhofer-Murray et al., 1997) while *Drosophila* MOF is involved in hyperactivation of the male X chromosome (Hilfiker et al., 1997). This prominent example of global acetylation is associated with dosage compensation in *Drosophila*. Here, a distinct chromatin domain, the male X chromosome, is denoted by specific H4K16 acetylation which correlates with roughly two-fold upregulation of gene expression (Bone et al., 1994). How this functions in detail is still elusive.

The localized effect of acetylation has become directly visible from studies on yeast. Chromatin immunoprecipitation (ChIP) studies have shown that the HDAC

Rpd3, which is part of a large multiprotein complex, is enriched at the INO1 promoter and deacetylates nearly all acetylation sites on histones H2A, H2B, H3 and H4. The two nucleosomes that are adjacent to the URS1 element are affected by this highly localized deacetylation (Kurdistani et al., 2003). It is thought that the binding of chromatin-remodelers and the TATA-binding protein (TBP) to this chromatin region is destabilized and hence transcription repressed (Deckert et al., 2002). This is consistent with the suggested repressive effect of Rpd3 on INO1, a gene involved in inositol biosynthesis.

Targeting to promoters via associated DNA-binding proteins is a mechanism that is also employed by HATs. An example is the HAT Gcn5, a member of the SAGA multiprotein complex. It is targeted to the promoter of the yeast HIS3 gene by the transcriptional activator Gcn4 (Kuo et al., 2000). The result is hyperacetylation of nearby H3 histones.

In addition to setting histone modifications, HAT complexes can acetylate other non-histone proteins, such as transcription factors (Gu et al., 1997), cytoskeletal proteins, molecular chaperones and nuclear import factors (Glozak et al., 2005).

5 Dosage Compensation – a model for the regulation of chromatin

Animal species have evolved elaborate systems to accomplish sex determination. This is not only reflected in the phenotypic appearance but is based on differences at the molecular level – more precisely the distinct regulation of chromatin.

The sexual dimorphism of the genders is linked to the presence of specialized chromosomes that diverged during evolution. Normally, the sexes differ in the number and type of sex-determining chromosomes. For example, in mammals the female possesses two X chromosomes whereas the male counterpart has one X and one Y chromosome. The reverse setup is realized in birds, in which the males show the homozygous state (ZZ) and the females are heterozygous (ZW) for the sex chromosomes. However, a common problem that all these species have to tackle is linked to the divergence of sex chromosomes: unequal distribution of chromosomes - or aneuploidy. This is the case in the heterozygous scenario

mentioned above. Without a mechanism to adjust expression levels, aneuploidy generally leads to developmental abnormalities and sex-specific lethality.

5.1 Different Species – Different Mechanisms

Therefore, species have independently evolved different mechanisms to ensure equal levels of sex chromosome-linked gene products in males and females. These processes are brought together in the term dosage compensation. Two common features are inherent to all these mechanisms: Dosage compensation takes place only in one of the sexes and therefore needs to reliably distinguish between the two karyotypes. Secondly, the compensation machinery has to recognize the chromosomes as it works only on one (or a set of two) sex-chromosomes but not on the autosomes.

Three different strategies thereof are presented briefly (compare Figure 5-1).

In humans, the females (XX) inactivate one X chromosome early in embryogenesis, resulting in a heterochromatic Barr body, which is genetically inactive. The X that is chosen for inactivation undergoes a series of events. The non-coding XIST RNA is hypertranscribed and coats the X chromosome from which it initiates. This is followed by silencing of the genes along this chromosome, marked by hypoacetylation, increased histone lysine methylation, H2A exchange to macroH2A and other changes to the chromatin. The result is a condensed and silenced X chromosome. This is reviewed in (Avner et al., 2001; Chow et al., 2005).

In contrast, male *Drosophila* flies (XY) double the expression of genes along the X chromosome leading to the same amount of X-linked gene products compared to the females. This will be described in more detail in the next chapter.

Yet, another strategy is applied by the nematode worm *C. elegans*. Hermaphrodites (XX) partially repress both X chromosomes thereby bringing the expression to the same level as in males (X0). This is achieved by a protein complex that is similar to the 13S condensin complex which in turn is involved in chromosome compaction during mitotic segregation. For a review see (Meyer, 2000). Therefore, it is likely that the dosage compensation complex in *C.elegans* is responsible for partial condensation of the X chromosomes and subsequent repression of X-linked genes.

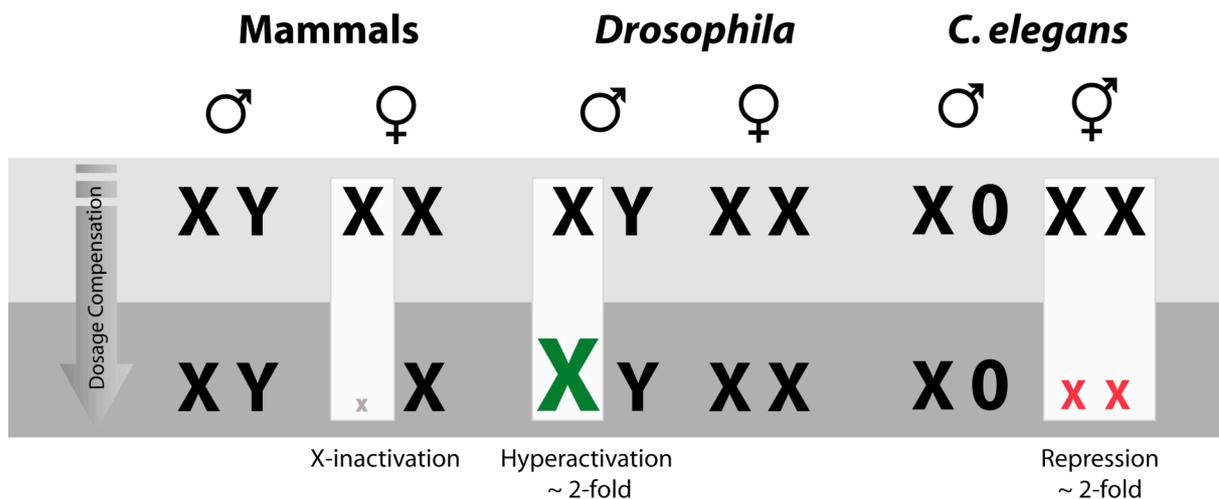


Figure 5-1: Dosage compensation mechanisms

The different strategies that are applied by mammals, *Drosophila* and *C.elegans* are shown

Although all of these mechanisms have divergent evolutionary origins, they result in balancing the relative gene expression between males and females.

5.2 A closer look at dosage compensation in *Drosophila*

The phenomenon of dosage compensation has been the subject of intensive research over the last decades and it has proven to be a very valuable model system for the epigenetic regulation of chromatin. There has been enormous progress in understanding which factors are the major players in this process. It turned out that the machineries involved in compensation are distinct but they all lead to the modulation of gene expression by altering chromatin structure.

One of the best studied model organisms in this respect is the fruit-fly *Drosophila melanogaster*. Genetic screens for sex-specific lethality in *Drosophila* have brought to light the factors responsible for sex determination and dosage compensation (Bashaw et al., 1995; Belote et al., 1980; Hilfiker et al., 1997; Palmer et al., 1993). By analysing these loss-of-function phenotypes it became clear that dosage compensation in the fly is achieved via the male organism. This is revealed by the lethality of males that are mutant for any of the genes identified. Therefore the genes were collectively named the male-specific lethals (MSLs). They code for five essential proteins that comprise Male-specific lethal -1 (MSL-1), MSL-2, MSL-3, Maleless (MLE) and Males absent on the first (MOF) (Bashaw et

al., 1995; Gorman et al., 1995; Hilfiker et al., 1997; Kuroda et al., 1991; Palmer et al., 1993). Together with the two non-coding RNAs, roX1 (RNA on the X 1) and roX2 (Amrein et al., 1997; Meller et al., 1997), they form the dosage compensation complex (DCC), alternatively called the MSL complex. This complex was also referred to as the compensasome (Franke et al., 1999).

The key to the DCC's exclusive action in the male fly lies in one of its constituents, the MSL-2 protein. It is under the control of the master sex-determining protein, Sex-lethal (Sxl). Sxl is only present in females and inhibits the translation of MSL-2 mRNA (Bashaw et al., 1997; Kelley et al., 1997). Therefore, the MSL-2 protein can only be found in male fruit flies. Additionally, the MSL-2 protein is required for the stabilization of MSL-1 by direct interaction thus eliminating another component of the MSL complex in females.

Apart from the core histone acetyltransferase MOF, the DCC contains at least one more enzyme. The MLE (maleless) protein is a RNA-DNA helicase (Kuroda et al., 1991). It possesses a helicase domain and two dsRNA-binding domains. The fact that MLE's helicase activity is essential for the stability of roX RNAs (Gu et al., 2000) points to its speculated function to integrate the roX RNAs into the DCC. However, dosage compensation does not seem to be the only process to involve MLE's action as it is also involved in mRNA splicing (Reenan et al., 2000).

Another interesting feature of the MSL complex is the incorporation of the two non-coding RNAs roX1 and roX2. Despite their largely differing size (roX1 ~3,6 kb, roX2~0,7 kb) and very low sequence similarity they are functional redundant. They do not only physically associate with the complex but they are also dependent on the MSL complex for stability (Franke et al., 1999; Meller et al., 2000). Intriguingly, the roX genes themselves provide nucleation sites for MSL complex assembly and can therefore contribute in DCC targeting to the X (Kageyama et al., 2001).

5.2.1 The DCC – confined to the X

The MSL complex assembles solely on the male X chromosome and does not localize to the autosomes. This can be beautifully visualized by using antibodies against the MSL proteins and staining polytene chromosomes from male *Drosophila* larvae (Figure 5-2). In these chromosome squashes hundreds of X-

chromosomal bands are lighting up at the positions where the DCC is localized, thereby painting the whole X chromosome.

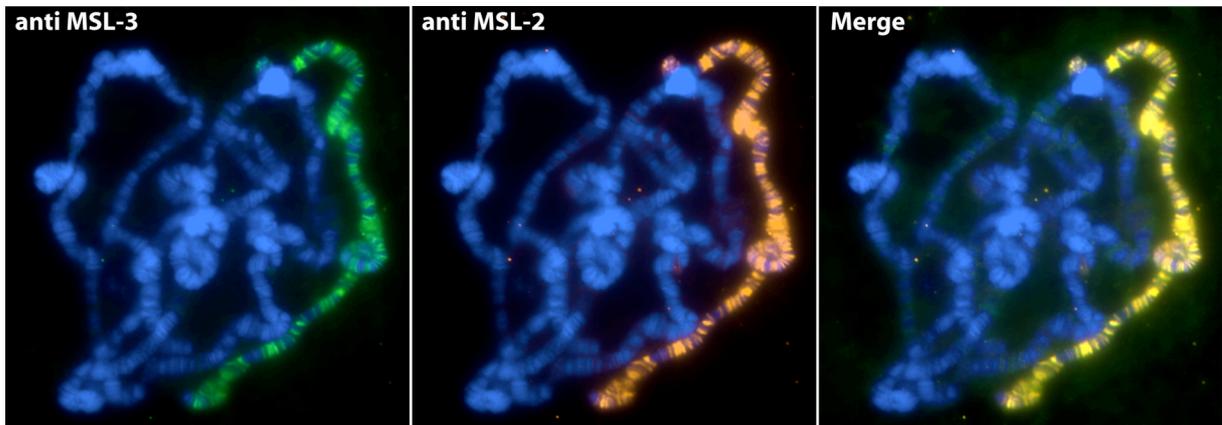


Figure 5-2: MSL localization on *Drosophila* polytene chromosomes.

MSL-3 (green) and MSL-2 (red) co-localize on the male X chromosome. DNA is counterstained with Hoechst (blue).

But how is this targeting of the male X chromosome by the MSL complex achieved? It is clear that this specific localization needs a mechanism for the reliable distinction of the X chromosome from autosomes.

Many studies have focussed on the targeting of the MSL complex to the male X chromosome. Earlier hypotheses proposed a targeting mechanism via specific DNA sequence elements that can be recognized by the dosage compensation complex. However, it was not possible to identify simple consensus DNA recognition elements that would make the X chromosome easily distinguishable from autosomes and allow specific binding of MSLs. The prediction of further binding sites could not be achieved by the analysis of three identified binding sites - two of them in the roX1 and roX2 genes, another at polytene band 18D10 (Kelley et al., 1999; Oh et al., 2004). A number of new DCC binding fragments were recently identified by the usage of chromatin immunoprecipitation but they yielded only short and degenerate sequences in bioinformatic analyses (Dahlsveen et al., 2006).

Another model proposed that there are around 35-40 so called chromatin entry sites (or high affinity sites) existing that are responsible for the initial targeting of the complex followed by further in *cis* spreading to coat the X chromosome (Kelley et al., 1999; Lyman et al., 1997).

However, Fagegaltier and colleagues (Fagegaltier et al., 2004) could show in elegant experiments that X-chromosomal regions without an entry site were still able to attract the MSL complexes when translocated to an autosome. Furthermore, they observed that spreading was not happening from X-chromosomal regions to autosomal material. Additionally, the translocation of the 18D10 high-affinity sites to autosomes did not necessarily result in spreading (Oh et al., 2004).

Therefore, it was proposed that the targeting is probably not only due to these chromatin entry sites and the spreading could occur via hierarchical affinities of individual binding sites (Dahlsveen et al., 2006; Fagegaltier et al., 2004). It is very likely that secondary binding sites might be characterized by other features, like actively transcribed sequences or transcription-associated epigenetic marks.

Recently performed studies on the question of MSL complex targeting to X-linked genes and its involvement in dosage compensation aimed at gaining a high resolution picture of MSL localization (Alekseyenko et al., 2006; Gilfillan et al., 2006; Legube et al., 2006). They therefore combined chromatin immunoprecipitations with DNA array analysis (ChIP-chip) to map the DCC binding sites along the X chromosome. The studies revealed that the DCC does not coat the whole X chromosome but rather binds to genes and not intergenic sequences. Binding was found to be enriched at the 3' end of active genes (Alekseyenko et al., 2006; Gilfillan et al., 2006). A number of degenerate sequence motifs were recently found to bear affinity for the DCC but proved to have noticeable, yet limited, potential for DCC binding prediction (Dahlsveen et al., 2006; Gilfillan et al., 2006). An interesting correlation could be drawn between promoter motifs of DCC target and non-target genes. A significantly higher incidence of a DNA replication element factor (DREF)-binding motif suggested an involvement of this sequence motif in dosage compensation (Legube et al., 2006). A very recent study (Kind et al., 2007) has provided evidence for a model that the MSL complex is targeted to transcriptionally active X-linked genes which contain a certain combination of small target sequences in their transcribed region. Importantly, these sequences alone are not sufficient for recruiting the MSL complex. Only in the context of the gene itself the sequences are revealed as

targets by the passage of RNA polymerase II through the gene and can attract the MSL complex (Kind et al., 2007).

5.2.2 MOF and the 2x X

Localization of the MSL-complex results in a hyperacetylation of the X chromosome, which is marked by specific acetylation of lysine 16 on histone H4 (H4K16) (Turner et al., 1992). The component responsible for this specific acetylation is the histone acetyltransferase MOF (males absent on the first). This 827 amino acid long (91 kDa) enzyme belongs to the MYST family of HATs (see above) and was shown to be H4K16-specific *in vitro* (Akhtar et al., 2000a; Smith et al., 2000) and *in vivo* (Mendjan et al., 2006; Smith et al., 2000). This mark is enriched on X-linked chromatin of the male fruit-fly and has been implied to be vital for the dosage compensation process. MOF contains domains that are clearly involved in chromatin regulation. Besides the HAT domain, MOF contains a chromo-barrel domain (CBD) and a Zinc-finger region.

The chromo-like chromo-barrel domain is distinct in its structure from the canonical chromodomains that were discussed before (Nielsen et al., 2005). The MOF CBD structure is composed of β -barrels, while the canonical chromodomains have an alpha/beta fold. A similar chromo-barrel domain can be found in the MSL-3 protein (Buscaino et al., 2006).

Unlike the HP1 chromodomain which can bind to methylated histone H3 lysine 9 (H3K9) binding of methylated residues is prevented in the MOF chromo-barrel domain because the three critical aromatic residues are not conserved (Nielsen et al., 2005). However, the CBD of MOF and its adjacent lysine-rich region was shown to bind to roX2 RNA *in vitro* and point mutations in the MOF chromo-related domain severely affected the interaction of MOF with RNA (Akhtar et al., 2000b). Structural studies revealed that the chromo barrel domain is necessary, but not sufficient, for the interaction of MOF with RNA (Nielsen et al., 2005).

Although, one could think of other HATs in *Drosophila* that can catalyze H4K16 acetylation, it is the MOF enzyme that - incorporated in the MSL complex – exerts its function on the X chromosome.

Inactivation of the acetyl transferase activity of MOF or complete depletion of MOF leads to the absence of H4K16 acetylation on the X chromosome and lethality at the larval stage (Akhtar et al., 2000a; Hilfiker et al., 1997; Smith et al., 2000). This classifies MOF as the essential enzymatic component of the MSL complex responsible for H4K16 acetylation. In *Drosophila*, specific acetylation of H4K16 on the male X chromosome correlates with approximately 2-fold transcriptional upregulation (Gu et al., 1998). How this is achieved in detail is still enigmatic. However, acetylating histones might not be MOF's sole function as it has been demonstrated that MOF can acetylate the MSL-1 and MSL-3 proteins as well (Buscaino et al., 2003; Morales et al., 2004).

6 The MSL complex and beyond

The biochemical purification of the MSL complex was recently performed in our lab and allowed for the first time the identification of the stably associating components of the MSL complex. We have shown that purified *Drosophila* dosage compensation complex, affinity-purified from *Drosophila* embryonic nuclear extract and the *Drosophila*-derived SF4 cell-line, not only contained the already known MSL proteins but also other proteins. Unexpectedly, nuclear pore components and associated proteins (Mtor, Nup153, Nup160, Nup98, and Nup154) were observed to co-purify with the MOF protein. Furthermore, depletion of the nuclear pore associated proteins Mtor and Nup153 by RNA interference manifested in a loss of MSL localization to the male X chromosome. Dosage compensation of a subset of X-linked genes was impaired as well (Mendjan et al., 2006).

Interestingly, apart from the nucleoporins, we found other novel proteins, which we termed the NSL (non-specific lethal) proteins. Underlining the significance of this finding, the mammalian orthologs of most of these proteins were also identified as co-purifying factors with the human MOF protein (Mendjan et al., 2006).

6.1 *Drosophila* MSL-1 and NSL1 – similar players

MSL proteins have previously been shown to increase the enzymatic activity of MOF *in vitro* (Morales et al., 2004). In particular, the MSL-1 protein exhibits a key role in coupling MOF and MSL-3. This interaction, brought about by the C-terminal domain of MSL-1, therefore contributes actively to the assembly of a stable core complex.

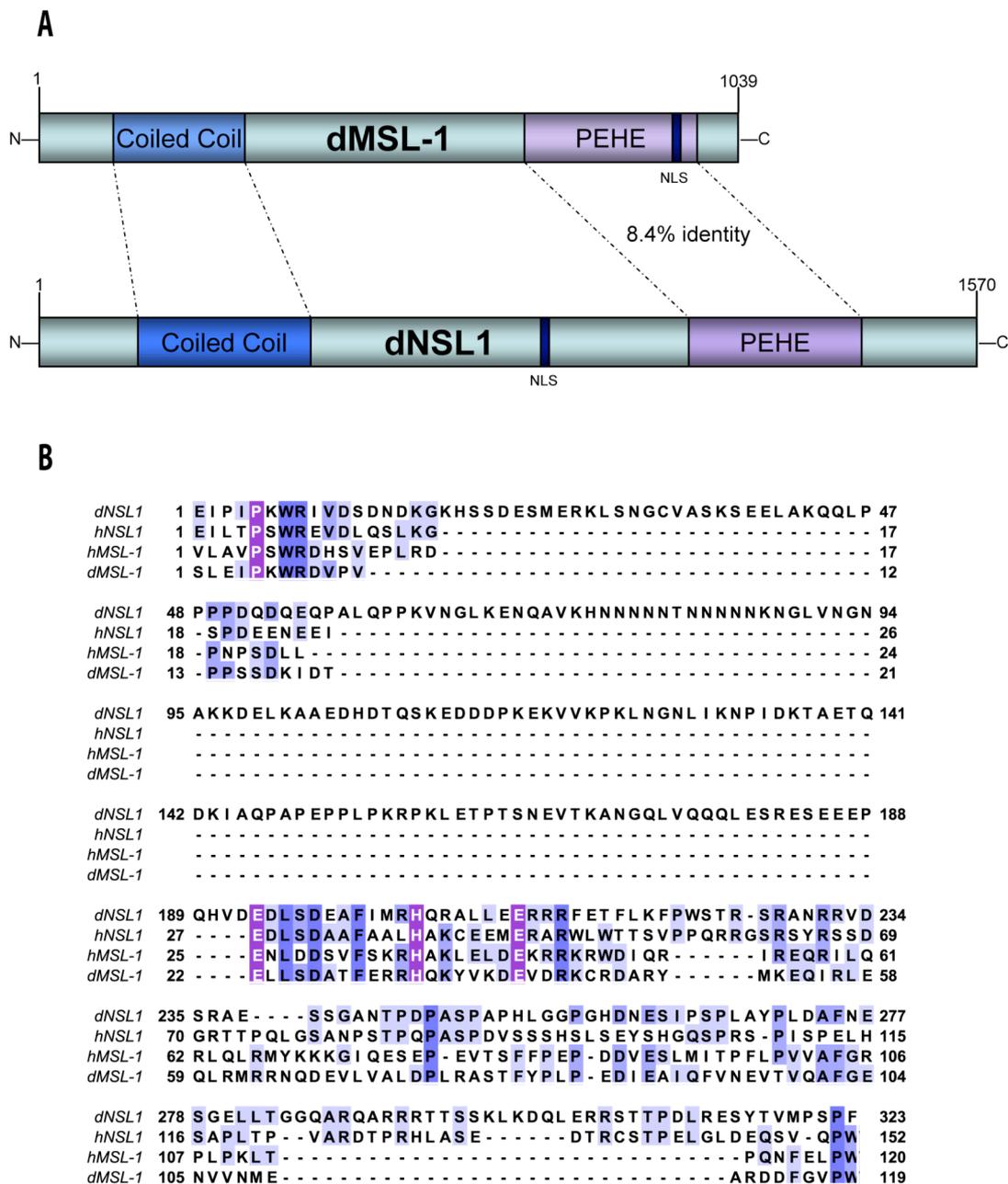


Figure 6-1: MSL-1 and NSL1 PEHE domains

A) Domain architecture of *Drosophila* MSL-1 and NSL1 proteins

B) Multiple sequence alignment of *Drosophila* and human MSL-1 - and NSL1 PEHE domains; conserved residues are shaded in purple, sequence identity in blue.

Interestingly, this interaction is mediated by the PEHE domain (Morales et al., 2004), a conserved amino acid motif (Marin, 2003) that is only present in MSL-1 and one of the newly identified proteins, NSL1 (compare Figure 6-1 B). These proteins are the only known PEHE domain containing proteins in *Drosophila* (Marin, 2003).

A few words to describe NSL1 (non-specific lethal 1), which is a 1570 amino acid (173 kDa) protein. *Drosophila* NSL1, an evolutionary conserved protein, shares several regions of homology with its orthologue in humans (KIAA1267, hNSL1). The predicted domain architecture consists of a putative N-terminal coiled-coil domain and the C-terminal PEHE domain. This arrangement is very similar to the *Drosophila* MSL-1 protein (Figure 6-1 A). As described above, the MSL-1 PEHE domain is a MOF-interacting motif.

AIMS AND OBJECTIVES

The main aim of my PhD thesis was the biochemical characterisation of NSL1 containing complexes. To address this question, I was using an affinity purification strategy to purify NSL1, with the goal to identify stable interaction partners by mass spectrometric analysis. Furthermore, the enzymatic activity and substrate specificity of the isolated NSL complexes was investigated.

In order to elucidate the role of the NSL1 protein *in vivo* and *in vitro*, its involvement in histone acetylation was studied. RNA interference strategy in male and female *Drosophila* cell lines was applied to test the effect of MOF, MSL-1 and NSL1 depletion on the global histone H4 lysine acetylation.

MATERIALS AND METHODS

7 Cloning of full-length NSL1 and NSL3 cDNAs

7.1 NSL1

As the available coding sequence for NSL1 (CG4699) was incomplete I decided to clone the full-length cDNA from adult flies. Therefore, total poly-adenylated mRNA was isolated from wild-type adults of *Drosophila melanogaster* by using the Dynabeads Oligo (dT)25 system (Dynal). mRNA was reverse-transcribed with the SuperScript II RT Kit (Invitrogen). This cDNA pool served as template for PCR-amplification (PCR Master, Roche) of the missing parts. The 5'-end of the CG4699 cDNA (bp1-1976) was synthesized with the primers 4699Ncolfwd + 4699Avallrev which added restriction sites for *NcoI* and *Avall* to the fragment. This fragment could be ligated to the 1.8 kbp *Avall*- and *XhoI*-cut fragment (bp1976-3838, 4699Avallfwd + 4699Xhorev primers). The 3'-end was similarly generated by a PCR reaction (primers 4699Xholfwd + 4699N2revNot) spanning base-pairs 3838 to 4738. The added *XhoI* site allowed ligation to the previous fragment. The *NcoI*- and *NotI*-sites at the ends of the resulting cDNA were used to clone the full-length cDNA into the vector pFastBac-HT(c) (Invitrogen). Multiple rounds of sequencing and elimination of point mutations (QuikChange Site-Directed Mutagenesis Kit, Invitrogen) yielded the correct coding sequence for NSL1.

Primers for NSL1 cloning:

4699Ncolfwd	5'-TTGCCATGGCCCCAGCGCTCACAG-3'
4699Avallrev	5'- CTGGGAGCAGAGCCAGGAC -3'
4699Avallfwd	5'- CGCTGGTCCTGGCTCTGCT -3'
4699Xholrev	5'- GCTCCTCGAGAAGAGCTCG -3'
4699Xholfwd	5'- GCTCTTCTCGAGGAGCGAC -3'
4699N2revNot	5'- GTTGCGGCCGCTTAGATGCGTCTGCTGCGAAC -3'

7.2 NSL3

The full-length cDNA for NSL3 was obtained from the BDGP Drosophila cDNA library. The 5'-end was taken from the construct that was used for NSL3 antibody production. This *NcoI* / *SphI* digested fragment was ligated together with the 3' *SphI* / *EcoRV* (blunt) fragment that was taken from the original pOT vector. The full coding sequence for NSL3 was cloned in the *NcoI* / *BamHI* (blunt) digested pFastBac-HT(c) (Invitrogen) vector.

7.3 Sequence alignments

Sequences of orthologous proteins were obtained from ENSEMBL database (www.ensembl.org) and aligned with MUSCLE 3.6 (Edgar, 2004). Alignments were further edited with the Jalview alignment editor software (Clamp et al., 2004).

8 Generation of NSL1- and NSL3-specific antibodies

8.1 Expression and purification of NSL antigens

Fragments of respective proteins were cloned into a modified pET9-derived expression vector (Gunter Stier, EMBL) and expressed as double-tagged glutathione S-transferase / 6xhistidine (GST-His₆) fusion proteins in *E.coli* BL21(DE)3 cells. Soluble recombinant proteins were purified to homogeneity via the His₆-tag on Ni-NTA resin (Qiagen), eluted with imidazole and dialyzed against PBS.

8.1.1 Protein expression

NSL antigens were expressed from *E.coli* strain BL21(DE3)pLysS that was transformed with the respective GST-His₆-NSL expression plasmids.

Overnight pre-cultures were used to inoculate the main cultures (LB broth + 20µg/ml Kanamycin). The volume for the main cultures was ranging from 0.5 to 2 liters depending on the protein that was expressed. The cultures were grown on a shaker at 37°C until they reached an OD₆₀₀ of around 0.4 - 0.6. Cultures were then transferred to 18°C and induced with 0.2 mM Isopropyl-β-D-thiogalactopyranosid (IPTG). Cells were left shaking for protein expression at 18°C overnight. Harvested bacterial pellets were shock-frozen in liquid nitrogen and stored at -80°C.

8.1.2 Affinity purification of GST-His₆-NSL fusion proteins

Cells were lysed by thawing the shock-frozen bacterial pellets and resuspending them in ice-cold lysis buffer (20 mM Tris pH 8.0, 10 mM Imidazole pH 8.0, 150 mM NaCl, 0.2% NP-40, 2 mM β-Mercaptoethanol, 1mM PMSF). For efficient lysis lysozyme (0.5mg/ml) and DNaseI were added and allowed to incubate for 5-10min. Sonication (Sorvall Omnimixer: 6x 30sec, 40% power) completed the cell lysis. Cell debris was separated from protein-containing supernatant by a centrifugation step in a cooled SS34 rotor of a Sorvall RC6 centrifuge (12.000rpm, 40 min., 4°C).

Purification of GST-His₆-tagged fusion proteins was done on preparative columns (Econo-Column, BioRad) equipped with Ni-NTA resin (Qiagen). This gravity-flow protocol allowed for a fast and efficient recovery of purified protein. All steps were performed at 4°C to preserve protein integrity. The affinity resin was equilibrated with >10 column volumes (CV) of lysis buffer. The extract was decanted on the column and the flow-through reloaded two times to ensure maximum binding of recombinant proteins. The column was washed according to the following scheme: 10 CV lysis buffer, 10 CV wash buffer (20mM Tris pH 8.0, 10mM Imidazole pH 8.0, 150mM NaCl, 2mM β-Mercaptoethanol, 1mM PMSF), 10 CV wash buffer (with 1M NaCl) and 10 CV wash buffer (with 25mM Imidazole pH8.0).

Proteins were eluted with 5 CV of elution buffer (20mM Tris pH 8.0, 330mM Imidazole pH 8.0, 150mM NaCl, 10% glycerol, 2mM β-Mercaptoethanol, 1mM PMSF). The whole purification procedure was tracked by sample analysis on Coomassie-stained SDS protein gels.

Before injection, elution fractions were dialysed against PBS.

8.2 Antigens

A C-terminal fragment of NSL1 (aa1019–1287, PEHE domain) was used to immunize 3 rats and 2 rabbits.

To generate a NSL1-specific antibody directed towards another epitope of the protein a 240amino-acid long polypeptide (no predicted domain) from the very N-terminus of NSL1 was expressed and purified the same way and injected in two more rabbits.

Similarly a N-terminal fragment of NSL3 (aa100-590, putative hydrolase domain) was injected in rats for antibody production.

Pre-sera and sera obtained after immunization were monitored for specificity by western blot analysis.

8.3 NSL1 and NSL3 antibodies

Below is a table that summarizes the antibodies that were generated and the dilutions that were used for the respective applications.

	animal	name	bleed	dilution 1:x		
				Western	IF	IP (μl)
αNSL1	rat	rat1	1	500	500	5
αNSL1	rabbit	FIN5	final	500	-	5
αNSL1	rabbit	3BUM	final	1000	50	-
αNSL3	rat	rat3	final	500	-	5
αNSL3	rat	rat3, purified	final	-	500	-

* IF (immunofluorescence), IP (immunoprecipitation)

9 Coimmunoprecipitation experiments (CoIPs)

For coimmunoprecipitation (CoIP) experiments, nuclear extract [25 mg/ml] from wild-type *Drosophila* embryos was used. CoIPs were performed in IP150 buffer [HEMG150 (25 mM Hepes, 150mM KCl, 0.2 mM EDTA, 12.5 mM MgCl₂, 10% (vol/vol) glycerol), 0.5% Tween-20, 0.2 mg/ml BSA, 0.2 mM PMSF, 0.5mM DTT, Roche COMPLETE protease inhibitor]. 100µl extract was mixed with 600µl IP150 buffer and pre-cleaned on protein G beads (Sigma) for 30min. at 4°C to remove unspecific resin-binding proteins. After this pre-cleaning step the supernatant was mixed with 6µl of the respective antibody serum or preimmune serum for 1 hour, rotating at 4°C. Following 4 washes with 700µl IP150-buffer each, the beads were resuspended in 50µl of 4xSDS-loading buffer, boiled for 5min. at 95°C and 40µl of the supernatant isolated. Fractions of this were loaded on SDS-PAGE for separation and subsequent western blot analysis with the corresponding antibodies.

10 Western Blot

The blotting procedure was optimized for large proteins like NSL1 as follows.

The system used for SDS-PAGE and subsequent transfer was the Mini Protean3 (BioRad). After separation on low-percentage (6%) polyacrylamide gels, proteins were transferred to 0.45µm pore-size PVDF membrane (Schleicher&Schuell) for 1.5-2 hours at 120 Volts. Only for the analysis of small histone proteins a membrane pore size of 0.22µm was used. The buffer was a standard Tris/glycine western transfer buffer with only 10% methanol.

Membrane was blocked with blocking buffer (PBS, 0.3% Tween20, 5% milk powder) for 1 hour. Depending on the protein to be detected the blot was probed with primary antibodies for 1 hour at room temperature (most of the antibodies) or at 4°C over night (NSL1 antibodies). Washes with PBS/0.3%Tween20 were done 3 times for 5min each. Secondary horseradish-peroxidase coupled antibodies were routinely used at a 1:10 000 dilution and incubated with the membrane for 45min. Washes as before.

Westerns were developed with LumiLite reagent (Roche) according to the manufacturer's instructions and exposed on BioMax MR film (Kodak).

11 Recombinant protein expression using the Baculovirus / SF9 system

11.1 Generation of NSL1 and NSL3 Baculoviruses

For the recombinant expression of full-length NSL1 and NSL3 proteins as well as for fragments of these proteins I used the Bac-to-Bac Baculovirus Expression System (Invitrogen). This is a site-specific transposition system to generate baculoviruses which are used for infection of SF9 insect cells and expression of the recombinant proteins therein.

cDNAs coding for NSL1 and NSL3 were cloned into the pFastBac-HT(c) vector (Invitrogen) as described above. This vector allows for the expression of His₆-tagged recombinant proteins under the control of a polyhedrin promoter. A TEV cleavage site is present between the tag and the protein of choice.

Briefly, bacmids were generated by transformation of the pFastBac-HT(c)-constructs into DH10Bac cells to initiate the transposition. Isolated bacmid DNA was transfected in triplicate into SF9 insect cells using the cellfectin transfection reagent. Three rounds of amplification were performed to produce the high-titer baculovirus stocks. Already after round 2 (P2 stock) clones were checked on a small scale for expression of the correct proteins.

11.2 Baculovirus protein expression and purification from SF9 cells

For larger scale expression of the recombinant proteins, 50ml of actively dividing SF9 cells at a density of 1×10^6 cells/ml were infected with 5ml of P3 baculovirus stock. Cells were placed on a shaker at 26°C for 2 days to allow for protein expression.

Prior to harvesting 48 hours post-infection, the cells were checked for successful infection. This could be judged by the presence of a majority of round and big cells and by the stagnant cell number because of the growth arrest that SF9 cells exhibit upon virus infection.

Cells were harvested by centrifugation (1000rpm, 4°C, 15min.) and resuspended in 4ml of His-purification buffer (30mM Tris pH8, 200mM NaCl, 1% Triton, 5mM β -

Mercapto, 0.2mM PMSF, 5mM Imidazole, 10% glycerol). Total cell lysate was prepared by 3 freeze-thaw cycles in liquid nitrogen, addition of 6ml of fresh His-purification buffer and centrifugation (3750rpm, 30min., 4°C). The supernatant was mixed with 350µl pre-equilibrated Ni-NTA resin (QIAGEN) in a 15ml Falcon tube and rotated for 2 hours at 4°C. After binding the beads were collected in a gravity-flow column and washed with 5ml each of ice-cold wash buffer 1 (WB1: 30mM Tris pH8, 200mM NaCl, 5mM β-Mercapto, 0.2mM PMSF, 5mM Imidazole, 1% Triton), WB2 (WB1 + 500mM NaCl) and WB3 (WB1 + 15mM Imidazole). Bound proteins were eluted with 4x 500µl fractions of elution buffer (WB1 + 315mM Imidazole).

An additional purification step was used for the eluted NSL3 protein. Protein containing elution fractions were pooled and mixed with recombinant His₆-tagged TEV protease (EMBL). The cleavage reaction was allowed to proceed over night at 4°C. Subsequently the buffer was exchanged to His-purification buffer on PD10 columns (Amersham) to get rid of the high imidazole concentration. This protein solution was passed over a second Ni-NTA column to remove the His₆-tagged TEV protease, the free His₆-tag and the uncleaved fusion protein. The free NSL3 protein was collected in the flow-through and concentrated on Centriprep YM-30 (Millipore).

The purification result was analysed by SDS-PAGE followed by Coomassie staining.

11.3 Co-expression and purification

For the purpose of testing particular protein-protein interactions two different strategies were used.

The first approach involved the expression of two proteins to be tested in the same cell. This was achieved by double-infection of SF9 cells with the two corresponding virus stocks – either both having different tags or one of them having no tag (example in Figure 29-1 A). Lysate preparation and purification was essentially done as described above unless other affinity-tags used for co-purification involved a different purification strategy.

To circumvent a possible problem of the first approach – namely varying infection efficiencies for the combined viruses – a second method was established.

Recombinant proteins were first expressed singularly by doing single infections. Individual lysates were prepared the same way as described above. Before passing the lysates over the affinity column they were mixed and incubated together for 2 hours on a rotating wheel at 4°C. This allowed complex formation of interacting proteins which were isolated together in the next step of affinity purification. See the Coomassie-stained protein gel in Figure 29-1 B as an example for the individual expression of MOF and NSL1 proteins and subsequent co-purification of the dimer.

12 Purification of the NSL complex from *Drosophila melanogaster* Schneider cells

12.1 Affinity-tagging of the NSL1 protein

The full-length coding sequence of the NSL1 protein was cloned into the multiple-cloning site of the pBSactshort-N-TAP vector (kind gift of Elisa Izzaurrealde). The main feature of this pBluescript-derived vector is the addition of an N-terminal Tandem-Affinity-Purification tag (TAP-tag) to the protein of choice. Expression is driven from a shortened Actin5C-promoter and transcription terminated by a BgH1 terminator sequence.

The NSL1 coding sequence was isolated from the pFastBac-HTc-NSL1 construct by digestion with *Nco*I(5') and *Not*I(3') enzymes. The 3' end was blunted with Klenow polymerase (NEB). Unfortunately the released insert and the remaining pFastBac-HT(c) vector exhibited the same size and could not be distinguished after agarose gel electrophoresis. Therefore an additional treatment of the fragments with *Sca*I enzyme was necessary to cut the donor vector in smaller pieces before isolation of the NSL1 sequence. The acceptor vector pBSactshort-N-TAP was prepared by restriction digest with *Nco*I and *Bam*HI enzymes. The *Bam*HI site was blunted and ligated to the blunt 3'-end of NSL1. In-frame cloning was assured by ligation of the 5' *Nco*I site which restored the first ATG of the NSL1 cDNA. The resulting construct was named pBSactshort-N-TAP-NSL1.

To tag the NSL1 with additional epitopes for purification (one FLAG-tag followed by an influenza hemagglutinin (HA) tag), the pBSactshort-N-TAP-NSL1 construct was used as starting material.

Oligonucleotides with coding sequences for FLAG and HA epitopes were designed such that they resulted in pre-formed *NcoI* sites on either side when hybridized (oligo sequences: *NcoI*-FLAG-HA-fwd 5'-catg gac tac aag gac gac gat gac aag tac cca tac gac gtc cca gac tac gct gg-3' and *NcoI*-FLAG-HA-rev 5'-gtacgggtcgcatcag accctgcag catacccatgaacagtagcagcaggaacatcag-3').

After temperature annealing and treatment with polynucleotide kinase (NEB) the hybridized oligonucleotides were used for a sticky-end ligation into the *NcoI* cut pBSactshort-N-TAP-NSL1 vector. Sequenced clones that showed the right orientation of the FLAG-HA insert were selected and named pBSactshort-N-TFH-NSL1 to describe the triple-tagged (TAP-FLAG-HA) NSL1 construct.

12.2 Generation of stable *Drosophila melanogaster* Schneider cell lines

Schneider (S2) cell lines were established for the stable expression of recombinant proteins. Briefly, S2 cells were transfected with the respective expression plasmid and the pUC-NEO resistance vector. The Effectene transfection reagent (Qiagen) was used according to the manufacturer's instructions and transfections carried out in triplicate for every line. MOCK transfections, where only the expression vector but not the resistance vector was transfected, were done in parallel. Cells were allowed to recover for 24 hours before the medium was exchanged. Selection with Geneticin (G418, Invitrogen) started 48 hours after transfection. Each of the 3 transfections per line was exposed to a different concentration of the antibiotic – normally 0.8, 1 and 1.2 mg/ml of G418. Selection was monitored by the death of MOCK transfected cells and colony formation of stably expressing cells. The cell line with the optimal G418 concentration - 1mg/ml G418 for stable NSL1 cell lines - was selected and tested for expression after 4 weeks of selection.

12.3 Preparation of nuclear extracts

Nuclear extracts were prepared from either wild-type or transgenic S2 cell lines by using a modified version of the standard protocol (Dignam et al., 1983). Cells were amplified in 175cm² cell culture flasks (Nunc) on a shaker at 26°C. Between 3 x10⁹ and 5x10⁹ cells were harvested from cell culture media by a 15min centrifugation step at 1300rpm in a cooled benchtop centrifuge (Eppendorf 5810R). Cells were washed once with ice-cold phosphate buffered saline (PBS). The following steps were performed at 4°C. Cell pellet was resuspended in 3 to 4 pellet volumes of Buffer B10 (15mM Hepes pH7.6, 10mM KCl, 5mM MgCl₂, 0.1mM EDTA, 0.5mM EGTA, 0.5mM DTT, 0.2mM PMSF, Complete protease inhibitors). The cells were allowed to stand for 10 min. and then lysed with 80 strokes on ice using the B type pestle of a 40ml dounce homogenizer. Lysis efficiency was checked under the microscope and the nuclei pelleted by centrifugation at 4500rpm for 10min in the SS34 rotor of a Sorvall RC6 centrifuge. The supernatant represented the cytoplasmic fraction and was discarded.

The nuclei pellet was washed in 30ml of Buffer B10 and centrifuged again for 10min at 4500rpm (SS34). This nuclei fraction was resuspended in 7ml of Buffer B10 and layered on the same volume of buffer B10 containing 0.8M sucrose to create a sucrose gradient. Separation of the nuclei from residual cell membranes and debris was achieved by centrifugation in a HB4 swing-out rotor (Sorvall) for 10min at 4000rpm. The supernatant was discarded and crude nuclei were resuspended in 3 pellet volumes of buffer B110 (15mM Hepes pH7.6, 110mM KCl, 5mM MgCl₂, 0.1mM EDTA, 0.5mM EGTA, 0.5mM DTT, 0.2mM PMSF, Complete protease inhibitors) and dounced with a glass homogenizer (60 strokes with B type pestle). The suspension was precipitated with 400mM final concentration of (NH₄)₂SO₄ (pH8.0) to break nuclei open. The mixture was rotated for 1 hour at 4°C before ultracentrifugation (35.000 rpm, 1hour) in Ti70 rotor of a Beckman L-70 ultracentrifuge. Isolated supernatant was again precipitated with an equal volume of 4M (NH₄)₂SO₄ (pH8.0) and rotated for 30min at 4°C. Precipitated proteins were pelleted by centrifugation in glass centrifuge tubes (Corex) in SS34 rotor at 12.000rpm for 30min.

Protein pellet was resuspended in HEMG120 buffer (25 mM Hepes, 120mM KCl, 0.2 mM EDTA, 12.5 mM MgCl₂, 10% (vol/vol) glycerol, 0.2 mM PMSF, 0.5mM

DTT, Roche COMPLETE protease inhibitor) and dialysed through Spectra/Por dialysis membrane (MWCO: 6-8kDa, Spectrapor) against 1 liter of HEMG0 buffer (25 mM Hepes, no KCl, 0.2 mM EDTA, 12.5 mM MgCl₂, 10% (vol/vol) glycerol, 0.2 mM PMSF, 0.5mM DTT, Roche COMPLETE protease inhibitor) for 2-3 hours with 2 changes of buffer. The dialysis was checked by conductivity measurements to achieve a final concentration of 120mM salt.

The total protein concentration was usually 2 to 4 mg/ml. Extracts with higher protein concentration (around 20 mg/ml) were obtained by omitting the second ammonium sulfate precipitation and immediate dialysis of the supernatant.

The nuclear extract was frozen as aliquots in liquid nitrogen and stored at -80°.

12.4 Tandem affinity purification (TAP)

The TAP procedure was performed as described in (Mendjan et al., 2006) with a few modifications to the protocol.

Nuclear extracts (5-10 mg/ml) were prepared from *Drosophila* S2 Schneider cells, stably expressing the TAP-NSL1 protein (as described above). The extracts were diluted in IgGBB150 (25mM Hepes pH7.6, 150mM KCl, 5mM MgCl₂, 0.5 mM EDTA, 20% glycerol, 0.5mM DTT, 0.2%Tween20, 0.2mM PMSF, Complete protease inhibitors) to about 5mg/ml protein concentration and spun down at maximum speed in a Eppendorf microcentrifuge for 15min.

Crosslinked IgG beads (Roche) were equilibrated in IgGBB150 before binding. Diluted extract was bound to IgG beads at 4°C for 60-90min on a rotating wheel. Supernatant was separated from the beads after binding and beads washed three times each with IgGBB150 and IgGBB200 (identical to IgGBB150 but with 200mM KCl). The last 2 washes were performed at RT for 5-10min each.

Beads were resuspended in TEV cleavage buffer CB150 (20mM Hepes pH7.6, 150mM KCl, 0.5 mM EDTA, 0.5mM DTT, 0.1%Tween20, 0.2mM PMSF, 10µg/ml TEV protease). The first wash was done with CB150 without TEV at RT and the second wash with CB150+TEV. Protease cleavage was allowed to take place for 2 hours at 18°C while rotating the reaction slowly. Supernatant was isolated and further centrifuged at maximum speed for 5min at 4°C. 3µl of 1M CaCl₂ was added per 1ml of cleavage supernatant. The supernatant obtained from the TEV cleavage reaction was diluted in a 1:3 ratio with calmodulin binding buffer

CalBB150 (20mM Hepes pH7.6 / 10mM Tris pH7.6, 150mM KCl, 2mM Mg-Acetate, 1mM Imidazole, 3mM CaCl₂, 20% glycerol, 10mM β-Mercaptoethanol, 0.2%Tween20, 0.2mM PMSF, Complete protease inhibitors). Calmodulin beads were equilibrated with CalBB150 before binding which was allowed for 2 hours at 4°C. Washes (each wash 5-10min) were performed according to the following scheme: 2x CalBB150 at 4°C, 2x CalBB150 at RT, 2x CalBB200 at RT. Final wash was done in CalBB150 (with Tris pH7.6 instead of Hepes).

According to the purification background, elutions were performed either directly in 1xSDS loading buffer (with β-Mercaptoethanol, no DTT) by boiling or in CalEI150 (20mM Tris pH7.6, 150mM KCl, 2mM Mg-Acetate, 1mM Imidazole, 3mM EGTA, 20% glycerol, 10mM β-Mercaptoethanol, 0.2mM PMSF) for 15-30min at 4°C on a shaker. The resulting eluate was analysed by SDS-PAGE and protein bands were revealed by silver staining.

12.5 FLAG/HA affinity purification

Isolation of the NSL1 complex was either done in a one-step purification format (αFLAG) or as a two-step purification (αFLAG followed by αHA). Purifications from S2 cells that were expressing tagged proteins were always accompanied by purifications from wild-type S2 cell nuclear extract to be able to compare specific enrichment.

The affinity resins used were anti-FLAG M2 agarose beads (Sigma) and monoclonal anti-HA agarose conjugate (Sigma). Beads were routinely stripped for 1min with ice-cold 0.2M glycine prior to use.

Immunoprecipitation from 1.5ml of Schneider cell nuclear extract was performed with 45µl compact beads that were pre-equilibrated with the binding buffer (IgGBB150: 25mM Hepes pH7.6, 150mM KCl, 5mM MgCl₂, 0.5 mM EDTA, 20% glycerol, 0.5mM DTT, 0.2%Tween20, 0.2mM PMSF, Complete protease inhibitors). Binding was allowed to take place for 1 hour rotating at 4°C. Unspecific binding proteins were reduced by 5 alternating washes (RT, 4°C) with 1ml IgGBB150 each.

For single-step purification the FLAG bound protein complex was eluted with 4x 100µl fractions of Tris-based elution buffer (TIgGEI150: 20mM Tris pH7.6, 150mM

KCl, 5mM MgCl₂, 0.5 mM EDTA, 20% glycerol, 0.2mM PMSF, Complete protease inhibitors, 400µg/ml FLAG peptide) for 10min each elution at RT.

For two-step purifications the FLAG elution was done with HEMG-based elution buffer (HiGGEI150: 25mM Hepes pH7.6, 150mM KCl, 5mM MgCl₂, 0.5 mM EDTA, 20% glycerol, 0.2mM PMSF, Complete protease inhibitors, 400µg/ml FLAG peptide).

The elutions were pooled and incubated with anti-HA agarose for 1 hour at 4°C followed by 3 alternating washes at RT or 4°C with 1ml IgGBB150 each. HA-elutions with 2x 120µl TIgGEI150 were done for 20min at RT each.

13 Analysis of the purified NSL complexes

13.1 Silver-staining of SDS polyacrylamide gels

Purification results were checked by separation on 5-15% SDS-polyacrylamide gradient gels and subsequent silver staining. A modified silver staining protocol from Shevchenko et al (Shevchenko et al., 1996) was used. All steps were performed on a shaking table at RT unless otherwise stated.

After electrophoresis the gel was fixed for 30min in fixing solution (40% methanol, 10% acetic acid). It was then rinsed several times with distilled water to remove the remaining acid and incubated in water o/n at 4°C. The next day the gel was sensitized by a 1min incubation with 0.02% (w/v) sodium thiosulfate, and then rinsed with two changes of water for 1 min each. After rinsing, the gel was submerged in chilled 0.1% (w/v) silver nitrate solution and incubated for 35 min at 4 °C. Two 1min washes with water preceded the developing step where the gel was incubated in developing solution (0.04% (v/v) formaldehyde, 2% (w/v) sodium carbonate) until sufficient staining was obtained. Development was quenched by exchanging the developing solution with 1% acetic acid.

13.2 Flamingo Staining of SDS polyacrylamide gels

This procedure is very similar to the silver staining but is taking advantage of a fluorescent dye called Flamingo (BioRad) that is used for protein visualization. Because of its completely different mode of action it can help to detect proteins that are insensitive to staining with silver.

The slab gel was fixed o/n with 40% ethanol and 10% acetic acid and immediately incubated in 1:10 diluted Flamingo staining solution for 3 to 5 hours. The gel was submerged for 10min in a 0.1% Tween20 solution before scanning on a PharosFX scanner (BioRad).

13.3 Identification of interacting proteins by mass spectrometry

Purified protein samples were prepared for mass spectrometry in two ways. Either individual silver-stained protein bands were digested in gel with trypsin as described (Shevchenko et al., 1996) or complex elution fractions were mixed with SDS-loading buffer and electrophoresed into the stacking gel of a 10% SDS-PAGE. The non-separated proteins were then stained by Coomassie blue and the total band excised from the gel and trypsin digested.

The samples were separated on a nano-flow 1D-plus Eksigent (Eksigent, Dublin, CA) HPLC system coupled to a qStar Pulsar i quadrupole time-of-flight MS (Applied Biosystems, Darmstadt, Germany).

The digest was loaded onto a 100 μm i.d. fused silica CapRod monolithic C18 pre-column and washed with Phase A (2% acetonitrile and 0.5% acetic acid in water) (all Merck, Darmstadt, Germany). The reverse-phase separation was performed on a column made by slurry packing 3 μm YMC C18 particles (YMC, Dinslaken, Germany) into a tapered 20 cm 100 μm i.d. fused silica capillary (Optronis, Kehl, Germany). The peptides derived digested samples were separated by a linear gradient which started at 100 % mobile phase A and increased the mobile-phase composition to 50% B (0.5% acetic acid in 98% acetonitrile) over a span of 45 minutes at a constant flow rate of 200 nl/min. Each run was followed by 15 minutes 100% mobile phase B. Peptides derived from digested standard proteins were separated by a similar gradient over 30 minutes. The MS was operated in data-dependent mode. MS spectra were acquired over m/z range from 350 to

1300 for 1 second and one subsequent MS/MS spectra from 80 to 1800 m/z for 1.5 seconds. Selected precursor ions were excluded for 50 seconds from the analysis. MS/MS data was extracted using the AnalystQS software v1.0SP8 and the vendor provided script Mascot.dll v1.6b16 (AppliedBiosystems, Darmstadt, Germany). The header information of the resulting peak list was modified based on the requirements of MsQuant v1.4a16 (29) using an in-house Practical Extraction and Report Language (PERL) script. Peptides were identified by searching the peak-list against the NCBI nr (v14_08_2006, 486696 mammalian entries) and SwissProt (vUniProt_Knowledgebase_Release_8.4, 41327 mammalian entries) database using the MASCOT v2.103 (Matrix Science, London, UK) algorithm. The taxonomy parameter was restricted to drosophila, trypsin cleavage specificity was allowed one missed cleavage, peptide tolerance was limited to 0.2 Da, fixed modifications were carbamidomethylation of cysteine, variable modifications were oxidation of methionine, and peptides with a score below 18 were excluded. All proteins were identified by MASCOT by at least two peptides in two independent samples with a summed ion score above 45.

14 Biochemical characterization of NSL complexes

14.1 *In vitro* assembly of polynucleosomes

Linearized plasmid pBS-SK was end-labelled with Biotin-C14-dATP (Invitrogen) and coupled to magnetic M280Streptavidin beads (Dyna). Polynucleosomes were reconstituted with pre-assembled recombinant *Xenopus* histone octamers by salt exchange from 1M NaCl to 100mM NaCl in SE buffer (10 mM Tris-Cl, pH8.0, 1 mM EDTA, 2mg/ml BSA). A histone:DNA ratio of 1:1.1 was used. 1.5µg of total histones went into one reaction.

14.2 Histone acetyltransferase assays (HAT assays)

The histone acetyltransferase assays were essentially performed as described earlier (Akhtar et al., 2000a).

For liquid HAT assays, 35µl of the elution fractions from the corresponding purifications were incubated together with the assembled polynucleosomes on beads for 80min at 26°C in HAT buffer (20mM Tris pH8.0, 1.5mM MgCl₂). The HAT buffer was supplemented with 0.05 µCi [¹⁴C]Acetyl-CoA (Amersham) as a cofactor for the reaction. Total salt concentration was adjusted to 40mM KCl. Moderate shaking assured the homogeneous distribution of beads in the reaction volume. For autoradiography, the reactions were stopped with SDS loading buffer and proteins resolved by 15% SDS-PAGE. The gel was shortly stained with Coomassie blue and subsequently treated with Amplify solution (Amersham). The dried radioactive gel was exposed at RT on BAS-TR2040 imaging plate (Fujifilm). Signals were read with BASreader (Fujifilm).

14.3 Site-directed mutagenesis of histone H4 tail lysine residues

To test for the specific enzymatic acetylation of single lysine residues in the tails of recombinant *Xenopus* histone H4, the lysines 8, 12 and 16 were mutated to alanine residues ((H4K8A, H4K12A, H4K16A). The pET3-HistoneH4 plasmid was used as a template for introduction of point mutations with the QuikChange site-directed mutagenesis kit (Invitrogen). Primers used for mutation as follows:

H4-K8A	5'-GTCTGGTCGTGGTAAAGGTGGT <u>G</u> CAGGTCTGGGTAAAG-3'
H4-K8A_antisense	5'-CTTTACCCAGACCTGCACCACCTTTACCACGACCAGAC-3'
H4-K12A	5'-TGGTAAAGGTGGTAAAGGTCTGGGT <u>G</u> CAGGTGGTGCTAAA-3'
H4-K12A_antisense	5'-TTTAGCACCACCTGCACCCAGACCTTTACCACCTTTACCA-3'
H4-K16A	5'-AAGGTCTGGGTAAAGGTGGT <u>G</u> CACGTCACCGTAAAG-3'
H4-K16A_antisense	5'-CTTTACGGTGACGTGCAGCACCACCTTTACCCAGACCTT-3'

* introduced mutations are underlined in sense oligos

All constructs were sequenced to verify the desired point mutations.

14.4 Preparation of wild-type and H4-mutant histone octamers

The protocol used for histone preparation and octamer reconstitution was taken from Luger et al (Luger et al., 1997).

14.4.1 Expression of recombinant histone proteins

Histones were expressed from *E.coli* strain BL21(DE3)pLysS that were transformed with the pET3-histone expression plasmids.

Large scale expression cultures (4 liter) were grown in LB medium to $OD_{600} \sim 0.5$ and induced by addition of Isopropyl- β -D-thiogalactopyranosid (IPTG) to a final concentration of 0.2 mM. Cells were left shaking for protein expression at 37°C for 2 hours (H3 and H4) or 3 hours (H2A and H2B) before harvesting by centrifugation. Pellet was resuspended in wash buffer (50mM Tris-HCl pH 7.5, 100mM NaCl, 1mM EDTA, 5mM β -mercaptoethanol, 0.2mM PMSF, Complete protease inhibitors) and shock-frozen in liquid nitrogen.

14.4.2 Inclusion Body Preparation

Thawed cell pellets were resuspended (20ml wash buffer per 500ml culture) and sonicated (Sorvall Omnimixer: 3x 30sec, 60% power) on ice to break the bacteria. Inclusion bodies were isolated by centrifugation at 15000rpm for 20min in a cooled SS34 rotor of a Sorvall RC6 centrifuge. Supernatant was discarded and the pellet washed two times with 20ml of Triton wash buffer (wash buffer + 1% (v/v) Triton X-100). Centrifugation steps as before. Two more washes with 20ml each of wash buffer (without Triton) and a last centrifugation completed this step.

14.4.3 Histone unfolding

Inclusion body pellets were resuspended in 15ml of unfolding buffer (7M guanidinium HCl, 20mM Tris-HCl pH7.5, 10mM DTT, sterile filtered) and incubated at RT stirring smoothly with a magnetic stirrer. After centrifugation at 15000rpm for 20min at 4°C (SS34 roto, Sorvall RC6 centrifuge) the supernatant was isolated

and the previous steps repeated with the remaining inclusion bodies. The supernatants were pooled at the end and dialysed against SAU200 buffer (7M urea, 20 mM sodium acetate pH5.2, 200mM NaCl, 5 mM β -mercaptoethanol, 1mM EDTA, sterile-filtered). Dialysis buffer was exchanged according to the following scheme: 2x 1liter for 1 hour each, 1x 1liter overnight).

14.4.4 Histone purification by ion exchange chromatography

Dialyzed histone proteins in SAU200 buffer were centrifuged again (10min., 4°C, 10.000rpm, SS34 rotor) to remove any insoluble matter. The protein solution was loaded on a preparative HighTrap SP FF (Amersham) ion exchange column that was equilibrated with SAU200 buffer. Using a flow rate of 4 ml/min, proteins were eluted with increasing salt concentration by a linear gradient to SAU600 buffer (7M urea, 20 mM sodium acetate pH5.2, 600mM NaCl, 5 mM β -mercaptoethanol, 1mM EDTA, sterile-filtered).

Peak-fractions were analyzed by SDS-PAGE and fractions containing pure histone protein pooled. After an o/n dialysis step against water the protein concentration was determined and histone proteins lyophilized in aliquots of 1mg and stored at – 20°C.

The presence of the correct mutations in the mutant H4 histones was verified by mass spectrometric analysis of the purified histone proteins.

14.4.5 Histone refolding and reconstitution of histone octamers

A 1mg aliquot of each lyophilized histone was dissolved to a concentration of approximately 2 mg/ml in unfolding buffer (7M guanidinium HCl, 20mM Tris-HCl pH7.5, 10mM DTT, sterile filtered) and was allowed to proceed for no more than 3 hours.

The absorbance of the unfolded histone proteins was measured at 276nm and concentration calculated for every histone. Four histone proteins were mixed to equimolar ratios and adjusted to a final protein concentration of 1 mg/ml using unfolding buffer. The mixture was dialysed at 4°C against three changes of 1 liter of refolding buffer (2M NaCl, 10mM Tris-HCl pH7.5, 1mM EDTA, 5mM β -

mercaptoethanol, sterile filtered). The third dialysis step was performed o/n at 4°C. Precipitated proteins were removed by centrifugation and the solution concentrated to approximately 250µl on a Centricon centrifugal filter unit (MWCO 10kDa, Millipore).

Gel filtration of assembled octamers was performed at 4°C on an äktaHPLC system (Amersham) equipped with an automatic fractionator. 250µl of concentrated histone octamer was injected in a Superdex200 HR10/30 gel filtration column that was pre-equilibrated with refolding buffer. Fractionation was done at a flow rate of 0.5ml/min and the octamer containing fractions collected.

Elutions were checked for purity and stoichiometry on 15% SDS-PAGE and fractions that contained equimolar amounts of the histone proteins pooled.

Protein concentration was determined ($A_{276} = 0.45$ for a solution of 1 mg/ml), octamers concentrated to 1mg/ml and adjusted to 50% (v/v) glycerol. Octamers were stored at -20°C.

15 Targeted quantitative mass spec analysis of histone lysine acetylation

Histone samples derived from acidic extraction of cell culture cells or from in vitro acetylated recombinant histones were separated on 15% 1D-SDS gels. Bands of interest were excised and modified in gel according to Peters et al (Peters et al., 2003). Modified proteins were digested in-gel with trypsin (Roche) according to standard protocols (Shevchenko et al., 1996). Peptide mixtures were analyzed by nanoflow capillary reversed phase chromatography (2D-NanoLC, Eksigent, Dublin, CA, USA, 50µm i.d. C18 columns prepared in-house) hyphenated to a Q-ToF1 mass spectrometer (Micromass/Waters, Manchester, UK).

16 RNA interference in *Drosophila* S2 and Kc cells

RNA interference was performed essentially as described before (Clemens et al., 2000) with the following modifications. S2 cells were propagated 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). They were cultured in solution in small screw-cap tubes (Sarstedt) that were placed on a shaker table. Gene-specific dsRNAs used for the knockdowns corresponded to fragments encompassing about 600 nucleotides of the coding sequences. They were amplified by PCR from corresponding cDNAs using T7-tailed oligonucleotides. The resulting PCR products were then transcribed using the T7 RiboMAX Express Large Scale RNA Production System (Promega).

For NSL1, several dsRNA fragments (see table below) targeting different regions of the NSL1 coding sequence were produced and tested for knockdown efficiency. A total of 6×10^6 S2 cells were incubated with 45µg dsRNA per 1×10^6 cells and harvested after 4 days for dMSL-1 RNAi. NSL1 and NSL3 were harvested on day 5 and day 6 respectively. For knockdown of MOF, 45µg dsRNA was added on day 1 and day 3 and cells harvested on day 8.

Primer pairs for generation of dsRNA fragments:

NSL1 A	T7CG4699-4TOP	5'-TTAATACGACTCACTATAGGGAGAATGGCCCCAGCGCTCACA-3'
	T7CG4699-5BOT	5'-TTAATACGACTCACTATAGGGAGATGAACTTGTGGCCACTGCC-3'
NSL1 B	T7CG4699shortTOP	5'-TTAATACGACTCACTATAGGGAGACTGCGCCAGGAGCGGTAACATCTAG-3'
	T7CG4699shortBOT	5'-TTAATACGACTCACTATAGGGAGAGCTGGGGCGTGCGGCTTTCTTGG-3'
NSL1 C	T7CG4699-3TOP	5'-TTAATACGACTCACTATAGGGAGATGTCGCATCAAAGTCAGAGG-3'
	T7CG4699-2BOT	5'-TTAATACGACTCACTATAGGGAGACTCGAGAAGAGCTCGCTGAT-3'
NSL1 D	T7CG4699-2TOP	5'-TTAATACGACTCACTATAGGGAGAGGTAACGCCAAAAAGGATGA-3'
	T7CG4699-2BOT	5'-TTAATACGACTCACTATAGGGAGACTCGAGAAGAGCTCGCTGAT-3'
NSL3	T7CG8233TOP	5'-TTAATACGACTCACTATAGGGAGA CCGCAGACCTCAGAGGCCAGAGGCTC-3'
	T7CG8233BOT	5'-TTAATACGACTCACTATAGGGAGACGAAAACCATCTCCTGCATGGGCGTC-3'
MOF	T7-MOF-TOP	5'-TTAATACGACTCACTATAGGGAGA ATGTCTGAAGCGGAGCTGGAACAG-3'
	T7-MOF-BOT	5'-TTAATACGACTCACTATAGGGAGA CGAAGTCGTCAATGTTGGAACCACTGCC-3'

17 Quantitative RT-PCR to assess knockdown efficiency

The quantification of the RNAi efficiency in S2 and Kc cells was done by quantitative real-time PCR (qRT-PCR). Total RNA was isolated from cells using the RNeasy kit (Qiagen). Reverse transcription of 1µg total RNA was achieved by using the SuperScript II RT kit (Stratagene) with random hexamer primers. The qRT-PCR reactions with SYBR Green PCR master mix (Applied Biosystem) were set up according to the manufacturer's protocol. Gene specific primers that were designed to span exon-exon junctions were used. Primer design was done with Primer3, a service provided by the Whitehead Institute for Biomedical Research.

Primer pairs used in the analysis were as follows:

NSL1	P1 5'-GGGATCTGGTCGAGACGAT-3' P2 5'-GGGCAACGGCCTCCAAGT-3'
NSL3	P1 5'-GAAAGATGGACGGTGGTTTAGA-3' P2 5'-CCATAGTCCTGGGCATCATT-3'
dMSL-1	P1 5'-CGGATGTAAACGCCAGAACT-3' P2 5'-AAGGCGCACAGGTCTTCTC-3'
MOF	P1 5'-CTCATCCGAACGGCAGAAG-3' P2 5'-TGCGGTCGCTGTAGTCATAG-3'

qRT-PCR was performed on an ABI real-time PCR cycler (Applied Biosystems) with SYBR detection, and the amplification curves were analyzed with the corresponding SDS software (ABi). Each qRT-PCR was done in duplicate and repeated at least three times from different biological replicates. Values were normalized to corresponding EGFP knockdown controls and to RNAPol II values. The standard error of mean within each experiment was calculated.

18 Acid histone extraction of dsRNA-treated *Drosophila* S2 and Kc cells

After treating the cells with dsRNA for the indicated time, aliquots of cells were taken and harvested by centrifugation. Normally, 6×10^6 cells were used for acid histone extraction. The cells were washed once with ice-cold PBS and the pellet resuspended in Triton extraction buffer (TEB: PBS, 0.5% TritonX100 (v/v), 0.2mM phenylmethylsulfonylfluoride (PMSF), 0.02% (v/v) NaN_3) at a cell density of 10^7 cells per ml. Lysis was allowed to proceed on ice for 10min with gentle flipping from time to time. A centrifugation step at 2000rpm (Eppendorf 5417C) for 10min at 4°C was done to form a small pellet at the wall of the tube. The supernatant was removed and discarded and the pellet washed again in half the volume of TEB as before. The washed pellet was subsequently resuspended in 0.2N hydrochloric acid (HCl) at a cell density of 4×10^7 cells / ml.

Histone extraction took place at 4°C overnight. The next morning, the histone containing supernatant was separated from the residual debris by an additional centrifugation at 2000rpm for 10 minutes at 4°C . Acid extracted histones were stored at -20°C until they were used.

19 Immunofluorescence staining of S2 and Kc cells

19.1 Fixation of cells

Round Coverslips were transferred to a 24-well plate and approximately 300 μl cells in growth medium were dispensed on the coverslips. Cells were allowed to settle down for 30min. Before the fixation, cells were washed twice with PBS and subsequently fixed for 10min in a 3.7% formaldehyde / PBS solution. In the next step cells were washed twice with PBS and then blocked with IF blocking buffer (PBS, 0.1% Tween20, 0.1% Triton-X100, 5% BSA). Blocking was done for 1 hour at RT or o/n at 4°C .

19.2 Immunostaining of cells

Fixed and pre-blocked cells were incubated for 1 hour with primary antibody at RT. Antibodies were 1:500 diluted (depending on the antibody) in IF blocking buffer. For coimmunostainings primary antibodies from different species were readily mixed. Slides were washed 3x 10min with IF wash buffer (PBS, 0.1% Tween20, 0.1%, Triton-X100) before a 1 hour incubation step with the fluorescent-labelled secondary antibodies (diluted in IF blocking buffer). DNA staining with Hoechst dye (1:2000, Invitrogen) was included in the second of the last 3 washes with IF wash buffer. Finally, round cover slips were mounted on microscopy glass slides with 2-3µl of FluoroMount-G (SouthernBiotech).

20 Immunofluorescence staining of polytene chromosome squashes

20.1 Preparation of chromosome squashes

Pairs of salivary glands were dissected in PBS from crawling 3rd instar larvae. Glands were fixed for 10min with 3.7% paraformaldehyde (in H₂O) on poly-L-lysine coated glass slide and covered with a Sigmacote (Sigma) treated cover slip. Tapping the coverslip with 15 strokes of a pencil was necessary to break up the cells and nuclei. Chromosomes were spread by pressing slide on blotting paper. After freezing the slide in liquid nitrogen the coverslip was removed with a razorblade and the slide washed two times for 15 min each in PBS slowly shaking the rack.

20.2 Immunostaining of polytene chromosome squashes

Preparation of polytene chromosomes from salivary glands of 3rd instar larvae was performed as described below. A detailed description can also be found on <http://www.igh.cnrs.fr/equip/cavalli/Lab%20Protocols/Immunostaining.pdf>.

Slides were blocked for 1 hour in blocking solution (PBS, 15% milk powder, 3% BSA, 0.2% NP-40, 0.2% Tween 20) at RT. Primary antibodies were diluted 1:50 to 1:500 in blocking solution (depending on the antibody). 40µl diluted antibody was added to the slide, covered with a coverslip and incubated at RT for 1 hour in a humidified chamber. After rinsing the slides in PBS they were washed 3 times in blocking solution (5min each wash). The same procedure was repeated for the secondary antibodies. Depending on the fluorophore they were applied in a dilution of 1:500 or 1:1000. To counterstain the chromosomes they were incubated with 40µl of a 1:2000 Hoechst [10mg/ml] solution in PBS. Final washes (15min each) with wash solution 1 (PBS, 300mM NaCl, 0.2% NP-40, 0.2% Tween20) and wash solution 2 (PBS, 400mM NaCl, 0.2% NP-40, 0.2% Tween 20) completed the staining procedure.

Coverslips were mounted on the slides with 10µl of FluoroMount-G (SouthernBiotech).

21 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. Images were processed with Adobe Photoshop and arranged with Adobe Illustrator.

22 EGFP-tagging of NSL1 for transfection in *Drosophila* S2 cells

For localization studies of the NSL1 protein within its endogenous environment, the protein was fused to an EGFP protein. The *NcoI/NotI*-fragment (NSL1+BgH1 terminator) was taken out from the pBSactshort-N-TAP-NSL1 construct and inserted in the *NcoI/NotI* digested pBSactshort-EGFP vector (kind gift of Elisa Izzaurre) thereby replacing the existing BgH1 terminator sequence.

The construct was named pBSactshort-EGFP-NSL1. It was used to transiently transfect S2 cells as well as for the generation of a S2 cell line stably expressing the N-terminal EGFP-NSL1 fusion.

23 EGFP-tagging of *Drosophila* histone H2B and generation of a stable S2 cell line

The live cell imaging of dsRNA-treated S2 cells required visualization of the chromosomes to track the formation of the observed segregation defects. Therefore an EGFP-tagged histone H2B was constructed. The sequence for the *Drosophila* histone H2B was amplified from genomic DNA isolated from wild-type flies. Restriction sites were introduced by the PCR-primers (H2BfwdNcoI 5'-TTGCCATGGGCATGCCTCCGAAAAGTAGTGG, H2BfwdBamHI 5'-TGCGGATCC ATGCCTCCGAAAAGTAGTGG) and allowed subcloning into the pBSactshort-EGFP vector (kind gift of Elisa Izzaurre). The construct was named pBSactshort-EGFP-H2B.

A stable cell line with the N-terminally EGFP-tagged H2B was established according to the above mentioned protocol. Only a small fraction of cells was expressing the fusion protein after the G418 selection. Therefore the EGFP-expressing population had to be greatly enriched by fluorescence activated cell sorting (FACS).

24 Live Cell Imaging

The live cell imaging experiments were performed on a Zeiss LSM510 META microscope (Zeiss) that was equipped with a motorised stage and a software macro for automatic cell tracking. The objective used was a 63x plan-apochromat 1.4 oil DIC. Pictures were processed with Zeiss LSM510 software and assembled using Adobe Photoshop and Illustrator.

Stable EGFP-H2B expressing S2 cells were treated with dsRNA for NSL1 and NSL3 as described in the corresponding section. On day 4 and day 5 of the

knockdown, cells were followed under the microscope for approximately 14 hours with pictures taken every 25 minutes.

In preparation for microscopy, dsRNA-treated cells were seeded into the wells of an 8-well LabTek chambered coverglass (Nunc) and left for 1 hour to settle down. The culture medium was novated and the chamber sealed off. Untreated EGFP-H2B S2 cells were used as a control, as the standard control of EGFP RNAi was inept for obvious reasons.

RESULTS AND DISCUSSION

In the results and discussion section of my thesis I will focus on essentially five topics. Below is a summary of what will be discussed.

The first one relates to the cloning of the full-length NSL1 and NSL3 cDNAs for protein expression and the generation of specific antibodies that were crucial for NSL localization, the initial characterisation of MOF-containing complexes and also for the later analysis.

Secondly, the strategy for successful NSL complex purification and identification of co-purified proteins by mass spectrometry will be presented.

The third part will deal with the biochemical characterization of the HAT-activity inherent to the NSL complex.

This will be followed by a fourth section about the analysis of bulk histone acetylation in the context of MSL and NSL complexes.

The fifth and last part will try to shed light on the cellular phenotype that emerges from the RNAi-mediated depletion of the NSL1 protein.

Results will be discussed in place and subsumed in a final discussion.

25 Getting started - cloning of the full-length NSL1 and NSL3 cDNAs

The analysis of an uncharacterized protein requires that the coding sequence of this gene is available for manipulation. By checking the public cDNA library of the Berkeley *Drosophila* Genome Project (BDGP) it became obvious that the predicted full-length cDNA for CG4699 (NSL1) did not exist. Only a minor part of the cDNA was assigned in the collection. Therefore, the NSL1 coding sequence had to be isolated from flies (see Figure 25-1). For this purpose, total RNA was extracted from adults of *Drosophila melanogaster* and mRNA purified via binding to magnetic Oligo(dT)₂₅ beads. A reverse transcription step yielded the cDNA pool from which the NSL1 sequence was amplified with gene-specific primers by PCR. The 3 fragments were subcloned in the pFastBac-HTc-vector for later baculovirus production and the point mutations eliminated by multiple cycles of site-directed mutagenesis. After generating the full-length cDNA it became clear that, compared to the database sequence, a short stretch of 63bp was missing. The base-pairs

26 Generating the tools for detection – antibodies against NSL1 and NSL3

Absolutely essential tools for the molecular characterization of a protein are antibodies that can specifically recognise it. They are very valuable for judging the protein's expression in a system, contribute to the analysis of protein localization in the cell and are indispensable for many biochemical assays. For these two so far uncharacterized *Drosophila* proteins – NSL1 and NSL3 – no antibodies were available. As it has been very difficult to obtain good antibodies for NSL1 and because of the necessity to use different antigens for immunization, this part will be elaborated a bit more in detail.

The first step was to generate antibodies in rats. The GST-fusions used in the first instance to immunize rats were fragments comprising part of the C-terminal PEHE domain (aa1019-1287) and the putative hydrolase domain (aa100-590) of NSL1 and NSL3 respectively (see Figure 26-1).

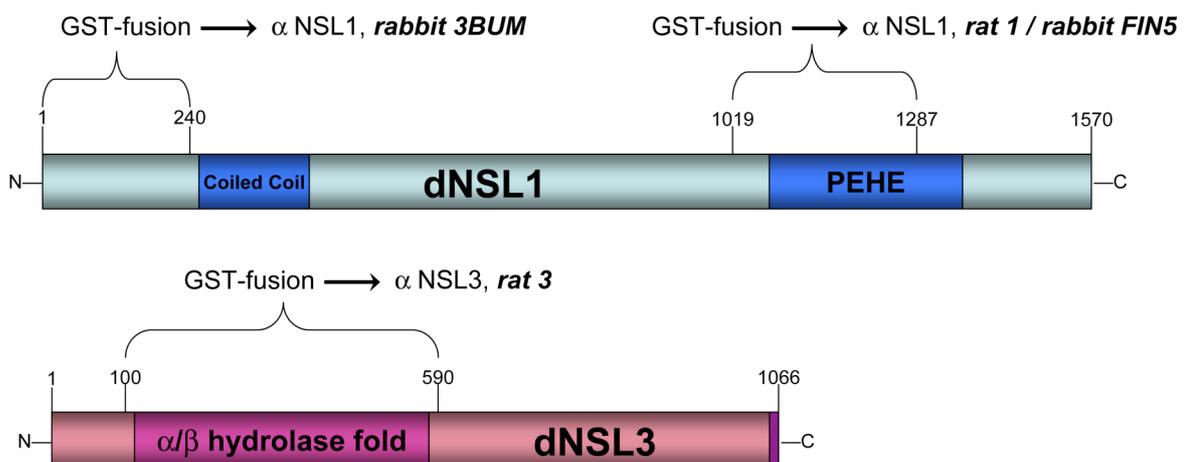


Figure 26-1: Domain structure and epitopes for the generation of NSL1- and NSL3-specific antibodies

Amino acid positions for fragments that were expressed as GST-fusions are denoted above the corresponding protein.

NSL3-specific antibodies were produced only in rats and showed a band at around 140kDa when used in western blot applications (see Figure 26-2 A).

The further description will focus on the generation of NSL1 specific antibodies as this turned out to be the more challenging task. It was noted that during the immunization procedure the antibody titers for NSL1 antibodies in rats were

decreasing already after the first bleed. Nevertheless, the serum could be used for detection of a western blot band around 240kDa (Figure 26-2 B, lanes 1-4). The contrasting calculated molecular weight of approximately 170kDa for NSL1 was not particularly astonishing as it can often be seen that chromatin-associated proteins are separated at a much higher molecular weight on SDS-PAGE. For example, MSL-1 (calculated 114kDa) produces a band at 170kDa on western blot and MSL-2 (calculated 85kDa) runs at 120kDa.

The detection of recombinant NSL1 protein purified from SF9 cells showed a single band at the indicated position (see Figure 26-2 B, lanes 3 and 4). Nevertheless, it was very surprising that the NSL1 western signal was not decreasing upon NSL1 RNA interference (RNAi) in the *Drosophila* S2 cell line – the knockdown being confirmed by qRT-PCR and its apparent cellular phenotype. Only after very extensive protein separation on low-percentage gels it could be realized that the NSL1 band on western of S2 total cell extract was masked by another band of nearly the same size which was cross-reacting with the NSL1 antibody (Figure 26-2 B lanes 1 and 2, Figure 27-1 upper blot). However, this problem of the rat antibody recognizing two protein bands at the same position was only apparent when total cell extracts were western blotted. The NSL1 rat antibody could also be used for coimmunoprecipitation (CoIPs) experiments from *Drosophila* embryonic nuclear extract. Reassuringly, this was resulting in a single NSL1 western blot band (compare Figure 28-1).

The same antigen (GST-PEHE) was as well injected in rabbits. The produced NSL1 rabbit antibody (FIN5) recognized a western band at the expected 240kDa position (Figure 26-2 C, lanes 1-4) when tested on nuclear extracts or recombinant protein. This rabbit antibody was as well suitable for selectively immunoprecipitating NSL1 from embryonic nuclear extracts.

The last attempt to make a multipurpose antibody was the immunization of rabbits with an N-terminal NSL1 fragment (aa1-240) which was also expressed as a GST-fusion (Figure 26-1 and Figure 26-2 F, lane 3). Therefore this antibody would recognize a completely unrelated epitope of the NSL1 protein.

Using the immune sera (rabbit 3BUM) on western blot resulted in a doublet band at the expected position for S2 and Kc whole cell extract (Figure 26-2 D, lanes 1-4). Kc cells were showing an additional band above 250kDa (Figure 26-2 D, lanes 1,2). The same pattern was observed for embryonic (lane 1) and nuclear (lane 2)

S2 cell extract (Figure 26-2 E). Interestingly, probed on recombinant protein the antibody detected not only a single 240kDa band but also another one around 130kDa (see Figure 26-2 D, lanes 3 and 4).

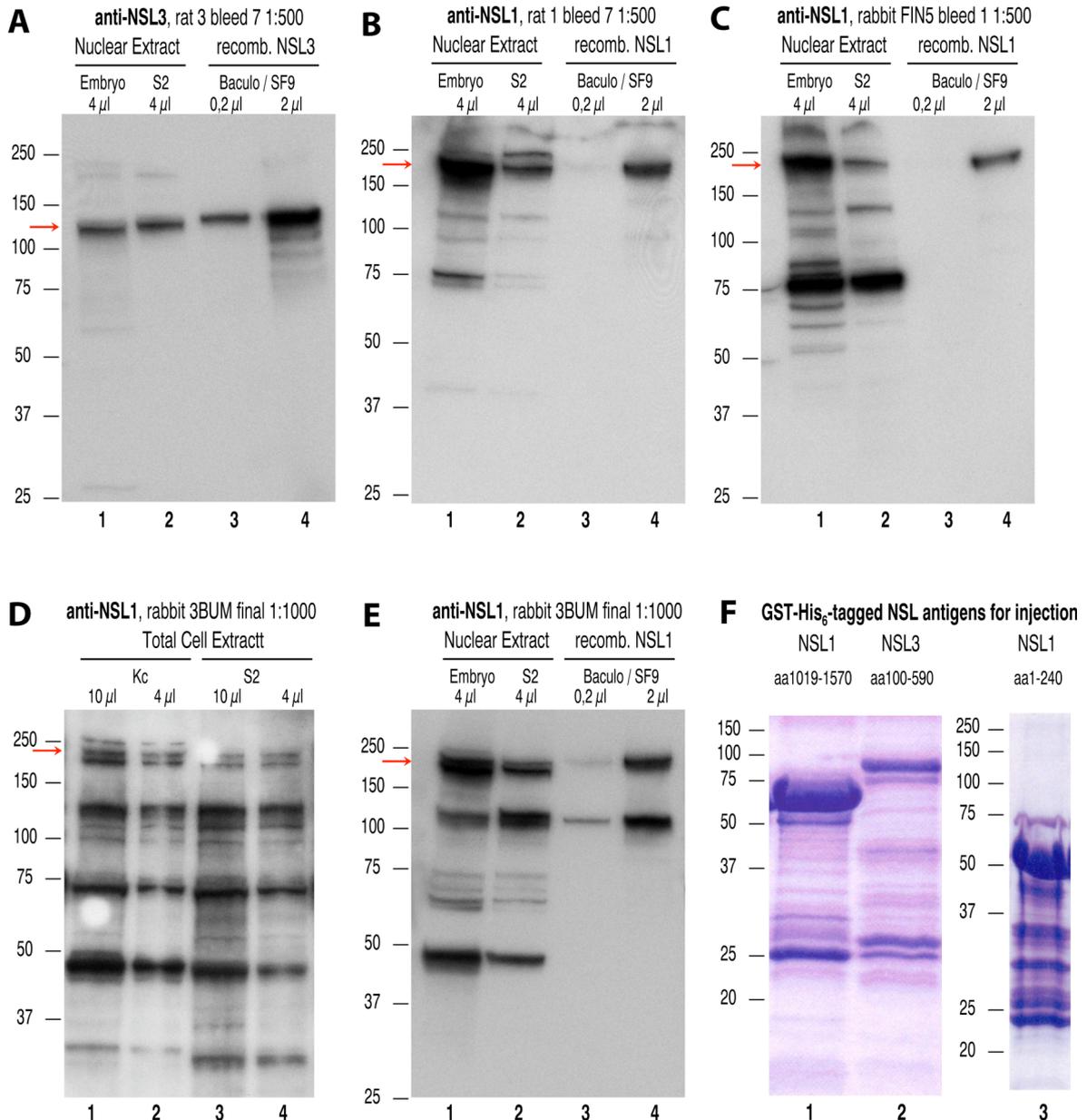


Figure 26-2: Western blot characterization of NSL1- and NSL3-specific antibodies

(A) anti-NSL3 antibody (rat3), probed on *Drosophila* embryonic nuclear extract, S2 cell nuclear extract and two different dilutions of recombinant full-length NSL3 protein from SF9 cells.

(B,C) anti-NSL1 antibodies (rat1, FIN5), probed on *Drosophila* embryonic nuclear extract, S2 cell nuclear extract and two different dilutions of recombinant full-length NSL1 protein from SF9 cells.

(D,E) anti-NSL1 antibody (rabbit 3BUM), probed on dilutions of Kc and S2 total cell extracts (D) and on *Drosophila* embryonic nuclear extract, S2 cell nuclear extract and two different dilutions of recombinant full-length NSL1 protein from SF9 cells (E).

(F) Coomassie-stained gels of purified GST-His₆-tagged NSL antigens for antibody production. Amino acid positions for NSL fragments are indicated (lane1: NSL1-PEHE, lane2: NSL3 hydrolase domain, lane3: NSL1 N-term). The banding pattern below the purified proteins shows typical degradation products of the GST-moiety.

→ Red arrows mark positions of full-length NSL1 (and NSL3) proteins as detected by the antibodies

It has to be noted that the rat and the rabbit (FIN5) antibodies that were raised against the PEHE domain of *Drosophila* NSL1 were also showing a major degradation band at 75-80kDa when used for total cell extract on western blots. In contrast, the rabbit antibody (3BUM) which was produced by injection of the N-terminal fragment of NSL1 regularly recognized additional bands at around 45kDa and approximately 130kDa, the latter one showing up even with recombinantly expressed full-length NSL1 protein.

It is very likely that these bands represented degradation products of the NSL1 protein. The distinct patterns that became apparent by the usage of antibodies with different epitope specificity could be produced by stable fragments of NSL1 degradation. This would be an interesting observation, because one could envisage a regulatory mechanism for the protein by self- or protease-mediated cleavage.

Other explanations for these 'contaminating' bands would imply impurities in the samples tested or the detection of other cross-reacting proteins.

27 Intracellular localization of the *Drosophila* NSL1 protein

The question of NSL1 protein's subcellular localization was tackled in different ways. First of all, Schneider S2 cell extracts were prepared and separated into nuclear and cytoplasmic fractions. The samples were western blotted and probed with NSL1-specific antibody (Figure 27-1).

27.1 NSL1 is a nuclear protein

NSL1-specific signals could only be detected in the nuclear fraction. The upper band that was also appearing in the cytoplasmic fraction was a result of an antibody crossreaction (denoted by a blue star). Probing the membrane with antibodies against other nuclear control proteins (WDS, NXF1) served as a quality control for the proper fractionation in nuclear and cytoplasmic extracts.

Moreover, NSL3 was also seen to exclusively localize to the nucleus with no sign of cytoplasmic NSL3 occurrence.

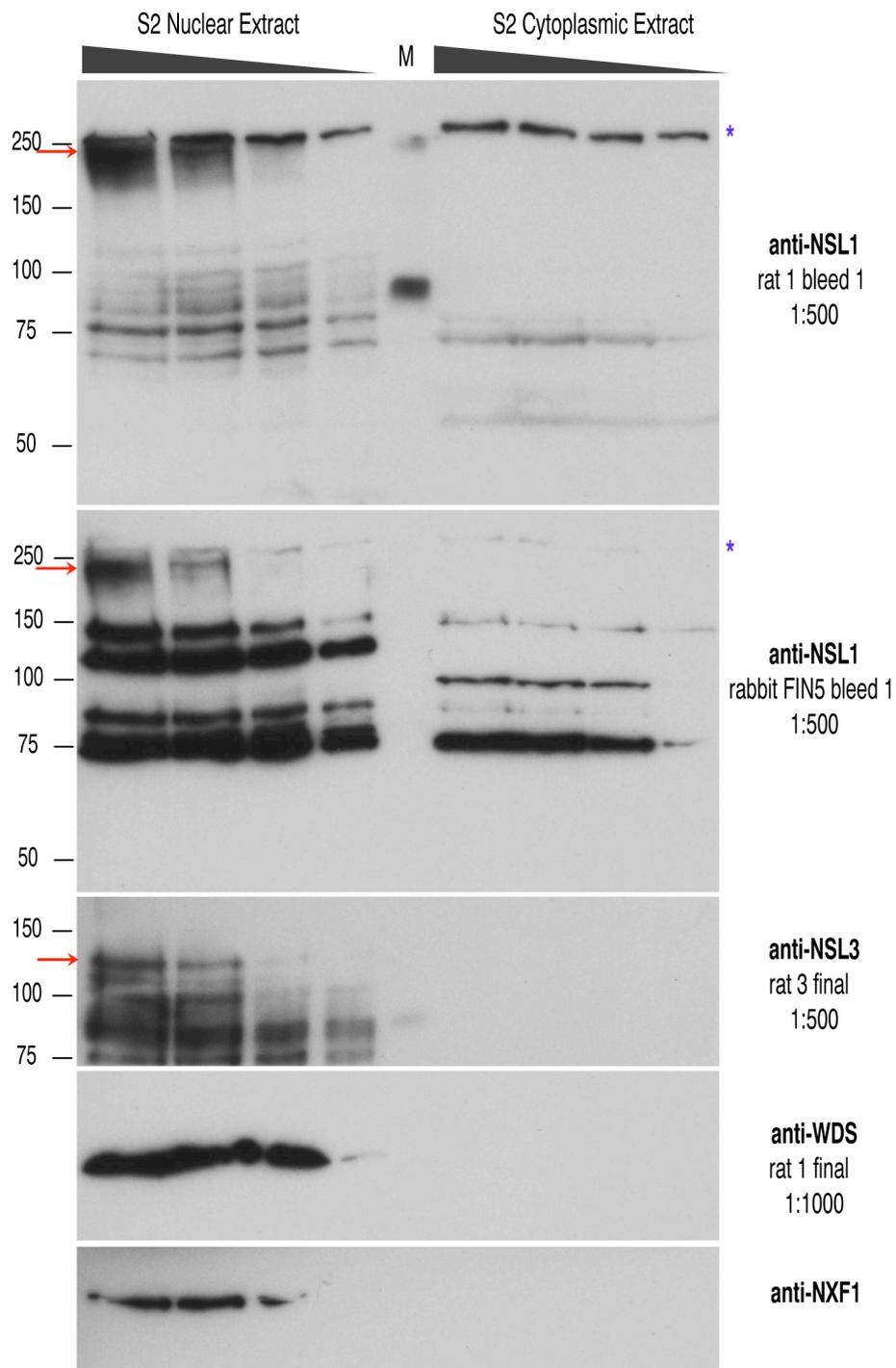


Figure 27-1: Analysis of NSL localization with S2 cell nuclear and cytoplasmic extracts

S2 cell extracts were separated in nuclear and cytoplasmic fractions and titrated amounts checked by western blot with the indicated antibodies. Red arrows (→) denote the position of NSL1 (two different antibodies) and NSL3 proteins. Blue stars (*) indicate cross-reaction with NSL1 antibody. Anti-WDS and anti-NXF1 serve as controls for nuclear localization.

To confirm this result, antibodies were used to immunostain fixed S2 and Kc cells and the pattern observed by confocal microscopy. Also here, the NSL1 protein localized to the nuclear compartment. The nucleolus was always excluded from the staining. As well, no cytoplasmic staining could be observed with the immunoaffinity-purified NSL1 antibody (Figure 27-2 A). The nuclear rim was counterstained with an antibody recognizing the nucleoporin NUP153.

Additionally, an N-terminal fusion of EGFP was made to the NSL1 full-length protein. The construct was introduced in S2 cells and the EGFP-activity monitored after two days of expression. The EGFP-NSL1 localized exclusively to the nucleus – with an exclusion of the nucleolus – whereas the EGFP transfection control showed nuclear and cytoplasmic staining (see Figure 27-2 B).

In conclusion, the results shown, argue unequivocally for a nuclear (not nucleolar) localization of the NSL1 protein.

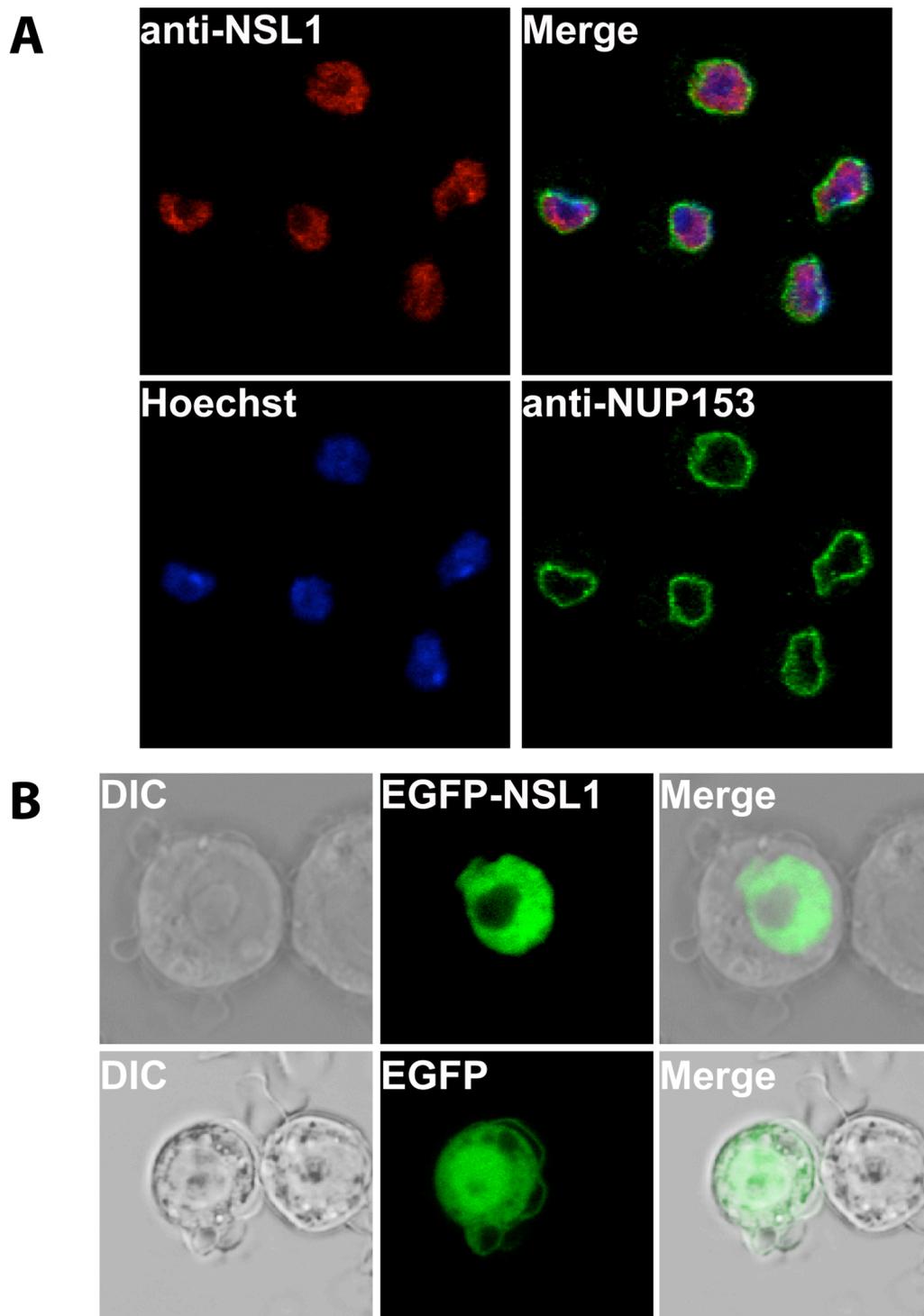


Figure 27-2: Subcellular localization of NSL1 protein in *Drosophila* S2 cells

(A) NSL1 staining with affinity-purified anti-NSL1 (rat 1) antibody (red). Nuclear rim staining with anti-NUP153 (green), DNA-counterstaining with Hoechst (blue)

(B) upper panel: EGFP-tagged NSL1 is restricted to the nucleus and is excluded from the nucleolus

lower panel: EGFP-transfected control cells show overall staining.

27.2 NSL1 localizes to chromosomes

The next question to answer, dealt with the subnuclear localization of the NSL1 protein. Whether it was bound to chromatin or distributed across the nucleoplasm. For that purpose, polytene chromosome squashes from 3rd-instar *Drosophila* larvae were prepared and proteins localizing to them detected with antibodies.

Firstly, the pre-serum for the NSL1 rabbit antibody did not show any background staining (Figure 27-3 A). The staining pattern that was generated by the NSL1 antibody was dispersed across all of the chromosomes (Figure 27-3 B). No X chromosome specific staining could be detected as it can be seen by staining for the MSL proteins. The NSL1 protein localized to a great number of interbands that were equally distributed across the entire length of the chromosomes. Telomeric staining was also observed (Figure 27-3 C) – mostly for one of the chromosomes.

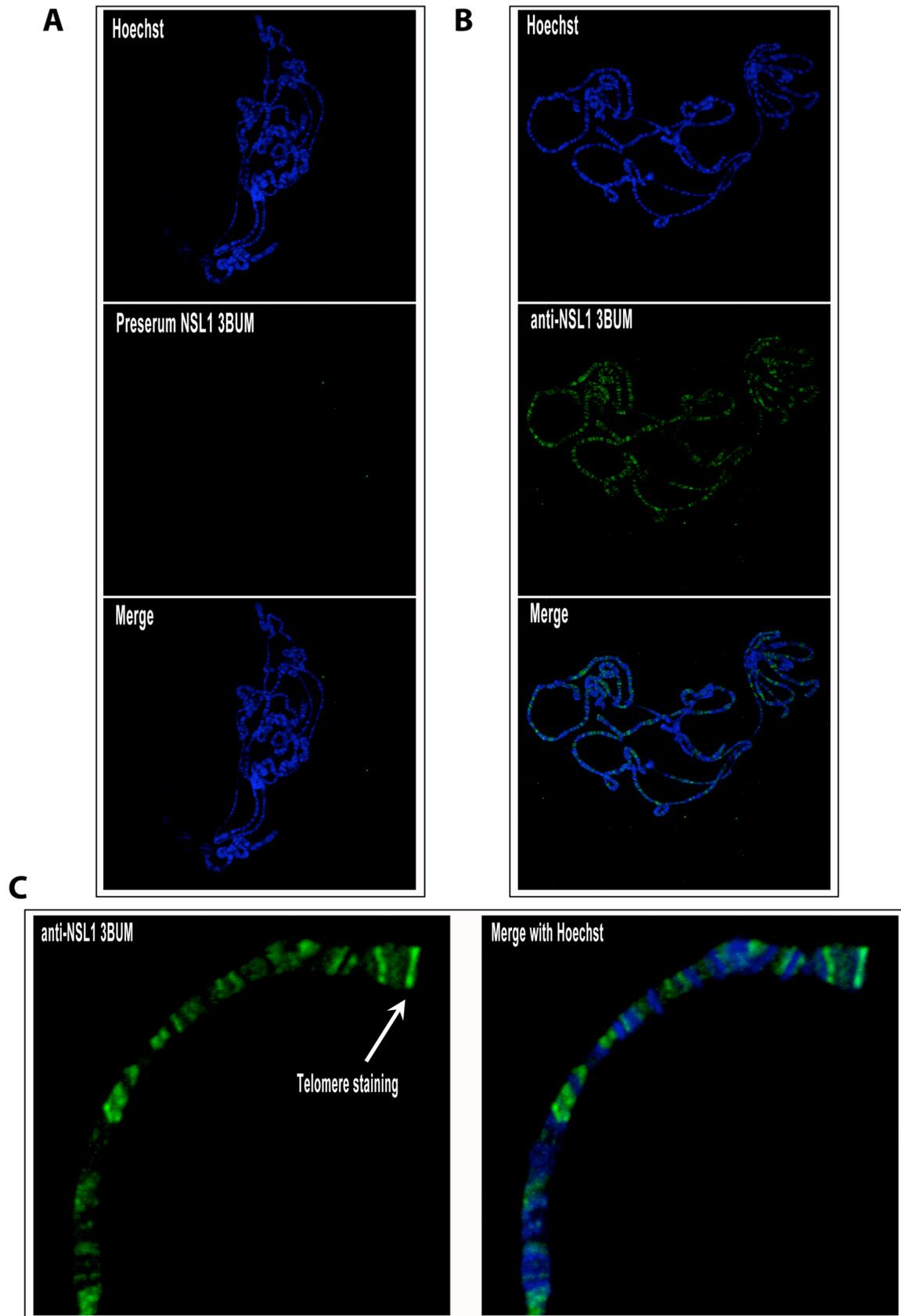


Figure 27-3: Localization of NSL1 protein on polytene chromosomes of *Drosophila* 3rd instar larvae

(A) No chromosome staining observable with NSL1 pre-serum (rabbit 3BUM) (B) Anti-NSL1 (rabbit 3BUM) immune serum stains all of the chromosomes (C) Zoom on chromosome end with telomeric staining.

28 Who interacts with whom – Coimmunoprecipitations reveal the dichotomy

The former purification of MOF-containing complexes in the lab identified not only the MSL proteins as MOF-interacting proteins but also a number of other proteins that have not been earlier breaking the surface in the context of dosage compensation complex (DCC) analysis (Mendjan et al., 2006). Amongst them the uncharacterized NSL proteins of which two are discussed here. It was not obvious from the beginning if these proteins would be part of the classical DCC or if they would only associate with the MOF protein to form (an)other complex(es).

A means to test for protein-protein interactions is to observe the coimmunoprecipitation pattern from cellular extracts. Antibodies that are directed against single proteins of interest are used to bind these proteins to a solid phase that can be washed. Provided that other proteins specifically interact with this protein bait they will be co-purified from the extract and can be detected by western blot.

Consequently, antibodies against NSL1, MOF and MSL-1 proteins were used for the coimmunoprecipitation experiments, using *Drosophila* embryonic nuclear extract as a source. Salt concentrations similar to the ones used previously for the MSL complex purifications were applied. The observed interactions are summarized in Figure 28-1. Parallel immunoprecipitations (IPs) with the corresponding pre-immune sera were performed to control for specific-enrichment. Cross-reactions with IgG-chains are marked with an asterisk.

For NSL1, the rat antibody which was raised against the PEHE domain was used. It reliably precipitated one single band for full-length NSL1 from the embryonic nuclear extract (lane 4, upper), whereas the pre-immune serum (lane 3, upper) did not show any background. The efficiency of NSL1 IP was estimated to be around 2.5% when compared to NSL1 signal in input lanes 1 and 2.

Cycles of repeated stripping and western blotting with different antibodies on the same blot revealed which proteins were interacting and which did not interact. MOF and MSL-1 IPs were performed in the same way and showed specific enrichment of 5 and 10% respectively.

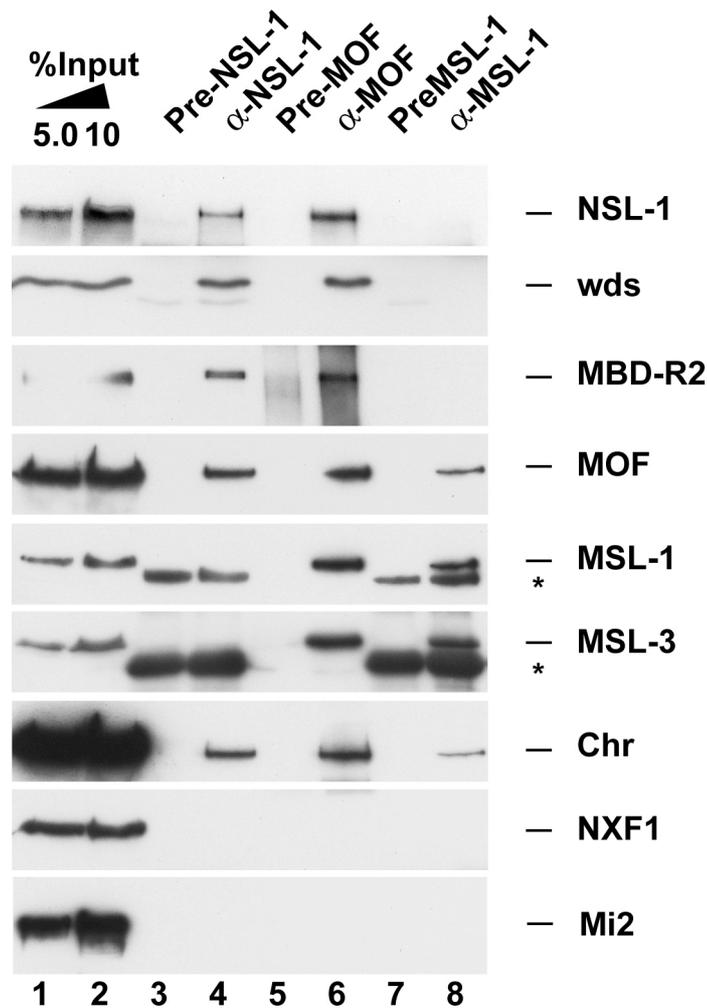


Figure 28-1: MOF associates with two distinct complexes

Immunoprecipitations (IPs) from *Drosophila* embryo nuclear extract using antibodies directed against NSL1 (αNSL1, lane 4), MOF (αMOF, lane 6) and MSL-1 (αMSL-1, lane 8). Specific enrichment was controlled for by performing parallel IPs with the corresponding pre-immune sera (lanes 3, 5 and 7).

IP efficiency can be estimated by comparison with titrated amounts of input nuclear extracts (lane 1 = 5% input, lane 2 = 10% input). * denotes IgG-band. Published in (Mendjan et al., 2006) – Figure 2B.

It became clear that MOF interacted with essentially all the tested candidates. This served as a confirmation of the previous purification that was done by affinity-tagging the MOF protein. However, for NSL1 and MSL-1 a non-overlapping set of interacting proteins was identified, apart from the already mentioned interaction with MOF. The NSL1 protein showed specific interactions with MOF, WDS (will die slowly), MBD-R2, Chromator (Chr) (Figure 28-1) and the NSL3 protein (not shown here). On the other hand MSL-3 (Figure 28-1) and MSL-2 (data not shown) coimmunoprecipitated with MSL-1 and a substoichiometric interaction between MSL-1 and Chr (<1% of input) could be observed as well.

This observation shaped the hypothesis that there might be more than one MOF-containing multi-protein complex in *Drosophila* and mammals. The idea was further strengthened by the separation of nuclear extracts on glycerol gradients where the intact complexes were separated by size. The isolated fractions were analyzed by western blot (Mendjan et al., 2006). Briefly, NSL proteins were fractionating in the higher-molecular weight range compared to MSL proteins. The MOF protein could be found in all the fractions (NSL and MSL) arguing for its involvement in both complexes.

Another piece of evidence that could also explain how the integration of MOF in both complexes could happen mechanistically, came from the computational analysis of the domain structures of NSL1 and MSL-1 (Marin, 2003). It was only these two proteins in the *Drosophila* proteome that seemed to possess a so called PEHE domain. The conserved region was called PEHE domain because of the identity of four characteristic amino acidic residues (P, E, H and E) in all the identified sequences (Marin, 2003). Interestingly, it was previously shown that MOF could directly interact with the PEHE domain of MSL-1 (Morales et al., 2004). The same holds true for the interaction of the hNSL-1 PEHE domain with hMOF (Mendjan et al., 2006).

It seems questionable if the PEHE should be called a 'domain'. The definition of a protein domain was described originally as the stable unit of protein structure that can fold autonomously (Wetlaufer, 1973). Others defined protein domains as units of compact structure (Richardson, 1981). The common theme is that a compact structural domain is likely to fold independently within its structural environment. Previously, I expressed the *Drosophila* MSL-1 PEHE domain in bacteria to explore its structure. The recombinant MSL-1 fragment was analyzed by nuclear magnetic resonance (NMR) and it turned out that the protein was not folded. It could only be speculated if this was a common feature of the PEHE or if it was only the case for the MSL-1 PEHE fragment. Yet, algorithm based predictions for disorder or globularity in protein sequences (GlobPlot) were also not assigning an intrinsically structured part for the protein's C-terminus where the PEHE is located. Due to the fingerprint-like presence of the proline, glutamic acid and histidine residues in the computationally identified region of homology one could speak of a PEHE motif. It is very likely that the PEHE 'motif' could get structured upon interaction with its

partner (MOF) and would adopt a so called induced-fit configuration. For simplicity reasons I will stay with Marin's nomenclature of the PEHE 'domain' in the following paragraphs although one should keep in mind the just mentioned considerations.

29 The *Drosophila* NSL1 PEHE domain mediates MOF interaction

To show the physical interaction of the *Drosophila* NSL1 PEHE domain with MOF, I co-infected SF9 cells with baculoviruses for the production of HA-tagged MOF and either full-length His₆-tagged NSL1 or His₆-NSL1-PEHE comprising fragments. It could be demonstrated that HA-MOF was not only binding to the full-length His₆-NSL1 protein (bait) but also to the His₆-PEHE domain fragment (aa1065-1415) alone (see Figure 29-1 A, lanes 1 and 2). Furthermore, an even smaller part of the His₆-PEHE (aa1065-1286) was sufficient to co-purify HA-MOF from SF9 cell extracts (lane 3).

The direct interaction between NSL3 and MOF protein could not be detected by these pull-down experiments (Figure 29-1 A, lane 4). Most probably they need a bridging factor – like NSL1 - to be incorporated in a complex. This result is substantiated by additional baculovirus reconstitution experiments that were performed in the lab by Herbert Holz and showed that MOF and NSL3 are not directly interacting.

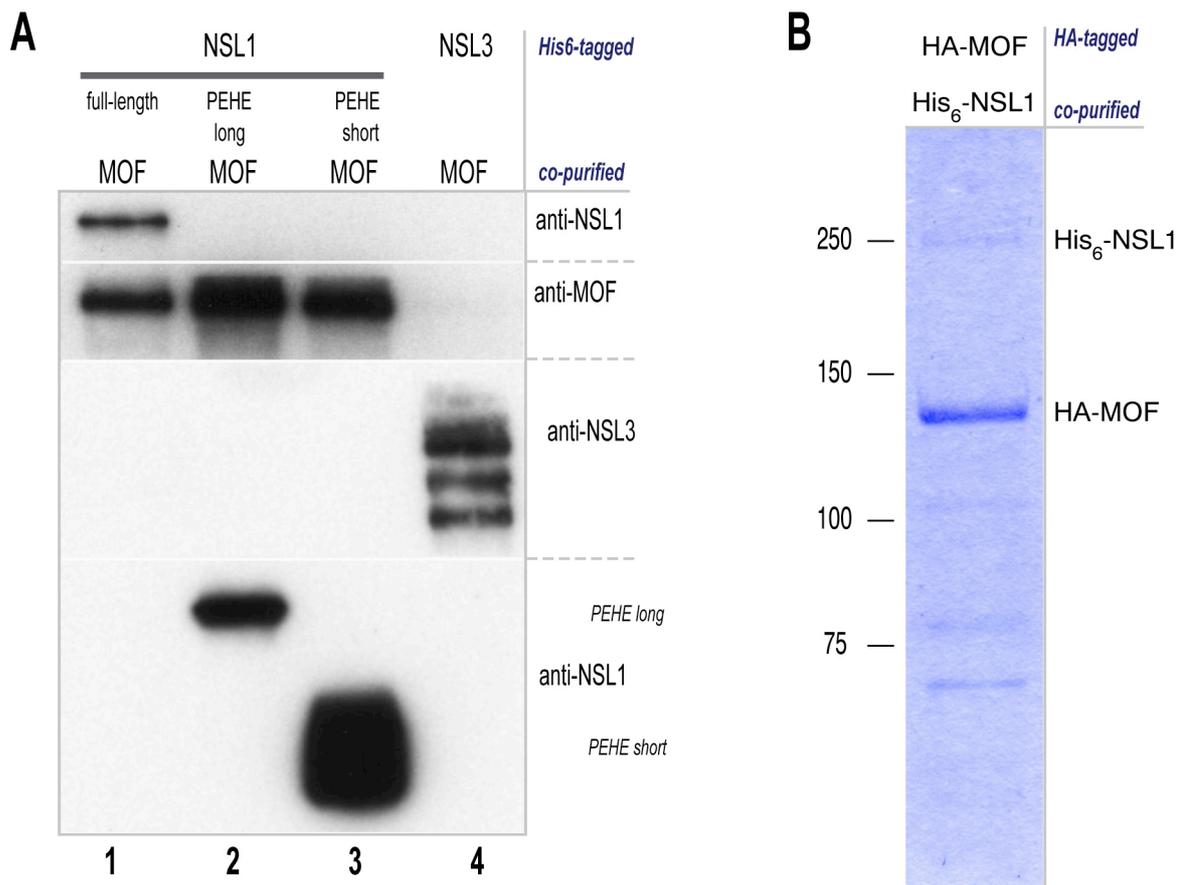


Figure 29-1: Pull-down experiments to test for NSL1 and MOF interaction

All recombinant proteins were expressed by infecting SF9 cells with the corresponding baculo viruses.

(A) Western of MOF pull-down with tagged full-length NSL1 (lane 1), long PEHE fragment aa1065-1415 (lane 2), short PEHE fragment aa1065-1286 (lane 3) but not with NSL3 (lane 4).

(B) Coomassie of full-length His₆-NSL1 co-purification with HA-MOF (bait)

The reverse experimental setup yielded the same results for MOF's interaction with the PEHE domain of NSL1. Here, HA-tagged MOF was used as the bait to co-purify His₆-NSL1-PEHE (data not shown).

In addition, Figure 29-1 B shows the co-purification of full-length His₆-NSL1 with HA-MOF (bait). The amount of purified MOF protein surmounts the quantities of co-purified NSL1 protein as MOF was used as the bait protein for affinity purification. This data was contributed by Herbert Holz, who was extensively testing the reconstitution of protein-protein interactions using our baculovirus-expressed proteins.

Considering the published interaction between MSL-1 and MOF (Morales et al., 2004), this suggests that the presence of the PEHE domain could represent a possible 'switch' for the exclusive interaction of either NSL1 or MSL-1 with the

MOF protein. Since MOF and the NSL proteins are present in males and females, it raised interesting questions about additional functions of MOF containing complexes in *Drosophila*.

30 Affinity-purification and analysis of the *Drosophila* NSL complex

The CoIP experiments were suggesting to us that MOF either interacted singularly with some of the newly identified proteins or that at least one additional MOF-containing multi protein complex existed which would be distinct from the MSL complex. In the light of the results from the glycerol gradient fractionations we favoured the hypothesis of an independent complex.

In an effort to further characterize the protein composition of this potential new complex and to study the enzymatic activity of MOF in the association with the new proteins we decided to purify the complex from *Drosophila* Schneider cells (S2). The initial purification strategy implied tagging the full-length NSL1 protein with a tag for Tandem Affinity Purification (TAP-tag). This method was originally developed for the purification of multi-subunit protein complexes from yeast (Rigaut et al., 1999) and was subsequently extended to complex purifications from other organisms, including *Drosophila* (Forler et al., 2003; Mendjan et al., 2006). The TAP-tag consists of a moiety of *Staphylococcus aureus* proteinA (zz-tag), a protease cleavage site for tobacco etch virus (TEV) protease and the calmodulin-binding peptide (CBP). The tandem affinity purification involves essentially three steps which will be shortly described in the following section: Extracts from cells that are stably expressing the TAP-protein fusion construct are incubated with immunoglobulin G (IgG)-sepharose beads which can bind the zz-tag. The bound and washed protein complexes are then released from the column by digestion with TEV protease and subsequently purified on the second column which consists of calmodulin-coated sepharose beads. Complexes can either be eluted from the beads natively (EGTA) or under denaturing conditions with SDS loading buffer.

I decided to fuse the TAP-tag to the N-terminus of the NSL1 protein and express the fusion protein under the control of a shortened actin5C promoter. S2 cells were

chosen for expression because of the ease of generating polyclonal cell lines which stably express the target protein.

After having generated the stable TAP-NSL1 expressing S2 cell line, nuclear extracts were prepared and small scale purifications performed to test for its experimental potential. The first step of the TAP procedure proved to be efficient as the fusion protein could be bound to the IgG-column in reasonable amounts (Figure 30-1 lane2). However, it turned out that the limiting steps of the purification were the TEV-elution of the complexes and the binding to the second column. Even after being able to release a small fraction of fusion protein from the IgG-column (lane 3) it was not possible to achieve binding to the calmodulin column. It turned out that the calmodulin binding step was not efficient to bind the TAP-NSL1 as this was tested also without prior purification on the first IgG column (Figure 30-1, lanes 4, 5 and 6). Extensive optimization trials for each step were not leading to a fundamental improvement of the purification results even when combined with glycerol gradient fractionations to pre-clear the extracts and to further reduce the background.

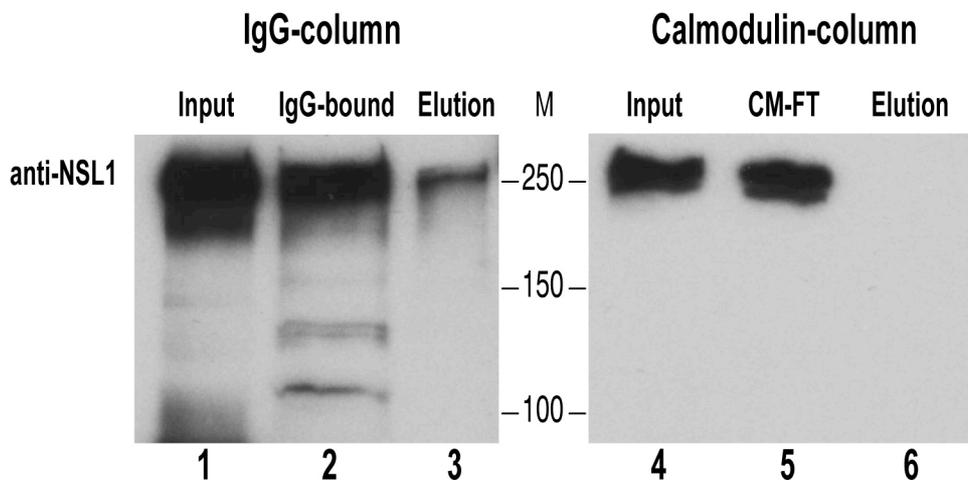


Figure 30-1: Western blot analysis of TAP-NSL1 binding to and elution from IgG- and Calmodulin affinity resins

TAP-NSL1 nuclear S2 cell extract (lanes 1 and 4) was incubated with IgG-beads (left) or Calmodulin-beads (right) to observe the binding efficiency of the TAP-NSL1 fusion protein (lanes 2 and 5). Elutions were performed with TEV protease (left, lane3) or with 2x SDS-loading buffer (right, lane6). CM-FT = Calmodulin flow through, M = protein size marker

A change of purification strategy was necessary to solve the problem and we decided to stay with a two-step purification method. The use of other affinity tags and the independence of a protease cleavage site for elution were desirable prerequisites. The tags chosen to be used in a tandem-like fashion were the

FLAG- and HA-tags. Therefore, an oligo coding for single FLAG- and HA-tags was inserted in between the existing TAP-tag and the NSL1 coding sequence. The resulting NSL1 construct was now N-terminally triple tagged with a TAP tag followed by one FLAG- and one HA-tag. I will refer to this construct as TFH-NSL1. This expression plasmid was transfected in S2 cells and a stable polyclonal cell line established. Firstly, the suitability of these newly introduced tags for purification was tested individually. Nuclear extracts from the TFH-NLS1 cell line were bound either to FLAG-M2 resin or to HA-agarose and eluted with the corresponding peptides. Consistent binding and elution was observed for both tags. The second aim was to combine the two purification steps in a linear setup. For the FLAG- and HA-tags it was possible to use either tag for the first column as the peptides used for elution from one affinity resin were not interfering with the binding to the next resin. The pilot experiments showed that it was possible to elute from each affinity column and thereby the subsequent purifications could be placed in an arbitrary order.

From there on, the complex purifications were performed as described in detail in the materials and methods section. A brief summary reads as follows. Anti-FLAG M2 beads were used to isolate the TFH-NSL1 and associating proteins from S2 cell nuclear extract. Retained complexes were eluted with FLAG peptide and the pooled elution fractions directly bound to anti-HA beads in the second purification step. Elution from this resin was done according to the intended use of the eluate: HA peptide was used to elute intact complexes for enzymatic assays. A more stringent SDS-elution was performed for the purpose of direct gel analysis. The purification proved to be effective at the ionic conditions (150mM KCl) that were applied throughout the whole purification procedure (binding, washes and elution). All the steps involved were shortened to a minimum in order to preserve the integrity and the potential (enzymatic) activities of the complexes.

Generally, enough material could be recovered from both columns to analyse the protein composition by mass spectrometry, by western blot or in enzymatic assays. As it turned out that a high degree of purity could be achieved already after elution from the first column (FLAG), both strategies were applied and samples analysed after 1-column or after 2-column purification.

30.1 Analysis of eluted proteins by western blot

A main goal of this project was to reveal the protein composition of the NSL complex. Having the affinity-purified NSL complex in hands we chose a combinatorial approach to identify its constituents.

The first part, which was accompanying the purifications, relied on tracking suspected interactors with available antibodies. A similar set of interacting proteins that were known from the CoIP experiments was followed by this means. For this purpose, FLAG-eluted complexes were subjected to western blot analysis and membranes probed subsequently with the indicated antibodies. Antibodies against abundant proteins, like RPD3, Tubulin and Lamin, were used to estimate the quality of the purifications.

The analysis revealed a comprehensive set of proteins which were eluting with the bait protein. First of all, the analysis confirmed the presence of the tagged NSL1 protein (Figure 30-2, lane 4 upper panel). Interestingly, at least by western blot analysis the endogenous NSL1 protein could not be found to co-purify with the bait. This might indicate that the NSL1 protein does not associate with other NSL1 molecules and could function as a monomer in the NSL complex. However, this is so far the only evidence for NSL1's monomeric existence and more detailed analysis will be necessary to investigate this further.

When comparing with the amounts of protein in the input material (Figure 30-2, lanes 1 and 2) the elution fractions of NSL2 (panel 3 from top) and NSL3 (not shown) proteins showed strong enrichment, arguing for their tight association with NSL1.

The MBD-R2 and WDS proteins (Figure 30-2, panel 4 and 6 respectively) could be detected on the western blot and were consistently found in other NSL complex purifications.

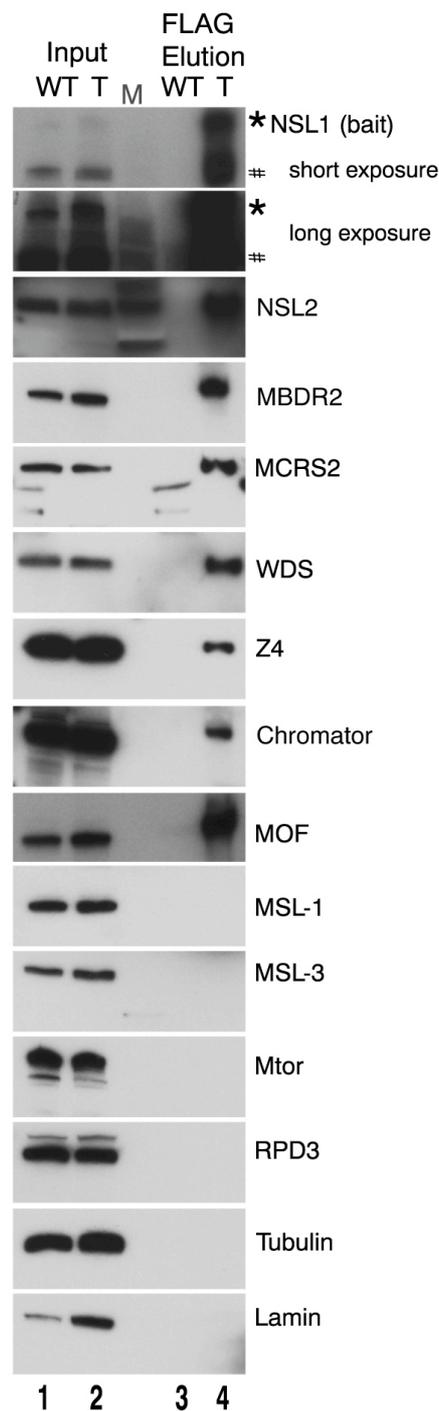


Figure 30-2: Western blot of FLAG eluted complexes probed with antibodies against identified complex members.

Lane 1: nuclear extract from wild-type S2 cells

Lane 2: nuclear extract from FLAG/HA-NSL1 S2 cell line

Lane 3: FLAG elution of purification from wild-type S2 extract

Lane 4: FLAG elution of purification from FLAG/HA-NSL1 S2 extract

M: Protein size marker; * indicates position of TFH-NSL1, # NSL1 degradation

Even though we could not conclude about the exact stoichiometric quantities of the interacting proteins by western blot, we found Z4 and Chromator co-purifying in amounts that were below their input signals (Figure 30-2, panels 7 and 8 from top). This substoichiometric association of Z4 and Chromator proteins that could be estimated from the western blot analysis most probably reflects their role as complex interacting partners rather than being core members of the NSL complex. Evidence for this is also contributed by immunofluorescence co-staining experiments of polytene chromosomes from salivary glands (see Figure 30-4).

The lack of the MSL proteins (MSL-1, MSL-3 in Figure 30-2, for MSL-2 data not shown) was opposed by the finding that MOF was consistently co-purifying with the NSL1 bait. This reinforced our hypothesis of the NSL complex being devoid of MSL proteins.

Interestingly, Mtor and NUP153, proteins that were previously found in MOF purifications (Mendjan et al., 2006), could not be detected in the TFH-NSL1 FLAG eluate. This suggested that this interaction is likely to occur via MOF or other members of the MSL complex.

30.2 Identification of NSL1 interacting proteins by mass spectrometry

The second part for the analysis of eluted proteins involved the use of mass spectrometry as a key technology for the identification of proteins from a mixture of unknown proteins. Two ways of mass spectroscopic analysis were selected: excision and analysis of individual silver-stained protein bands or analysis of total complex elutions. Results of both approaches will be discussed below. The identification of NSL-complex associating proteins by mass spectrometry was performed by Sven Fraterman (EMBL Heidelberg) and Adrian Cohen (NCLMS, Netherlands).

To ensure the consistency and quality of individual TFH-NSL1 purifications, electrophoresed elution fractions were visualized by silver-staining. The stained protein bands could then be excised, the proteins trypsin digested and analyzed by the very sensitive LC-MS/MS (Liquid Chromatography - Tandem Mass Spectrometry) method. Figure 30-3 (left) shows the silver-stained gel of a typical NSL1 complex purification. The eluted material was obtained by double-affinity

purification via FLAG- and HA-resins and was run side-by-side with the control purification from wild-type (WT) nuclear extract. Proteins identified by mass spectrometry are assigned next to the corresponding bands.

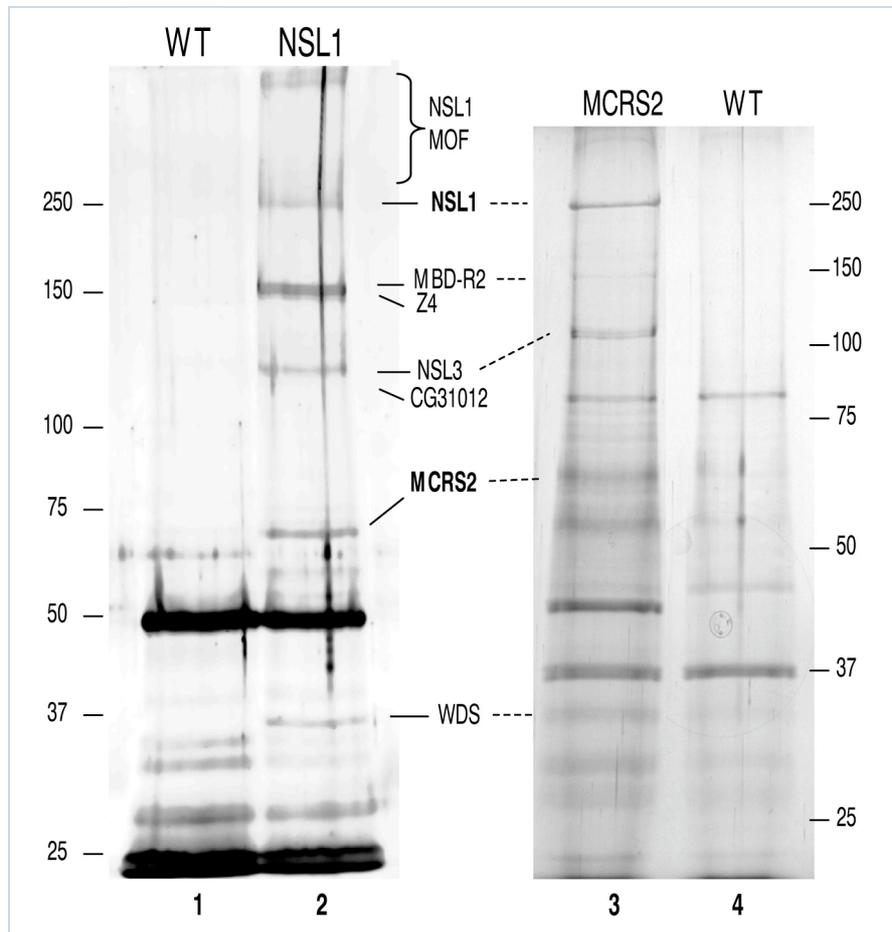


Figure 30-3: Silver stained gels and the proteins identified in NSL complex purifications from nuclear extracts of *Drosophila* S2 cells

Elution fractions were separated on 5-15% SDS-PAGE gradient gels and silver-stained

Lane 1: wild-type control for FLAG/HA purification

Lane 2: HA-elution fractions of TFH-NSL1 purification from a stable cell line expressing TAP-FLAG-HA-NSL1 protein

Lane 3: Elution fractions of TAP-MCRS2 purification from a stable cell line expressing TAP-MCRS2 protein

Lane 4: wild-type control for TAP purification

Figure 30-3 (right) additionally shows a complex purification via TAP-tagged MCRS2 that was performed by Iryna Zhloba, another PhD student in the lab. Both, NSL1 and MCRS2 purifications, led to the identification of the same proteins by LC-MS/MS. Dashed lines are pointing to the positions of proteins identified with tagged MCRS2.

The occurrence of mostly the same proteins co-purifying with NSL1 and MCRS2 led to the conclusion that these proteins were part of the same complex. Identification of MCRS2 protein in NSL1-derived complexes and vice versa further

substantiated this hypothesis and helped us to classify some of the proteins as core complex members (see Table 30-1).

	Identified proteins	number of experiments	LC-MS/MS (bands excised)	LC-MS/MS (total elution)
			Mascot scores (number of peptides)	Mascot scores (number of peptides)
Core NSL Complex	NSL1 (bait)	+++++	205(3), 260 (6), 122 (4)	275 (5), 36 (8)
	CBP (→ NSL1)	++	48 (1)	57 (1)
	NSL2	+		105 (1)
	NSL3	++++	68 (2), 666 (18)	114 (3), 52 (4)
	MOF	++	51 (1)	57 (1)
	MBDR2	+++++	424 (8), 694 (12), 168 (3), 405 (12)	328 (6), 28 (15)
	MCRS2	+++	60 (2), 97 (2), 422 (10)	
	WDS	+	306 (7)	
Associated Factors?	Z4	+++	424 (8), 135 (3)	95 (2)
	Chromator	+	104 (2)	
	Actin	+++	560 (10), 55 (2)	27 (13)
	CG5381	++	27 (1), 56 (1)	
	CG15415	++	617 (10), 451 (10)	
	CG31012	++	166 (4), 44 (1)	

Table 30-1: NSL complex members identified by mass spectrometry from TFH-NSL1 purifications

Mascot scores and number of identified peptides are shown. CBP = calmodulin-binding protein (signifies the presence of the bait, TAP-FLAG-HA-tagged NSL1)

In order not to miss out on potential interacting proteins, whole elution fractions from NSL complex purifications were analysed by LC-MS/MS without prior separation on a gel. This procedure allowed for identification of proteins that were not staining (or weakly) with the silver staining protocols.

Every sample analyzed was paralleled by a control purification sample from wild-type nuclear extract. By comparison with the wild-type control the specific enrichment could be demonstrated and contaminating proteins ruled out. Multiple

rounds of analyses were performed with independent purifications from different batches of S2 cell nuclear extract. The proteins identified with both approaches (gel-separated complexes and total elutions) are summarized in Table 30-1.

Not only the bait, TFH-NSL1, was detected with several peptides and high Mascot scores but also the other NSL proteins (NSL2, NSL3) were reliably identified. The analysis confirmed the interactions with MBD-R2, WDS, Chromator and Z4 that were observed with the western blot and previously by CoIP experiments.

The co-purification of the MOF enzyme with NSL1 was also affirmed by mass spectrometry. Reassuringly, the tight association of the microspherule protein 2 (MCRS2) with the NSL complex members could be validated.

Strikingly, already after FLAG-elution from the first column, no MSL proteins could be detected anymore in the eluate (Figure 30-2 and Table 30-1). This strengthened the hypothesis of an MSL-independent NSL complex.

A number of additional proteins were regularly co-purified in different purifications (Figure 30-3, Table 30-1). We were judging them as potential bonafide interactors of the NSL complex for several reasons, such as they were present with many peptides, exhibited good Mascot scores and were isolated repeatedly. This group comprised actin and three uncharacterized proteins: CG5381, CG15415 (predicted SMC domain) and a protein with three predicted SH3 domains (CG31012). So far they could not be ruled out as mere contaminants of the purification but will still have to prove their association with the complex in different experimental setups.

30.2.1 *Drosophila* MSL-1 and NSL1 – MOF's exclusive friends

In summary, the prominent feature was the complete absence of MSL proteins from this apparently distinct MOF-containing complex. This was arguing for the model of MOF interaction via the PEHE-domain of MSL-1 or NSL1 in a mutually exclusive manner. It looks like MOF can only bind one of them – either MSL-1 or NSL1 – at one time and therefore 'selects' the set of proteins it interacts with. Therefore, it was interesting to speculate if the integration of MOF in the NSL complex would modulate its activity and could potentially influence the enzyme's specificity.

30.3 Description of identified *Drosophila* NSL complex members

The following paragraphs are dedicated to a short description of the proteins that were identified to be part of the NSL complex. Characteristics of NSL1, the bait protein for the complex purification, and MOF have been already discussed in the introduction and will therefore be omitted.

30.3.1 Confirmed members of the NSL complex

NSL2 (non-specific lethal 2) can be found under the accession number CG18041. The protein consists of 484 amino acids (53 kDa). Its molecular function and the biological processes in which it is involved in *Drosophila* are not known. The evolutionary conserved protein is expressed ubiquitously in mice and in *Xenopus laevis* embryos in the animal pole from stage IV onwards (Mata et al., 2003; Shim et al., 2000). Conventional software-based prediction of domains failed to identify any domains. Still, sequence analysis identified a conserved motif rich in cysteine and histidine residues (Taipale et al., 2005a).

NSL3 (non-specific lethal 3) is an evolutionary conserved protein that has three annotated polypeptides with differing length in *Drosophila melanogaster* (1001aa/110kDa, 1066aa/117kDa, 934aa/103kDa). Its function is still unknown. The protein contains a domain that has a fold which is similar to α/β -hydrolases. This α/β -hydrolase fold is common to several hydrolytic enzymes of widely differing phylogenetic origin and catalytic function (Ollis et al., 1992). Enzymes like esterases, lipases and peptidases are examples for members of this structural superfamily. The catalytic triad that is characteristic for these enzymes is altered in the NSL3 protein and it is therefore very likely that the protein is not enzymatically active. Taipale (Taipale, 2005) hypothesized that despite its non-functionality the catalytic center of the enzyme might still serve to bind the former 'substrate'. Of what nature the binding partner might be and if the protein really binds to it still has to be investigated.

This present work gives a hint that the NSL3 protein might be unstable without the NSL1 protein. The putative interdependency is described more in detail in the corresponding chapter 33.1.1.

MCRS2 (microspherule protein 2, CG1135) is a homologue of the human MCRS2 / MCRS1 protein and consists of 558 amino acids (61 kDa). It contains a FHA (fork-head associated) domain which has been implicated in phosphopeptide binding (Durocher et al., 1999; Durocher et al., 2002; Li et al., 1999).

The human MCRS2 protein was shown to be involved in telomerase inhibition. Together with its cell cycle dependent expression it was speculated that the protein might play a role in linking telomere maintenance to cell cycle regulation (Song et al., 2004). Its exact molecular function in *Drosophila* is still unclear.

MBD-R2 is a 1081 amino acids (119 kDa) comprising protein that can be found under the accession number CG10042. The protein has not been extensively characterized yet. Its interesting domain structure links to a variety of possible molecular functions. The protein possesses a MBD (methyl CpG binding) domain, two Tudor domains, a C2H2 zinc finger, a PHD finger and a THAP domain that was found to be a zinc-dependent sequence-specific DNA-binding domain (Clouaire et al., 2005).

The analysis of protein domains suggested a role in gene expression regulation for most of the MBD domain containing proteins in mouse and man (Roloff et al., 2003). MBD domains, for example, have been shown to bind methylated DNA and are involved in transcriptional repression in mammals (Bird, 2002).

However, in *Drosophila* DNA methylation is rare and it predominates only during early embryonic stages with a decrease at later stages (Lyko et al., 2000).

The Tudor domain was originally identified in the Tudor protein encoded by *Drosophila*. Subsequently it was found among other proteins involved in binding to RNA (Ponting, 1997) and to methylated lysines (Brahms et al., 2001; Huang et al., 2006).

WDS (will die slowly, CG17437) is a 361 amino acid (47 kDa) protein which contains seven WD-40 repeats (Hollmann et al., 2002). These repeats (also known as WD or beta-transducin repeats) are short motifs consisting of about 40 amino acids. Structural studies revealed that the seven repeats form a circularised beta-propeller structure (Sondek et al., 1996) that is frequently used to coordinate the assembly of multi-protein complexes.

WD-repeat containing proteins form a large family found in all eukaryotes. They are implicated in a plethora of functions as diverse as signal transduction, transcription regulation, cell cycle control and apoptosis (Neer et al., 1994).

WDS is evolutionary conserved and can be found from *Arabidopsis* to humans. The mammalian ortholog, WDR5, has been found to be able to bind to dimethylated H3K4 (Wysocka et al., 2005). Additionally, the purified human WDR5 protein complex was shown to contain not only the MLL1 methyltransferase but also hMOF was found to be part of the same complex (Dou et al., 2005). Due to the identified association of these two enzymes in one complex, Dou et al. (2005) speculated that this might serve as a molecular explanation for the closely correlated distribution of H3K4 methylation and H4K16 acetylation on active genes.

Chromator/Chriz (CG10712) is a protein of 926 amino acids (102 kDa) that contains a chromodomain. The protein is essential in *Drosophila* as could be seen from the analysis of P-element insertions (Rath et al., 2004).

During interphase, Chromator localizes to interbands on polytene chromosomes and co-localizes with the zinc-finger protein Z4 (Eggert et al., 2004; Gortchakov et al., 2005). Interestingly, Z4 was also found to co-purify with the NSL complex.

However, Chromator detaches from the chromosomes during mitosis and forms a spindle-like structure together with the Skeletor protein (Rath et al., 2004).

Z4 (CG7752) consists of 996 amino acids (110 kDa) and exhibits seven zinc fingers in its structure. Like Chromator, it localizes to interbands on polytene chromosomes. The protein is essential for fly development and it is believed to play a role in chromatin compaction as Z4 mutants show an overall decompaction of chromosomes and loss of interbands on chromosomes. (Eggert et al., 2004). The described interaction and co-localization with Chromator adds a degree of confidence that these proteins were specifically co-purified with the NSL complex. Still, the substoichiometric presence of these two proteins in the NSL complex purifications could hint towards an interaction with the holo-complex rather than being core complex members.

30.3.2 Newly identified members of the NSL complex

Another protein that was found in different purifications of the NSL complex is the **CG31012** protein. This uncharacterized protein is annotated in Flybase with four polypeptides (545aa, 537aa, 882aa, 635aa). It contains three predicted SH3-domains. These predicted domains are the only hint to what the function of this protein could be.

SH3 (src Homology-3) domains are small protein modules consisting of approximately 50 amino acids. They can be found in a variety of proteins such as intracellular or membrane-associated proteins, proteins with enzymatic activity, adaptor proteins and cytoskeletal proteins. It is not yet clear what the function of the SH3 domain is, but it is thought that it may contribute to diverse processes, e.g. increasing local protein concentration, changing protein's subcellular location or liaising proteins into large multiprotein complexes (Mayer et al., 1995; Morton et al., 1994).

Unfortunately no antibody was available to further characterize CG31012's association with the NSL complex. But, according to the nature of its predicted SH3 domains, it could be deduced that it fulfils a variety of roles. It is therefore very likely that it could take part in the regulation of proteins or the assembly of the multiprotein NSL complex. Further investigation will be necessary to assess its role in this context.

The 686 amino acids (75 kDa) long **CG5381** is a completely uncharacterized protein that was found in 1-column and in 2-column NSL complex purifications. No recognizable domains are predicted for this protein. Due to the non-existence of CG5381-specific antibodies the association of this protein with NSL core complex members could not be investigated further.

Actin, a protein that is ubiquitously expressed in all eukaryotic cells was found with multiple peptides in several purifications of the NSL complex. Although actin is often present as a common contaminant in purifications of nuclear complexes it is still an interesting candidate due to its involvement in various cellular processes. The Actin5c (376 amino acids, 41kDa) belongs to a family of highly conserved proteins that are involved in various types of cell motility. The cytoplasmic protein

is expressed in several isoforms that undertake multiple cellular functions. It is not only involved in cell motility (Pollard et al., 2003) and cytoskeleton organization but also in chromosome movement (Lenart et al., 2005) and cytokinesis (Pelham et al., 2002).

Taking into account the observations from the live cell imaging experiments with NSL1 RNAi cells (described later in chapter 33.3), it seems an interesting possibility that Actin might have a function in conjunction with the NSL complex, since the NSL1 knockdown causes severe chromosome segregation defects and proper cytokinesis is also impaired.

The **CG15415** protein (807 amino acids, 89kDa) was found in FLAG-purifications of the NSL1 complex. It is so far mostly uncharacterized. Its predicted SMC (structural maintenance of chromosomes) domain in conjunction with the observed NSL1 RNAi phenotype makes it a very interesting protein as it might be involved in different processes like chromatin condensation, chromosome cohesion or even DNA repair. The link is described in more detail in the section about NSL1 knockdown in cells (chapter 33.3.1).

30.4 NSL1 co-localizes with NSL complex members on polytene chromosomes

Immunofluorescence co-stainings with antibodies against proteins (e.g. MCRS2) that were found to tightly associate with the NSL1 protein revealed an almost perfect overlap in terms of localization to chromosome bands (Figure 30-4 A). For sure, the polytene banding pattern can only be indicative of protein co-localization as the bands themselves can comprise long stretches – up to megabases – of DNA. However, for some of the proteins, e.g. Z4, the co-staining pattern was not completely overlapping (Figure 30-4 B). Again, this argues for the hypothesis that Z4 might be one of the complex-associated proteins rather than being an integral part of the NSL1 complex.

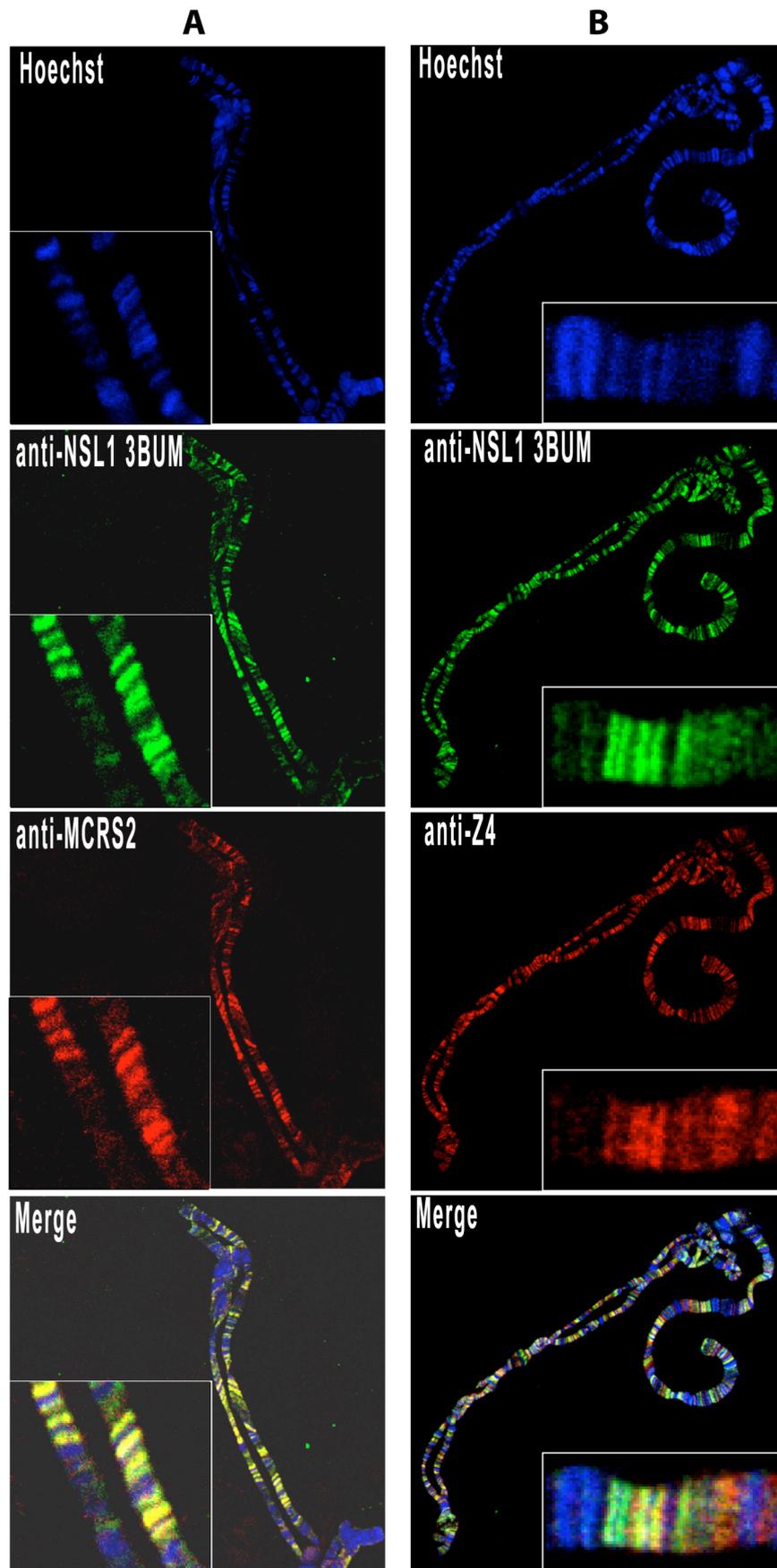


Figure 30-4: Fig. NSL1 co-localization on salivary gland polytene chromosomes

(A) Co-staining with NSL1 (green) and MCRS2 (red) antibodies. **(B)** Co-staining with NSL1 (green) and Z4 (red) antibodies. Zoomed regions are surrounded by a white box.

31 Biochemical characterisation of the histone acetyltransferase activity in the NSL complex

The next step that followed the successful purification was the biochemical characterisation of the NSL complex. Equipped with the knowledge about the protein composition of the complex we could have a look at known functions of the associated proteins. The first focus was on enzymatic activities that could be associated with this complex. A prerequisite for further testings *in vitro* was the fact that the NSL complex could be eluted from the purification column in an intact manner. Keeping the enzymatic assays in mind already at the beginning, the purification procedure was designed to be as gentle as possible. Taken together, these steps were promising to be able to preserve delicate biochemical properties such as enzymatic activities.

Most of the identified candidates did either not display any recognizable enzymatic activity or no biochemical functions were attributed to them yet because they were so far uncharacterized. At first glance the most obvious enzyme that was contained in the complex was the co-purified MOF protein. MOF is the histone acetyltransferase (HAT) that was shown to specifically acetylate lysine 16 of histone H4 (H4K16) in *Drosophila* (Akhtar et al., 2000a; Hilfiker et al., 1997; Smith et al., 2000). This remarkable selectivity of the enzyme was also shown by using the eluted MSL complex for histone acetylation *in vitro* (Mendjan et al., 2006).

31.1 MOF retains histone H4 specificity upon integration in the NSL complex

The experimental setup for carrying out the histone acetyltransferase assays (HAT-assays) was chosen such, that the substrate for the enzyme reflected the natural substrate of MOF as close as possible. *In vivo*, MOF acts on nucleosomes that are packed in higher order chromatin structures. Consequently, for the HAT-assay polynucleosomes were reconstituted. A linear DNA array served as the scaffold for the loading with histone octamers that were produced from recombinant *Xenopus* histones.

The eluted NSL complexes were incubated with this substrate in the presence of radioactive tritium (^3H) labelled acetyl-CoenzymeA (acetyl-CoA). The acetyl-CoA

served as the donor of the acetyl group to be transferred on the histone substrate. The treated histones were separated by SDS-PAGE and subsequently autoradiographed to reveal the enzyme's specificity.

Firstly, reconstituted polynucleosomes were tested with nuclear extracts. Crude nuclear extracts produced strong acetylation of all histones (Figure 31-1, lanes 1 and 2) due to the likely presence of multiple HATs with differing specificities. Furthermore, the intensities of signals produced by wild-type and TFH-NSL1 cell line derived nuclear extracts were comparable.

The quality of the purification was reconfirmed by the fact that elutions from wild-type purifications did not result in measurable acetylation (Figure 31-1, lane 3).

When testing the eluted NSL complexes in the described assay, they were indeed showing enzymatic activity. HAT-assays on polynucleosomes with wild-type histone tails were yielding a strong signal for histone H4 and weaker signals for histones H2A/B and H3 (Figure 31-1, lane 4).

Thus, MOF clearly exhibited a preference for acetylation of histone H4 even when part of the NSL complex. The acetylation pattern was very similar to the one that was observed with the eluted MSL complex (Mendjan et al., 2006). However, this result was not giving insight in the exact specificity of the MOF protein in the context of the NSL proteins.

In contrast, when free histone octamers were supplied as a substrate for modification by the enzymatically active NSL complexes, no preference for a single histone could be detected (data not shown) – all histones were acetylated to the same extent. This reflects the already mentioned specificity of histone acetyltransferases which increases with the choice of the most original substrates.

31.2 The NSL complex preferentially acetylates H4K16

The question was whether the observed acetylation was exclusively targeting lysine 16 of H4 or possibly other residues. I tried to address this with the following experiments.

31.2.1 Histone acetyltransferase assays with mutant polynucleosomes as substrate

The first approach to demonstrate the specificity of the purified complex *in vitro* was taken by using an altered form of the substrate. To this end the histone tails of H4 were genetically modified. Single lysine to alanine (K→A) mutations were introduced by site-directed mutagenesis. Namely, lysine 16 (K16A) and lysine 12 (K12A) were changed to alanine. It was hoped that in a reaction, depleted of the suspected target lysine, the enzyme's original specificity would possibly lead to a decrease or even complete elimination of the signal.

For this purpose, the mutated histones were expressed in bacteria and purified to homogeneity. The presence of the mutations (H4K16A and H4K12A) was confirmed by mass spectrometric analysis. Together with wild-type histones H2A, H2B and H3 the mutant H4 histones were assembled into octamers. Octamer fractions displaying the correct histone stoichiometry were used to reconstitute polynucleosomes. These were subsequently used as the substrate in the HAT assay.

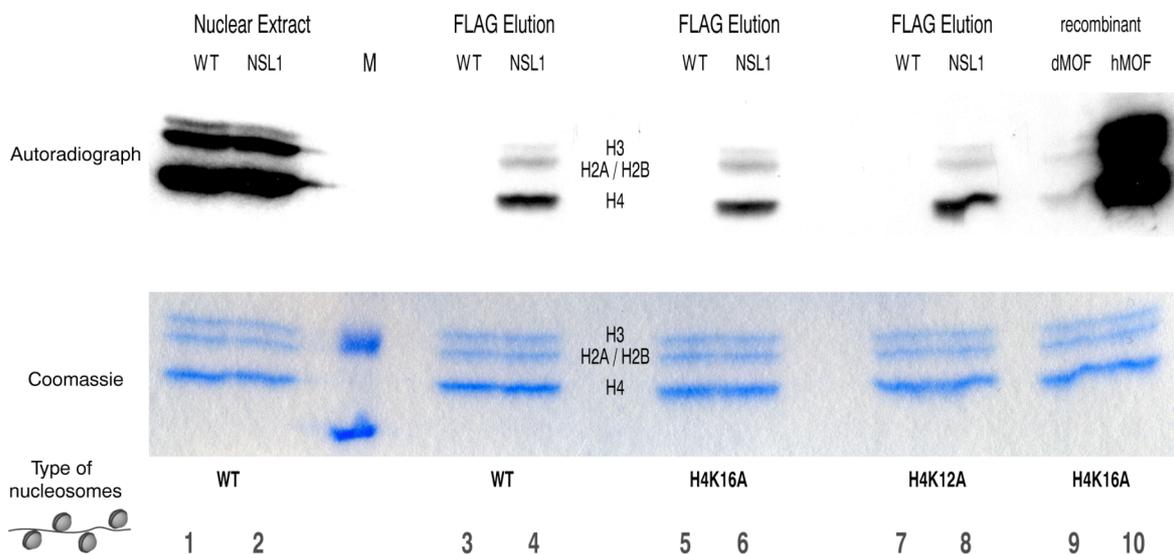


Figure 31-1: HAT-activity of TFH-NSL1 eluates on reconstituted polynucleosomes. Autoradiograph (top panel) and corresponding Coomassie gel (lower panel)

Lane 1 & 2: Nuclear extract from wild-type (WT) and stable TFH-NSL1 (NSL1) S2 cells on wild-type polynucleosomes

Elutions from wild-type (WT) and TFH-NSL1 (NSL1) FLAG-purifications assayed on wild-type polynucleosomes (**Lane 3 & 4**), on H4K16A mutant polynucleosomes (**Lane 5 & 6**) and on H4K12A mutant polynucleosomes (**Lane 7 & 8**).

Lane 9 & 10: recombinant MOF proteins, *Drosophila* MOF (dMOF) and human MOF (hMOF) tested on H4K16A mutant polynucleosomes.

It was surprising to see, that despite the absence of the suspected K16 target lysine the overall H4 acetylation levels were unaffected. When incubated with the NSL complex, H4K16A (Figure 31-1, lane 6) as well as H4K12A (lane 8) mutant polynucleosomes were both showing H4 acetylation to the same degree as wild-type polynucleosomes (lane 4). This result could be explained in several ways: either the complex was not strictly specific for a single lysine residue or, most likely, the specificity was lost upon removal of the original target. This would imply that the complex could turn towards the acetylation of other lysines in the histone H4 tail when the target lysine was not present anymore. Yet another explanation for this result would be the existence of some other histone acetyltransferase that copurified with the other proteins but was not identified by mass spectrometry.

It is noteworthy, that the bulk of the acetylation was still observed on H4 whereas the 'background' levels for H2A/B and H3 acetylation were not changed significantly.

One has to take into account that these conclusions are based on *in vitro* experiments that might have their limitations. It cannot be ruled out that the *in vivo* situation could still be different from the observed *in vitro* effects. Furthermore, it is known that histone acetyltransferases can exhibit limited substrate specificity whereas histone methyltransferases and kinases tend to be the most specific of the chromatin modifying enzymes (Kouzarides, 2007). For example, the recombinant human MOF protein potently acetylates all of the four core histones *in vitro* (Taipale et al., 2005b). The same was true when tested on K16-mutant polynucleosomes (Figure 31-1, lane 10). H4K16-specificity could only be detected when the enzyme was incorporated in the affinity-purified MSL-complex (Taipale et al., 2005b).

Additionally, *in vitro* acetylation experiments that have been done recently in the lab with a reconstituted trimer (MOF; MSL-1, MSL-3) from baculovirus-expressed proteins did show comparable results. When this minimal MSL-complex was tested on the same mutant polynucleosomes (H4K12A, H4K16A) a big drop in acetylation activity could be observed (data of Herbert Holz). However, the H4-specific enzymatic activity was not completely abolished, arguing for the loss of histone acetyltransferase specificity upon removal of the original target lysine. This, most probably, parallels the situation with the *in vitro* tested NSL complexes.

31.2.2 Mass spectrometric analysis reveals the specificity of the NSL complex

In order to get more conclusive data on the specificity of MOF in the context of NSL proteins, a second experimental approach was taken. The HAT-assay was performed under the same conditions as described but this time in the presence of 'cold' (non-radioactive) acetyl-CoA. Histones were separated after the reaction on a protein gel and stained with Coomassie blue. The histones were then excised from the gel, trypsin digested and analysed by targeted quantitative mass spectrometry. The mass spectrometry part was done by Marc Gentzel from the bioanalytical research group at EMBL Heidelberg.

The method is described in more detail in the section on the quantification of bulk histone acetylation (see chapter 32). Briefly, HPLC-MS/MS (high pressure liquid chromatography tandem mass spectrometry) was used to acquire sequence-specific information on the acetylation sites and to quantitate the signal intensities. Importantly, the high-sensitivity of this method allows for fast and site-specific identification of histone modifying activities. This is contrasted by conventional methods, like autoradiography, that are more time-consuming and that do not provide sequence-specific information.

It turned out that in HAT assays performed with eluted NSL complexes indeed a preference for lysine 16 of histone H4 could be observed (Figure 31-2). This was the case for one-step (FLAG-eluted) and two-step (HA-eluted) affinity-purified NSL complexes. The activity seemed to be higher in the fractions where the complexes were used already after elution from the first affinity resin - simply because of better recovery of NSL complexes.

The analysis indicated that, as already expected, other lysines in the H4 tail were also acetylated to some extent. This is depicted for lysine 5 (Figure 31-2 right panel). Very likely, these measurable 'side-products' were showing up as a consequence of the *in vitro* acetyltransferase reaction (compare 31.2.2.1).

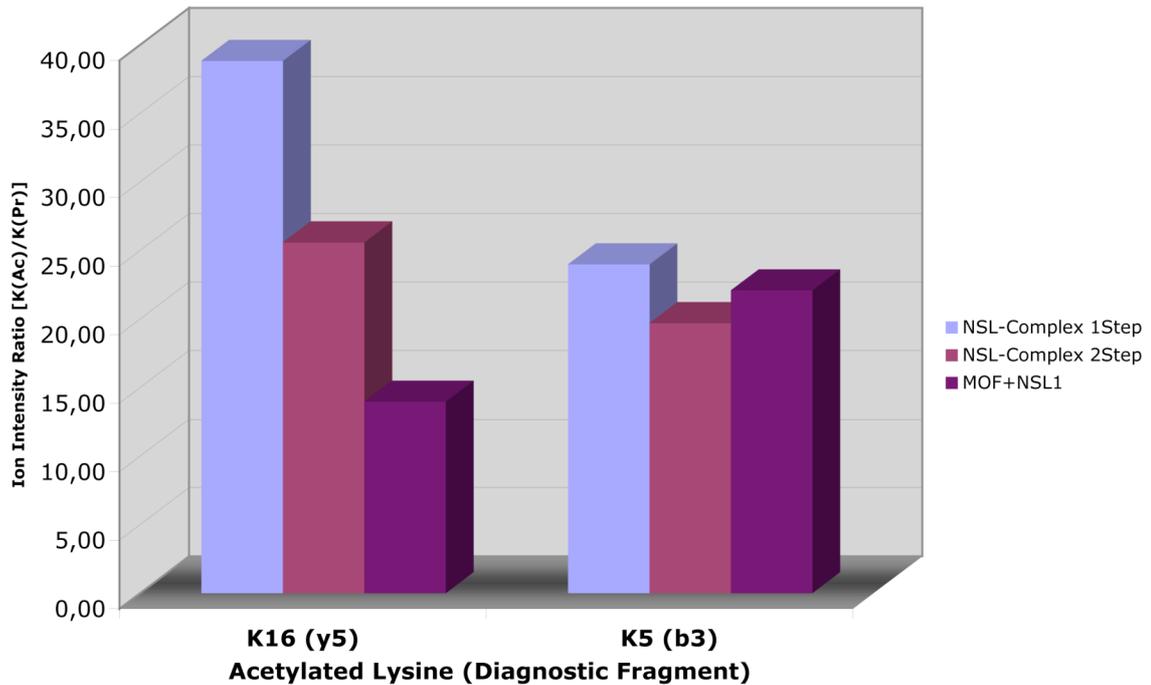


Figure 31-2: Analysis of acetylation specificity of eluted NSL complexes by targeted quantitative mass spectrometry

The diagram shows the H4K16- and H4K5-specific acetylation activity of eluted NSL complexes from 1-step and 2-step affinity purifications. For comparison the acetylation activity of a recombinant purified MOF/NSL1 dimer from the baculo/SF9 system is depicted. The ion intensity ratios are plotted for mono-acetylated peptides.

In parallel, the same experiment was performed with recombinant proteins – MOF and NSL1 – which were co-purified from SF9 cells. However, the reaction with the MOF/NSL1 dimer yielded lower H4K16 acetylation levels as seen with purified NSL1 complexes. It has to be noted that this part of the experiment was measured only once and the observed decrease in H4K16 acetylation could be an artefact of the analysis. Another explanation could be, that factors which normally contribute to acetyltransferase-specificity in the NSL complex, were missing in this dimer and the observed result was a consequence of this.

Nevertheless, the latter result was especially interesting, as we have not seen any enzymatic activity with the recombinant MOF protein (produced in the baculovirus/SF9 system) when assayed alone in radioactive HAT assays. Baculovirus reconstitution experiments showed that MOF was activated when present in combination with the MSL-1 plus MSL-3 proteins similarly as it was shown earlier (Morales et al., 2004).

We could further demonstrate that partial complexes of MOF and NSL1 sufficed as well to enhance the enzymatic activity of MOF. Bearing in mind MOF's interaction

via the PEHE domain of NSL1 and MSL-1 proteins might give hints to explain this phenomenon. Binding via the PEHE domain of these 'co-factors' could cause structural changes in the MOF protein that lead to activation and unleash the enzyme's acetylase function. However, equipping the enzyme with specificity for a target residue would require the association with other factors that confer this capability.

31.2.2.1 A change in enzyme kinetics for *in vitro* lysine acetylation?

An interesting observation was made with all of the measured *in vitro* acetylated samples. When looking at the distribution of unacetylated, single-, double-, triple- or four-fold acetylated fragments it became apparent that the unacetylated state represented the major fraction. The singly-modified fragments were clearly detectable and the levels for di- and tri-acetylated peptides dropped-down further. Surprisingly, the fraction of peptides harbouring all four acetylated residues was increased as well. This phenomenon indicated that the *in vitro* acetylation was an end-point reaction. Most probably, the MOF enzyme modified its original target residues in the first instance and then became promiscuous. Interestingly, the di- and triple-acetylated states were underrepresented which pointed to a change in enzyme kinetics after the first residue had been acetylated and consequently the peptide fragments were fully acetylated at all lysine positions.

32 Targeted quantitative analysis of lysine acetylation by MOF

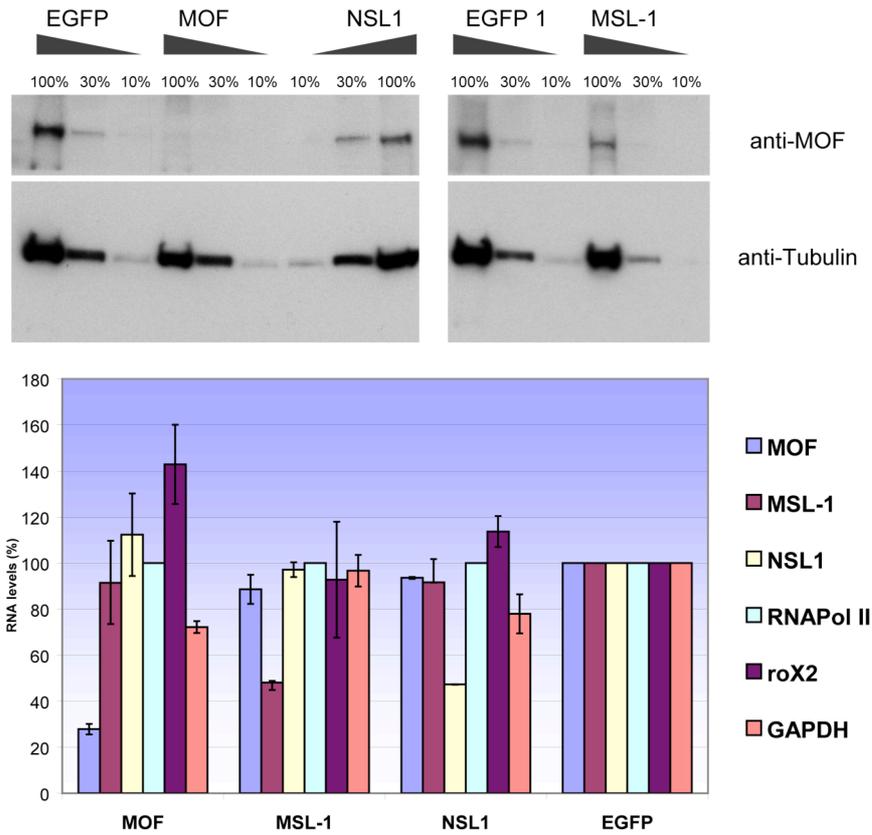
Post-translational modifications, such as acetylation, are modulating biological protein activity as described earlier. The information that is contained in these regulatory messages – especially on histones – can be generated by the number, the position and the combinatorial occurrence of modified residues. They are, in a manner of speaking, coding for the biological output. It has become of increasing interest to gather not only site-specific but at the same time also quantitative knowledge on post-translational modifications.

To address the effect of the MOF protein on the levels of global histone H4 tail lysine acetylation we examined the acetylation pattern by mass spectrometry. For that purpose, we made use of the enzyme's involvement in varying molecular contexts, namely the MSL- and the NSL-complexes.

The targeted quantitative approach used here, combined the methodical principle of nano-electrospray peptide sequencing to achieve high sensitivity with the separation strength of HPLC-MS/MS (high pressure liquid chromatography tandem mass spectrometry) experiments to quantitate bulk histone acetylation on different lysine residues.

In brief, the following experimental setup was developed during the course of this analysis. RNAi knockdowns of MOF, MSL-1 and NSL1 proteins were performed in two *Drosophila* cell lines: phenotypically male S2 and female Kc cells. Kc cells were shown to be female as they express SXL (sex lethal), the female-specific regulator of sex determination, but express very little MSL-2 protein (Mendjan et al., 2006). After successful knockdown, checked by quantitative RT-PCR and western blot (Figure 32-1), total histones were isolated from cells by acid extraction. The histones were separated on 1-D SDS gels, visualized by Coomassie staining, in-gel trypsin-digested and processed to obtain a measurable protein source.

A RNAi in Kc cells



B RNAi in S2 cells

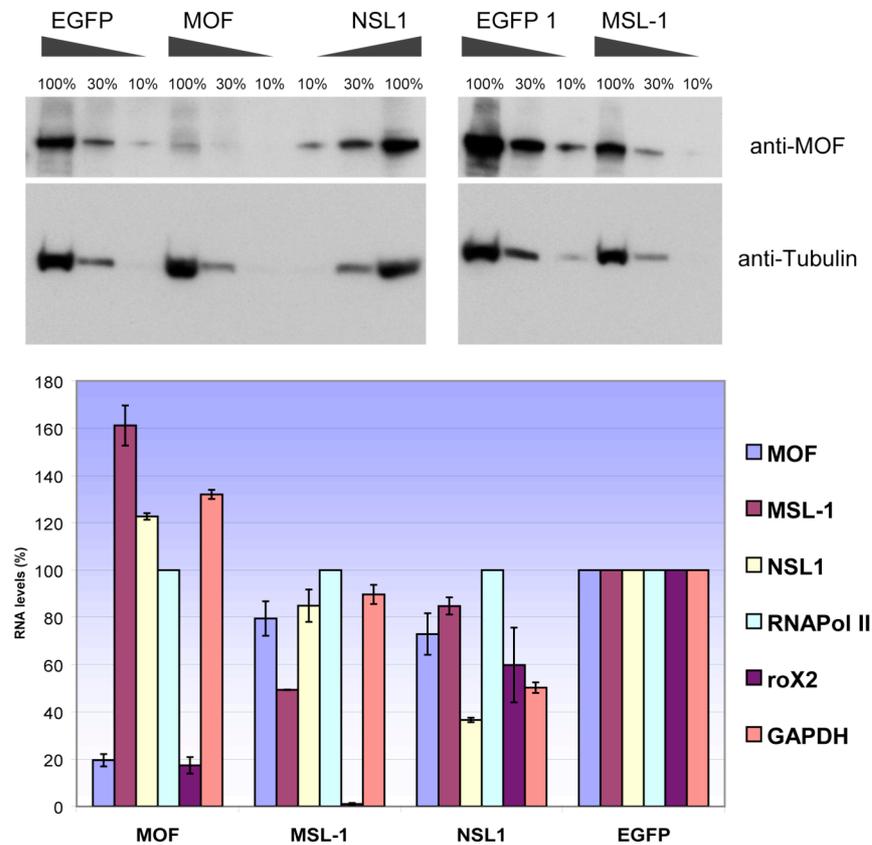


Figure 32-1: Analysis of knockdown efficiency by western blot and qRT-PCR

(Figure legend for Figure 32-1)

RNAi knockdown of MOF, NSL1 and MSL-1 in **(A)** Kc cells and in **(B)** S2 cells.

Cells were harvested on day 4 (MSL-1), day 6 (NSL1, EGFP) and day 8 (MOF) of knockdowns.

Western blots show titrations (100%, 30%, 10%) of whole cell extracts after dsRNA treatment.

MOF knockdown efficiency can be estimated by comparison with EGFP-RNAi control cells. Tubulin

serves as internal loading control. Quantitative RT-PCR was used to measure RNA levels (Y-axis)

after knockdown of individual targets (X-axis). Error bars reflect standard error of mean of 3

independent experiments. Values are normalized to RNA-Pol II RNA-levels.

An experimental trick was used to distinguish the naturally acetylated and the non-acetylated positions in the histone H4 4-17 peptide. Chemical conversion of naturally unmodified lysine residues to propionyl-lysine created a mass shift for fragment ion pairs which allowed the distinction. Importantly, this chemical derivatization was ideally suited as it preserved the residues of interest within a single peptide during tryptic digestion of the protein. The different peptide isoforms could be separated by Reverse Phase High Performance Liquid Chromatography (RP-HPLC) according to the number of naturally acetylated lysines.

The actual determination of relative site-specific lysine acetylation quantities was done by conventional quantitative HPLC-MS(/MS) followed by a targeted HPLC-MS/MS experiment. The peptides that were generated by the trypsin digest were harbouring K5, K8, K12 and K16 acetylation sites (Figure 32-2).

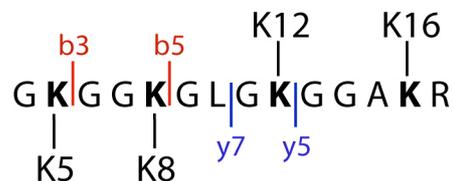


Figure 32-2: Nomenclature of the histone H4 peptide 4-17

The four acetylation sites and the diagnostic fragment ion pairs are indicated.

The diagnostic fragment ion pairs for K16 (y5) and K5 (b3) allowed direct quantification of acetylation sites. Values, expressed as ion intensity ratios (IIR), were calculated as follows for K16 and K5: ion intensity (in counts) of the signal for the acetylated peptide fragment divided by the ion intensity of the signal for the propionylated fragment, e.g. $y5[\text{K}(\text{Ac})/\text{K}(\text{Pr})]$ for H4K16.

The fragments with a combination of two acetylation sites, K5/K8 or K12/K16, required an additional relative comparison to reveal the site where a change of the acetylation occurred. Here, relative changes for K5 versus K8 (or K12 versus K16) were calculated as ion intensity of the acetylated fragment signal of K5 divided by

the sum of ion intensities of acetylated fragment signal for K5+K8 (same for K12/K16). Triple and quadruply acetylated fragments were not analysed because of their low abundance and the associated technical sensitivity restrictions.

For the reliable monitoring of quantitative changes of the acetylated lysines, each set of RNAi knockdowns was done in multiple biological ($n \geq 3$) and technical ($n = 4-9$) replicates (Figure 32-1). The error bars reflect the standard deviation (SD) of mean. EGFP dsRNA treated cells served as the control for naturally occurring acetylation levels.

32.1 Quantification of MOF-dependent H4K16 acetylation

Interestingly, a comparison between wild-type cells (MOCK) and EGFP-treated control cells revealed a slight decrease for global acetylation levels in the control cells (e.g. difference in K16 acetylation $\sim 0.5\%$ Kc, $\sim 3\%$ S2). We reasoned that the double-stranded RNA (dsRNA) itself, even if completely unrelated, exhibited a mild effect on overall acetylation levels. The finding that a big proportion of the H4 peptides from control cells were not acetylated at all – vaguely estimated to be in the range of 70-80% of total histones - demonstrates the great challenge for analytical sensitivity that was required to analyse the knockdown scenario.

Knowing that MOF is responsible for H4K16 acetylation led us to have a closer look at the acetylation status of this residue. The amount of acetylated peptides that carried a single modification on lysine 16 of histone H4 varied between S2 cells and Kc cells. Kc cells exhibited ~ 22 IIR units H4K16 acetylation (Figure 32-3 A, green) this being only around 2/3 of the H4K16 acetylation observed in S2 cells (~ 30 IRR units) (Figure 32-3 A, blue). This difference in acetylation could be due to cell line properties. However, assuming that this is not a cell line difference, the more likely explanation for this is the contribution of the male X chromosome in S2 cells, which is hyperacetylated by the MSL complex.

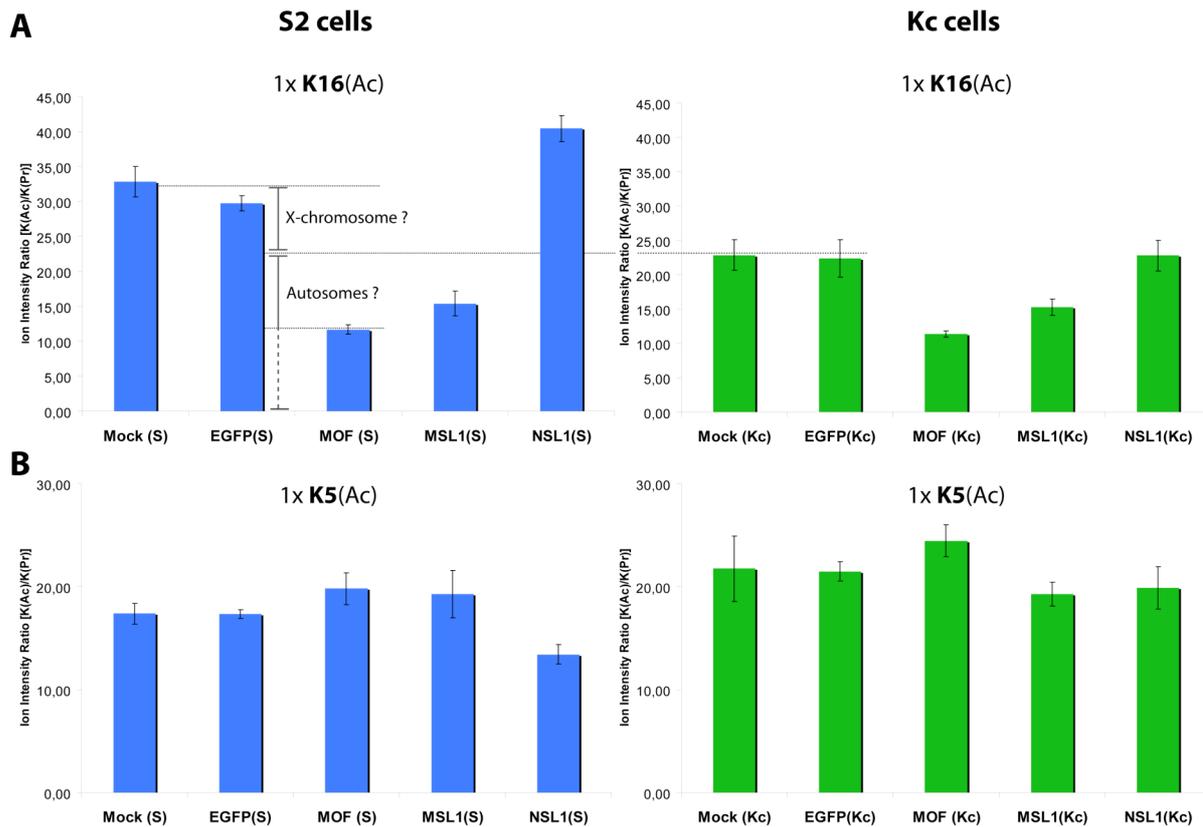


Figure 32-3: Relative quantification of histone acetylation in S2 (blue) and Kc (green) cells

Diagrams show ion intensity ratios (IRR, Y-axis) obtained for **(A)** K16 and **(B)** K5 mono-acetylated peptides after RNAi-mediated knockdown (X-axis) of MOF, MSL-1 and NSL1. MOCK- and EGFP-treated cells served as controls for original acetylation levels. The error bars reflect the standard deviation (SD) of mean of independent knockdown experiments ($n_{\text{biological}} \geq 3$, $n_{\text{technical}} = 4-9$).

The strongest reduction in H4K16 mono-acetylated peptides could be seen, as anticipated, in the MOF knockdowns. First of all, the MOF protein was successfully depleted from S2 and Kc cells with greater than 90% knockdown efficiency (see Figure 32-1). The ion intensity ratios for H4K16Ac after MOF RNAi in S2 (11.7 units) and Kc (11.4 units) cells were very similar which led us to the conclusion that H4K16 acetylation was reduced to basal levels. Therefore, it was very interesting to infer that H4K16 acetylation was not only detectable in female cells but it was mostly a result of MOF activity.

The remaining acetylation after MOF RNAi was either not contributed by the MOF enzyme - and consequently the result of other histone acetyltransferases - or the turnover for H4K16 acetylation on histone tails was rather slow. Coinciding with the latter explanation, it was earlier suggested that different pools of acetylated histones are present in the cell which exhibit distinct acetylation kinetics (Waterborg, 2002). Therefore, it is possible that the global acetylation (e.g. the

reminder that can be seen after MOF knockdown) corresponds to a slow turnover fraction whereas the fast turnover is involved in gene-specific regulation. However, apart from kinetic analyses of acetylated histones in various organisms and pulse-chase labelling, this issue has not been addressed directly.

Subsuming the acetylation patterns observed in the wild-type control and the MOF knockdown, Figure 32-3 A depicts the contribution of MOF on X chromosome and autosomes. The difference between wild-type K16 acetylation levels in S2 and Kc cells might constitute the hyperacetylation on the male X chromosome (marked by lines in the figure). Applying the same logic, acetylation on autosomes could be visualized as the difference between wild-type K16 levels in Kc cells and basal levels after MOF RNAi in S2 or Kc cells. According to the above mentioned considerations about acetylation turnover, it can not be excluded that the acetylated fraction on autosomes might be even bigger (Figure 32-3, dashed lines).

Comparably to MOF knockdown, the effect of MSL-1 RNAi led to a drop (S2 ~52% reduction, Kc ~32% reduction) in H4K16 acetylation levels (Figure 32-3). This effect could be explained – at least in S2 cells – with a disintegration of the MSL complex and a resulting decrease of MOF activity on the male X chromosome. The relevant quantitative RT-PCR supports the thought of MSL complex disassembly by showing a severe reduction of roX2 RNA concomitant with the MSL-1 knockdown (see Figure 32-1). Reasoning on this effect in Kc cells is a little bit more complicated as MSL-1 is thought to be destabilized in female cells by the absence of MSL-2 protein. However, MSL-1 RNA and protein is still detectable in female cells by RT-PCR (Figure 32-1) and western blot (not shown). Even though it does not seem to be targeted to the X chromosome it could still modulate MOF's activity in a genome-wide manner. MOF protein itself was slightly reduced in MSL-1 RNAi samples as seen by western.

The NSL1 knockdown could be quantitated by RT-PCR (Figure 32-1) but unfortunately not by western blot. However, the very apparent growth phenotype of NSL1-knockdown cells was an additional sign for efficient RNAi. Strikingly, the values for peptides monoacetylated at H4K16 were not diminished in the case of the NSL1 RNAi both in S2 (~40 IRR units) and Kc (~23 IRR units) cells (Figure 32-3). They were at wild-type levels - if anything was changed they were slightly increased in S2 cells.

The same method was applied to analyze the acetylation patterns on K5 (K8 and K12) in wild-type and knockdown situations. Due to the above mentioned coupling of two acetylation sites in the peptide, values for K8 and K12 could only be expressed as relative changes compared to K5 and K16 respectively.

However, a direct measurement of K5 acetylation was possible. The analysis of this data revealed an overall similar situation in S2 (blue) and Kc (green) cells (Figure 32-3 B). Additionally, the values did not change significantly when MOF, MSL-1 or NSL1 proteins were knocked down individually.

In summary, it can be noted that the developed mass spectrometric method could be successfully applied to the quantitative measurement of lysine acetylation on histone H4 tails. First of all, an overall reduction of H4K16 acetylation could be detected in Kc cells when compared to S2 cells. Secondly, when challenging the cells with dsRNA to knockdown MOF and MSL-1, reductions in H4K16 acetylation could be observed in both cell lines. This specific reduction did not relay to the acetylation status of H4K5 or H4K8 and was assigned to the action of the MSL complex. However, the knockdown of NSL1 did not result in a change of histone H4 lysine acetylation. Even though the MOF histone acetyltransferase was shown to be part of the NSL complex and to exhibit H4K16 specificity, the acetylation status of this lysine upon NSL1 RNAi was not affected.

33 Effect of NSL1 knockdown in S2 cells – the NSL1 RNAi phenotype

The RNA interference (RNAi) phenomenon is experimentally exploited to artificially regulate gene expression and study the function of genes in cell culture systems and even in whole model organisms. To achieve the targeted inhibition of gene expression, double-stranded ribonucleic acids (dsRNA) are synthesized with complementary sequence to the gene to be silenced. The uptake of the dsRNA into the cell activates the RNAi pathway and can lead to a drastic decrease in expression of the target gene (Elbashir et al., 2001). In contrast to a complete “knockout”, the RNAi-mediated knockdown may not totally abolish target gene expression. Therefore, the knockdown efficiency has to be quantitated by RT-PCR or western blotting.

Using this mechanism, I was studying the effects of NSL1 RNAi on *Drosophila* Schneider (S2) and Kc cells. The goal was to learn more about the physiological role of the NSL1 protein.

33.1 RNAi-mediated knockdown of NSL1 results in a retarded growth phenotype of *Drosophila* cell culture cells

Four separate dsRNA fragments were designed to target different regions of the NSL1 sequence (Figure 33-1). It was very reassuring to see that all four resulted in the same phenotype when transfected into S2 cells. For consistency reasons the fragment A was used in all the presented RNAi knockdowns of NSL1.

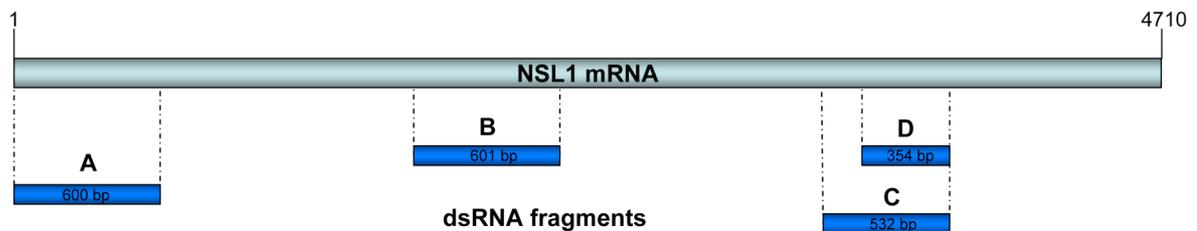


Figure 33-1: dsRNA fragments used for NSL1 RNAi

The four dsRNA fragments and their relative targeting position on the NSL1 mRNA are depicted. Presented RNAi experiments were performed with fragment A.

Cells that were treated with either of the dsRNA fragments were exhibiting strongly retarded growth rates (Figure 33-2). This was compared to control cells that were incubated with an unrelated dsRNA complementary to the coding region of the EGFP (enhanced green fluorescent protein) gene.

Cell numbers were reaching a plateau on day 3 of NSL1 knockdown. The drastic drop in cell numbers during the following days (days 4 to 6) was pointing to a growth arrest with subsequent cell death. The same effect was observed with dsRNA-treated Kc cells (see insert Figure 33-2).

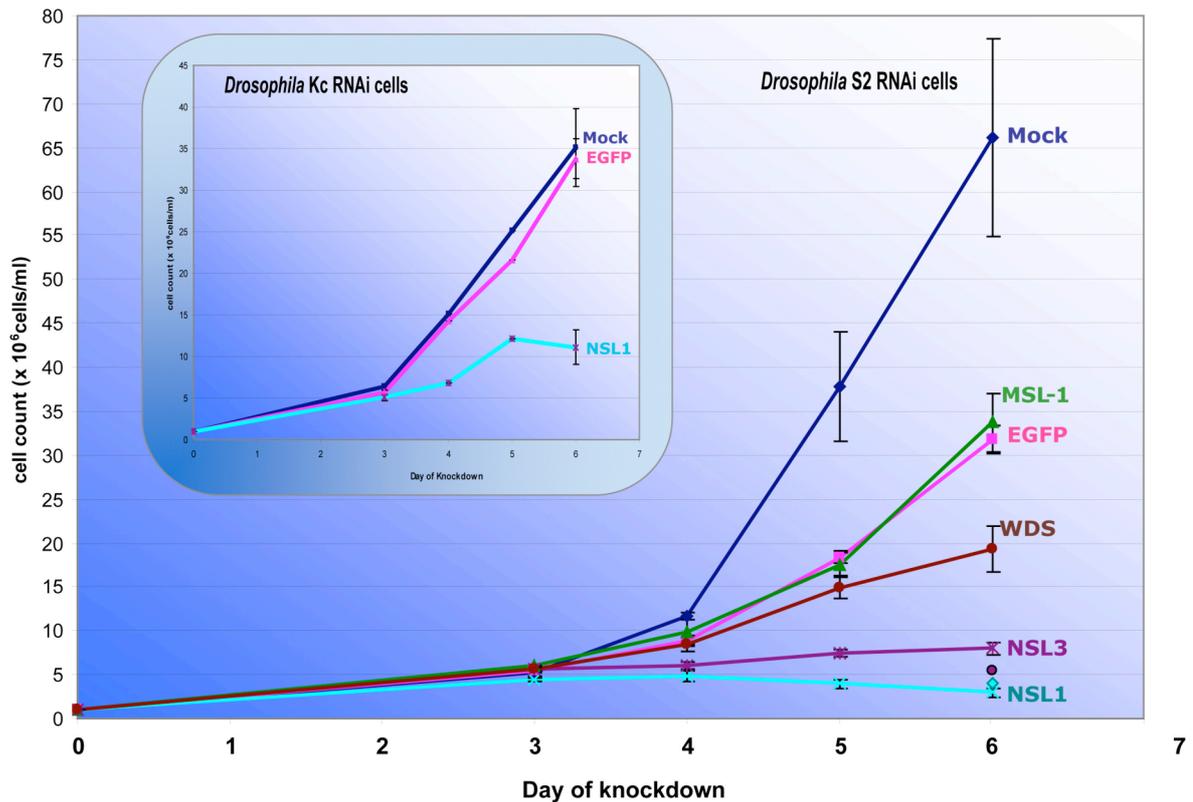


Figure 33-2: Growth curves for *Drosophila* S2 RNAi knockdown cells (insert: Kc RNAi cells)

dsRNA-treated cells were counted on days 0, 3, 4, 5 and 6 of knockdown. Error bars reflect standard error of mean of independent RNAi experiments ($n \geq 3$).

Interestingly, parallel knockdowns with dsRNAs targeting the NSL3 gene were leading to a comparable growth phenotype, suggesting that the two proteins might be naturally working together in the same process. On the other hand, reduction of MSL-1 transcripts did not result in decreased cell numbers which was again indicating that the MSL and NSL proteins might take over distinct functions.

In order to quantitate the NSL1 depletion by RNAi, RNA was extracted from the knockdown cells and used to perform quantitative Real-Time-PCR. Data from three independent biological replicates is presented (Figure 32-1). Values were normalized to RNA-polymerase II RNA levels and demonstrated that the NSL1 transcripts could be knocked down to about 35% of EGFP RNAi control levels.

It was difficult to demonstrate the knockdown by western blot as the NSL1 antibodies were not suitable for quantitation on whole cell lysate.

33.1.1 Knockdown of NSL1 impacts on NSL3 protein stability

A preliminary analysis of the effect of NSL1 on the stability of other NSL proteins yielded a surprising result. Notably, the depletion of the NSL1 protein had a pronounced impact on NSL3 protein stability. Several independent experiments showed a major reduction (up to complete depletion) of the NSL3 protein in NSL1 knockdown cells whereas in MOF knockdowns no NSL3 protein reduction could be observed (Figure 33-3). However, the levels of NSL3 RNA, as assessed by RT-PCR, were not influenced by either NSL1, MSL-1 or MOF knockdowns (not shown).

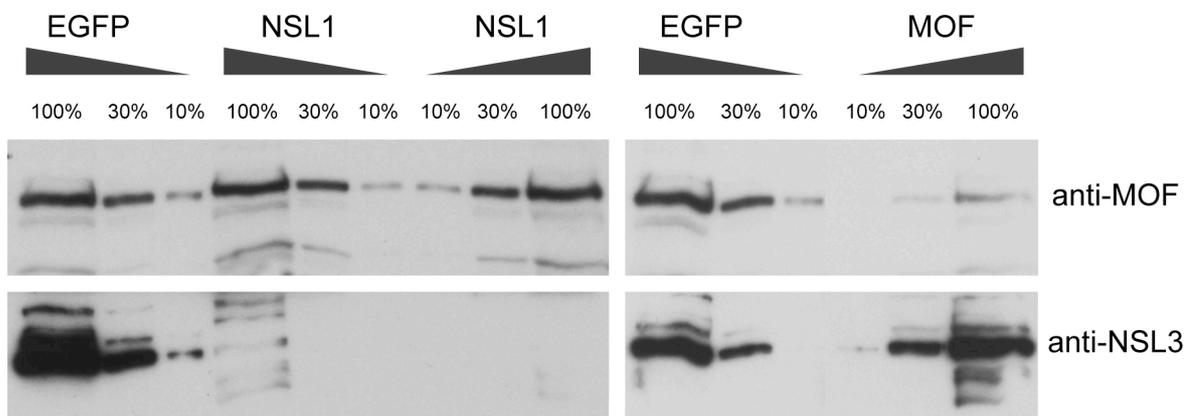


Figure 33-3: RNAi depletion of NSL1 impacts on NSL3 protein stability

Knockdowns for NSL1 (2x) and MOF were performed in Kc cells. RNAi cells were harvested after 6 days (NSL1) and 7 days (MOF) and whole cell extracts western blotted with anti-MOF and anti-NSL3 antibodies as indicated. The RNAi efficiency was judged by comparison with EGFP RNAi cells. NSL1 protein levels could not be checked by western as NSL1 antibodies were not suitable for quantitation on whole cell lysates.

If the observed phenomenon was a direct consequence of NSL1's absence or an indirect effect could not be judged by this experiment but it was reminding of the functional interdependence of MSL-1 and MSL-2 proteins (Chang et al., 1998). For confirmation it will be important to further evaluate the efficiency of NSL1 depletion and its effect on NSL3 stability.

33.2 NSL1 knockdown cells exhibit apoptotic features

A closer look at the cell morphology revealed a proportion of bigger cells in the NSL1 RNAi samples. They seemed to have vacuoles inside and most strikingly they were multinucleated, meaning that they exhibited two or more nuclei. Immunofluorescence co-staining with Hoechst dye (DNA staining) and antibodies against NUP153, a nuclear pore protein that can be used to stain the nuclear rim, indicated that in addition to membrane-surrounded nuclei, big chunks of DNA without nuclear membrane were distributed in the cell.

An antibody directed against the activated form of Caspase-9 was used to screen the NSL1 RNAi population for apoptotic events. Cells exhibiting the most extensive nuclear fragmentation stained positive for activated Caspase-9 and were therefore judged as apoptotic cells (see Figure 33-4).

This finding could explain the drastic decrease in cell numbers – affected cells being removed by apoptosis – but it was still unclear whether this was directly caused by the reduction of NSL1 protein levels or if it was a consequence of accumulated secondary effects originating from the loss of NSL1 protein.

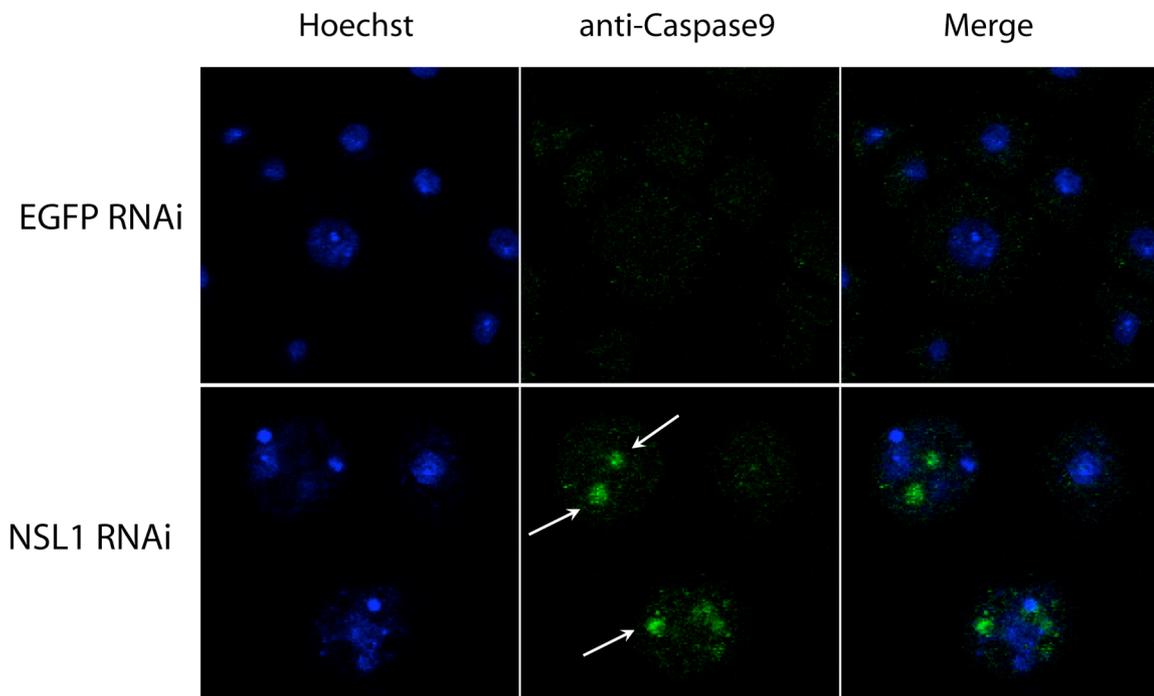


Figure 33-4: Staining of EGFP- and NSL1-RNAi S2 cells with antibody against activated Caspase 9

EGFP- and NSL1-RNAi cells were fixed on day 6 of knockdown and stained with Hoechst (blue) and antibody against activated Caspase 9 (green). EGFP control cells do not show apoptotic staining whereas NSL1 knockdown cells exhibit accumulation of activated Caspase-9 (white arrows point to green dots) which is typical for apoptotic events.

33.3 Live cell imaging

To further investigate this issue, I took advantage of a fluorescence microscope (Zeiss LSM510 Meta) that was equipped with an automated cell-tracking system. Placing the early-stage knockdown cells on this microscope and recording their growth behaviour over an extended time-period allowed a more detailed analysis.

A prerequisite for the visualization of the developing RNAi phenotype was the fluorescent staining of the genomic DNA. During the first trials, Hoechst 33342 DNA stain was used at a very low concentration (~ 100ng/ml) (Beaudouin et al., 2002) to follow the chromosomes.

Yet, Hoechst is a DNA-intercalating compound and can therefore interfere with mitotic condensation and normal chromosome segregation (Mora-Bermudez et al., 2007). Moreover, cells are very sensitive to light, especially in mitosis. The UV phototoxicity that emerges from the laser irradiation has to be confined to the minimum to rule out possible adversary effects. The experimental setup for live-cell imaging is always a trade-off between highest possible resolution and minimum illumination. This is difficult to achieve with conventional DNA stains.

It was therefore necessary to select another method for chromatin visualization. For that purpose I fused the EGFP sequence to the 5'-end of the *Drosophila* histone H2B cDNA. A S2 cell line stably expressing the EGFP-H2B fusion protein was established and the population selectively enriched for cells that were fluorescing at a medium intensity. The selection was done by fluorescence activated cell sorting (FACS). Finally, EGFP-H2B expressing cells were presenting approximately 60% of the population. This cell line was used for RNAi knockdown of NSL1 and parallel live-cell imaging.

33.3.1 RNAi-mediated knockdown of NSL1 causes severe chromosome segregation defects

The most obvious feature of the NSL1 knockdown could be readily observed: a strong reduction in cell numbers and an almost complete absence of cell divisions. This was not an artefact of the imaging setup as the control cells divided normally and reached confluency. Another feature of the control cells was their high

mobility. They were sometimes floating off the surface, leaving the observed field or settling down in the recorded area. This is very typical for actively dividing S2 cells but it also complicated the imaging process. Cells affected by NSL1 RNAi were less mobile.

Additionally, chromosome segregation defects were observed very frequently in the NSL1 knockdown cells, where the cell's chromosomes could not be separated and equally divided between the future daughter cells. Bridging chromatin, probably anaphase bridges, was inhibiting the normal mitotic course (Figure 33-5). Sometimes also much smaller nuclei were discovered (see Figure 33-5 B). These micronuclei possibly formed because of the loss of chromosomes during the attempted segregation process. These lost chromosomes were most probably represented by smaller pieces of free chromatin that could be observed in the cytoplasm. Strongly affected cells were showing large cytoplasmic vacuoles. Eventually, the polyploid/aneuploid cells went apoptotic and died.

Furthermore, the cells were unable to divide and form daughter cells, probably because of the unresolved chromosome segregation issue. Daughter nuclei – often still connected – were instead forming in the original cells and this was leading to the multinucleation phenotype. This is illustrated in the top panel of Figure 33-5 C (3.75 hours extract) where NSL1 knockdown cells were followed over a time-course of fourteen hours. Cells exhibited severe chromosome segregation defects and could not divide at all. In contrast, normal mitotic divisions could be observed in MOCK RNAi cells (Figure 33-5 C, lower panel).

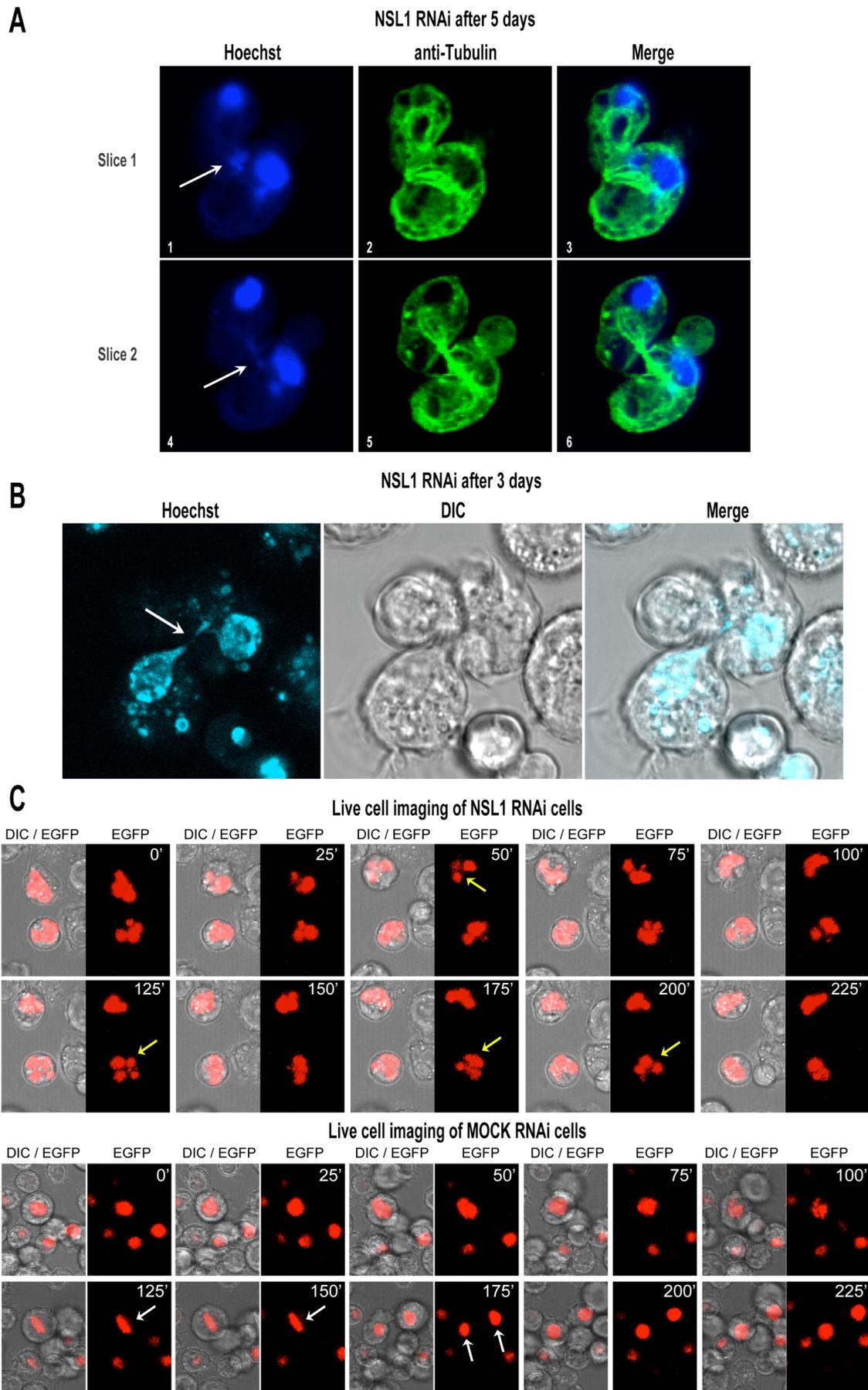


Figure 33-5: NSL1 RNAi causes segregation defects

(Figure legend for Figure 33-5)

(A) Immunofluorescence staining with anti-Tubulin (green) and DNA staining (blue) show unresolved connection between daughter cells. Two imaging slices are presented (pictures 1,2,3 and 4,5,6)

White arrow points to anaphase bridges in (A) and (B).

(B) Anaphase bridges as revealed by live cell imaging of Hoechst-stained (blue) NSL1 knockdown cells. DIC channel shows morphology of affected cells.

(C) Excerpt of live cell imaging recording of NSL1 and MOCK RNAi in EGFP-H2B S2 cells

Cells were followed for approximately 14 hours (day 5 to day 6 of knockdown), pictures were taken every 25 minutes. Timepoints are indicated in minutes from start of recording. DIC/EGFP and corresponding EGFP channels are shown next to each other. Yellow arrows are pointing to nuclei with chromosomes unable to separate. NSL1 RNAi cells were unable to complete mitosis even when observing longer time intervals (over-night imaging). White arrows are pointing to a normal mitosis in MOCK RNAi cells.

The live cell imaging experiment was repeated several times - always with the same outcome. Due to time constraints it was unfortunately not possible to increase the number of replica experiments in order to get final statistically significant values.

However, knowing these results, speculations on the underlying mechanisms for the observed phenotypic consequences of NLS1 knockdown can be made. For sure, the RNAi of NSL1 resulted in a mitotic defect as the lagging chromosome pieces, the anaphase bridges and the micronuclei were indicative of displaced mitotic chromosomes (Figure 33-5).

Various reasons could be responsible for the observed problems with proper chromosome segregation, ranging from difficulties in resolving sister chromatid cohesion to abnormal chromosome condensation or mitotic spindle defects.

Sister chromatid cohesion, for example, is mediated by a chromosomal protein complex, named the cohesin complex. It holds replicated sister DNA strands together after their synthesis (Uhlmann et al., 1999). The cohesion is essential for the alignment of chromosomes in metaphase but must be given up to allow the start of sister separation during anaphase (Uhlmann, 2004). Proteins implicated in this process are the DNA-binding SMC (structural maintenance of chromosomes) proteins (Strunnikov et al., 1999). SMC proteins represent a large family of ATPases that are conserved from bacteria to humans. They participate in many aspects of chromosome dynamics and depending on their interaction partners they can fulfil a variety of functions. They are as well part of the condensin complex, which is responsible for chromosome condensation. A specialized case in *C.elegans* is the mediation of gene dosage compensation by a SMC2/SMC4 type complex which represses transcription of sex chromosomes in hermaphrodites. Another function of this nematode complex is the proper

chromosome segregation in mitosis and possibly in meiosis (Chuang et al., 1996; Lieb et al., 1996).

Very interestingly, a predicted SMC protein (CG15415) was found in the NSL1 complex purification. Mass spectrometric analysis identified the protein with ten different peptides and a high Mascot score in bands cut out from silver-stained gels. Most likely, CG15415 is rather a complex-associating protein than a core complex member as it was only found in the 1-step purification but it could point towards an involvement of the complex in the presented processes. Unfortunately, it was not possible to further assay the presence of this protein in the HA-eluted NSL1 complex as no antibody for CG15415 was available.

Strikingly, RNAi directed to this gene caused as well a cell growth and viability phenotype when assayed in Kc167 and S2R+ cells in a genome-wide RNAi screen (Boutros et al., 2004).

Linking together the observed chromosome segregation defects and the presence of a putative SMC protein in the NSL complex is tempting for speculation about the involvement of this complex in mitosis and chromosome segregation. Yet, more investigations will need to be done to characterize a possible connection.

FINAL DISCUSSION AND PERSPECTIVES

The data presented in this work provides new insights in the involvement of MOF in different multiprotein complexes. The former purification of MSL-complexes via the MOF protein and the identification of the novel NSL proteins (Mendjan et al., 2006) demanded clarification on the protein composition of putative other MOF-containing complexes. Especially the integration of the histone acetyltransferase in the newly defined NSL complex could be characterised.

It could be demonstrated that bifurcation of the MOF-containing complexes is achieved by the mutually exclusive interaction with either MSL-1 or NSL1 (Figure 28-1). Extending the findings of Morales et al (Morales et al., 2004) that the PEHE domain of MSL-1 is the interaction module for MOF binding, we determined the requirement of this domain for the interaction of NSL1 with MOF (Figure 29-1).

Even though initial coimmunoprecipitation experiments were suggesting the existence of several MOF-containing complexes the ultimate answer about the detailed composition of such complexes could only come from affinity-purifications involving one of the novel interactors – in this case NSL1 - as bait proteins (Figure 30-3). The successful purification of the NSL complex opened new doors to explore its possible functions in conjunction with MOF. By this means, formerly identified proteins could be grouped whether they belonged to the MSL complex or the NSL complex. For example, the association with nucleoporins (Mendjan et al., 2006) could be clearly attributed to the MSL complex whereas no such link could be established with the NSL complex. However, the strong association of NSL1 with MOF, NSL2, NSL3, MBD-R2, MCRS2 and WDS became apparent in multiple ways and led to the classification of these proteins as core members of the NSL complex (Table 30-1).

An important step towards the determination of MOF's role in the NSL complex was taken by analysing the enzymatic activity. Different *in vitro* approaches could finally present that MOF was not only exhibiting acetylation activity when incorporated in the NSL complex (Figure 31-1) but also the specificity of the enzyme was still directed towards lysine 16 of the histone H4 tail (Figure 31-2).

In a second step, the involvement of MOF in global histone acetylation was determined. RNAi-mediated reduction of MOF, MSL-1 and NSL1 protein levels in cell culture cells was used to create different scenarios. Native histones were extracted from these samples and analysed for relative changes in bulk histone acetylation by a very sensitive mass spectrometric method. These data were able to pinpoint MOF's additional contribution in terms of H4K16 acetylation in male cells whereas the overall H4K16 acetylation levels were lower in female cells and the other acetylatable residues on the histone H4 tail did not show significant alterations in their acetylation pattern (see Figure 32-3).

The MSL-1 knockdown – most probably destabilizing the whole MSL complex – confirmed that the dosage compensation complex was responsible for this effect. In this regard the measured effects were congruent with the currently accepted hypothesis of dosage compensation in *Drosophila*. But more importantly, this could be proven for the first time by a quantitative analysis.

Remarkably, the removal of NSL1 did not result in a comparable outcome. H4K16 levels were not reduced when the NSL complex was targeted for disruption similar to the MSL complex (Figure 32-3). An explanation for the observed differences could be derived from recent high-resolution studies on MSL and NSL localization (unpublished data). ChIP-chip data from another PhD student, Jop Kind, indicate that in the default state MOF is restricted to promoter regions of its target genes on autosomes and the X chromosome. The difference in male cells is made by proteins like MSL-1 that allow the complex to spread along the coding regions. The H4K16 acetylation was seen to 'follow' the distribution of MOF and thereby generating a hyperacetylation pattern on the male X chromosome.

In contrast, preliminary data of another PhD student in the lab, Iryna Zhloba, shows that NSL1 seems to be restricted to the 5'-end of genes, to co-localize with MOF but it has not been observed to be located on the body of the genes. Therefore, it is very likely that acetylation contributed by this complex is not enriched on the coding regions of genes, neither on autosomes nor at the X chromosome. These recent findings could be reconciled with the observed differences in histone acetylation depending on which complex members were targeted by RNAi (Figure 32-3). The putatively smaller contribution of acetylation by MOF in the NSL complex could have different reasons. Either it was too little to quantitatively assess it by mass spectrometry or the NSL1 knockdown does not affect the acetylation of H4K16. One has to keep in mind that due to the limited possibilities of checking the NSL1 protein levels in knockdown cells, it cannot be ruled out that the NSL1 RNAi did not result in complete depletion of the protein and thus no effect was visible in the NSL1 knockdown scenario.

The NSL complex performs essential functions which are required for both sexes. On the one hand, this is reflected in the early larval lethality that disruptive P-element insertions in NSL genes cause in males and females (Mendjan et al., 2006). On the other hand, the essential nature of the NSL proteins was observed to manifest in the RNAi knockdown cells. Depletion of NSL1 caused severe mitotic segregation defects with the result of decreased cellular growth rates and cell death by induction of apoptosis (Figure 33-2, Figure 33-4, Figure 33-5). It is clear that future studies need to address whether these defects are a consequence of failure in mitotic condensation, chromosome segregation or involve other

processes that are not yet evident. Again, the knowledge about NSL-interacting proteins could give hints about the causality of the aforementioned processes. One such link is represented by the identification of the SMC-domain containing CG15415 protein in the NSL complex purification.

Another intriguing link to the presented purification was contributed by a genome-wide RNAi screen in *Drosophila* cells to identify new components of the hedgehog signalling pathway (Nybakken et al., 2005).

The Hedgehog Signaling Pathway is an essential cellular pathway required during the embryogenesis of various organisms. It is involved in cell fate determination, pattern formation, proliferation, and differentiation in multiple tissue types.

Very surprisingly, the study identified most of the NSL complex members as potential negative regulators of the hedgehog signaling pathway. Even though there were around 60 proteins classified in this category, NSL1, NSL2, NSL3, MOF, MCRS2, MBD-R2 and Chromator were amongst them. This important finding further reassured us that we found the true protein composition of the NSL complex and that these proteins most probably work in a concerted fashion in some yet unidentified process. In addition the eukaryotic translation initiation factor (eIF4e / CG8277) was both on the list of negative regulators and identified in one of our NSL purifications. The unforeseen link between NSL proteins and hedgehog signaling is currently being investigated in the lab.

The presented results unequivocally substantiated and further defined the novel link between NSL proteins and the histone acetyltransferase MOF.

By purification of the NSL complex we learned more about its distinct protein composition and were able to characterise its enzymatic activity. In addition, the closer analysis of the NSL1 protein suggested possible functions for the NSL complex in mitotic chromosome segregation.

It is clear that future studies will need to further address the involvement of the NSL complex in gene-specific regulation. Keeping in mind the extensive banding pattern that is produced by staining polytene chromosomes with NSL1-specific antibodies, it will be interesting to see if the localization of NSL1 and associated proteins to the promoter region of genes is reflecting the involvement of these proteins in gene regulation on a genome-wide scale. Furthermore, it will be

exciting to explore the presented links to mitotic chromosome segregation and possibly to cell signaling pathways.

The *in vivo* function of NSL1 remains elusive. Future work will therefore be required to elucidate the mechanism of action of these novel proteins in *Drosophila*. The foundations for this were laid by this work.

REFERENCES

- Ahmad, K., and Henikoff, S. (2001). Centromeres are specialized replication domains in heterochromatin. *J Cell Biol* 153, 101-110.
- Akhtar, A., and Becker, P. B. (2000a). Activation of transcription through histone H4 acetylation by MOF, an acetyltransferase essential for dosage compensation in *Drosophila*. *Mol Cell* 5, 367-375.
- Akhtar, A., Zink, D., and Becker, P. B. (2000b). Chromodomains are protein-RNA interaction modules. *Nature* 407, 405-409.
- Alekseyenko, A. A., Larschan, E., Lai, W. R., Park, P. J., and Kuroda, M. I. (2006). High-resolution ChIP-chip analysis reveals that the *Drosophila* MSL complex selectively identifies active genes on the male X chromosome. *Genes Dev* 20, 848-857.
- Allfrey, V. G. (1966). Structural modifications of histones and their possible role in the regulation of ribonucleic acid synthesis. *Proc Can Cancer Conf* 6, 313-335.
- Amor, D. J., and Choo, K. H. (2002). Neocentromeres: role in human disease, evolution, and centromere study. *Am J Hum Genet* 71, 695-714.
- Amrein, H., and Axel, R. (1997). Genes expressed in neurons of adult male *Drosophila*. *Cell* 88, 459-469.
- Arents, G., Burlingame, R. W., Wang, B. C., Love, W. E., and Moudrianakis, E. N. (1991). The nucleosomal core histone octamer at 3.1 Å resolution: a tripartite protein assembly and a left-handed superhelix. *Proc Natl Acad Sci U S A* 88, 10148-10152.
- Avner, P., and Heard, E. (2001). X-chromosome inactivation: counting, choice and initiation. *Nat Rev Genet* 2, 59-67.
- Avvakumov, N., and Cote, J. (2007). The MYST family of histone acetyltransferases and their intimate links to cancer. *Oncogene* 26, 5395-5407.
- Balasubramanian, R., Pray-Grant, M. G., Selleck, W., Grant, P. A., and Tan, S. (2002). Role of the Ada2 and Ada3 transcriptional coactivators in histone acetylation. *J Biol Chem* 277, 7989-7995.
- Bannister, A. J., and Kouzarides, T. (1996). The CBP co-activator is a histone acetyltransferase. *Nature* 384, 641-643.
- Bashaw, G. J., and Baker, B. S. (1995). The *msl-2* dosage compensation gene of *Drosophila* encodes a putative DNA-binding protein whose expression is sex specifically regulated by *Sex-lethal*. *Development* 121, 3245-3258.
- Bashaw, G. J., and Baker, B. S. (1997). The regulation of the *Drosophila msl-2* gene reveals a function for *Sex-lethal* in translational control. *Cell* 89, 789-798.
- Bauer, U. M., Dujat, S., Nielsen, S. J., Nightingale, K., and Kouzarides, T. (2002). Methylation at arginine 17 of histone H3 is linked to gene activation. *EMBO Rep* 3, 39-44.
- Beaudouin, J., Gerlich, D., Daigle, N., Eils, R., and Ellenberg, J. (2002). Nuclear envelope breakdown proceeds by microtubule-induced tearing of the lamina. *Cell* 108, 83-96.
- Belote, J. M., and Lucchesi, J. C. (1980). Male-specific lethal mutations of *Drosophila melanogaster*. *Genetics* 96, 165-186.
- Bernstein, B. E., Kamal, M., Lindblad-Toh, K., Bekiranov, S., Bailey, D. K., Huebert, D. J., McMahon, S., Karlsson, E. K., Kulbokas, E. J., 3rd, Gingeras, T. R., *et al.* (2005). Genomic maps and comparative analysis of histone modifications in human and mouse. *Cell* 120, 169-181.
- Bernstein, E., Duncan, E. M., Masui, O., Gil, J., Heard, E., and Allis, C. D. (2006). Mouse polycomb proteins bind differentially to methylated histone H3 and RNA and are enriched in facultative heterochromatin. *Mol Cell Biol* 26, 2560-2569.
- Bird, A. (2002). DNA methylation patterns and epigenetic memory. *Genes Dev* 16, 6-21.

- Bone, J. R., Lavender, J., Richman, R., Palmer, M. J., Turner, B. M., and Kuroda, M. I. (1994). Acetylated histone H4 on the male X chromosome is associated with dosage compensation in *Drosophila*. *Genes Dev* 8, 96-104.
- Borrow, J., Stanton, V. P., Jr., Andresen, J. M., Becher, R., Behm, F. G., Chaganti, R. S., Civin, C. I., Distech, C., Dube, I., Frischauf, A. M., *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, 33-41.
- Bouazoune, K., Mitterweger, A., Langst, G., Imhof, A., Akhtar, A., Becker, P. B., and Brehm, A. (2002). The dMi-2 chromodomains are DNA binding modules important for ATP-dependent nucleosome mobilization. *Embo J* 21, 2430-2440.
- Boutros, M., Kiger, A. A., Armknecht, S., Kerr, K., Hild, M., Koch, B., Haas, S. A., Paro, R., and Perrimon, N. (2004). Genome-wide RNAi analysis of growth and viability in *Drosophila* cells. *Science* 303, 832-835.
- Brahms, H., Meheus, L., de Brabandere, V., Fischer, U., and Luhrmann, R. (2001). Symmetrical dimethylation of arginine residues in spliceosomal Sm protein B/B' and the Sm-like protein LSm4, and their interaction with the SMN protein. *Rna* 7, 1531-1542.
- Braunstein, M., Sobel, R. E., Allis, C. D., Turner, B. M., and Broach, J. R. (1996). Efficient transcriptional silencing in *Saccharomyces cerevisiae* requires a heterochromatin histone acetylation pattern. *Mol Cell Biol* 16, 4349-4356.
- Brownell, J. E., Zhou, J., Ranalli, T., Kobayashi, R., Edmondson, D. G., Roth, S. Y., and Allis, C. D. (1996). Tetrahymena histone acetyltransferase A: a homolog to yeast Gcn5p linking histone acetylation to gene activation. *Cell* 84, 843-851.
- Buscaino, A., Kocher, T., Kind, J. H., Holz, H., Taipale, M., Wagner, K., Wilm, M., and Akhtar, A. (2003). MOF-regulated acetylation of MSL-3 in the *Drosophila* dosage compensation complex. *Mol Cell* 11, 1265-1277.
- Buscaino, A., Legube, G., and Akhtar, A. (2006). X-chromosome targeting and dosage compensation are mediated by distinct domains in MSL-3. *EMBO Rep* 7, 531-538.
- Cao, R., Wang, L., Wang, H., Xia, L., Erdjument-Bromage, H., Tempst, P., Jones, R. S., and Zhang, Y. (2002). Role of histone H3 lysine 27 methylation in Polycomb-group silencing. *Science* 298, 1039-1043.
- Carrozza, M. J., Li, B., Florens, L., Suganuma, T., Swanson, S. K., Lee, K. K., Shia, W. J., Anderson, S., Yates, J., Washburn, M. P., and Workman, J. L. (2005). Histone H3 methylation by Set2 directs deacetylation of coding regions by Rpd3S to suppress spurious intragenic transcription. *Cell* 123, 581-592.
- Chang, K. A., and Kuroda, M. I. (1998). Modulation of MSL1 abundance in female *Drosophila* contributes to the sex specificity of dosage compensation. *Genetics* 150, 699-709.
- Chen, D., Ma, H., Hong, H., Koh, S. S., Huang, S. M., Schurter, B. T., Aswad, D. W., and Stallcup, M. R. (1999). Regulation of transcription by a protein methyltransferase. *Science* 284, 2174-2177.
- Chow, J. C., Yen, Z., Ziesche, S. M., and Brown, C. J. (2005). Silencing of the mammalian X chromosome. *Annu Rev Genomics Hum Genet* 6, 69-92.
- Chuang, P. T., Lieb, J. D., and Meyer, B. J. (1996). Sex-specific assembly of a dosage compensation complex on the nematode X chromosome. *Science* 274, 1736-1739.
- Clamp, M., Cuff, J., Searle, S. M., and Barton, G. J. (2004). The Jalview Java alignment editor. *Bioinformatics* 20, 426-427.
- Clemens, J. C., Worby, C. A., Simonson-Leff, N., Muda, M., Maehama, T., Hemmings, B. A., and Dixon, J. E. (2000). Use of double-stranded RNA interference in *Drosophila* cell lines to dissect signal transduction pathways. *Proc Natl Acad Sci U S A* 97, 6499-6503.
- Clements, A., Rojas, J. R., Trievel, R. C., Wang, L., Berger, S. L., and Marmorstein, R. (1999). Crystal structure of the histone acetyltransferase domain of the human PCAF transcriptional regulator bound to coenzyme A. *Embo J* 18, 3521-3532.

- Clouaire, T., Roussigne, M., Ecochard, V., Mathe, C., Amalric, F., and Girard, J. P. (2005). The THAP domain of THAP1 is a large C2CH module with zinc-dependent sequence-specific DNA-binding activity. *Proc Natl Acad Sci U S A* 102, 6907-6912.
- Corona, D. F., and Tamkun, J. W. (2004). Multiple roles for ISWI in transcription, chromosome organization and DNA replication. *Biochim Biophys Acta* 1677, 113-119.
- Cosgrove, M. S., Boeke, J. D., and Wolberger, C. (2004). Regulated nucleosome mobility and the histone code. *Nat Struct Mol Biol* 11, 1037-1043.
- Costanzi, C., and Pehrson, J. R. (1998). Histone macroH2A1 is concentrated in the inactive X chromosome of female mammals. *Nature* 393, 599-601.
- Craig, J. M. (2005). Heterochromatin--many flavours, common themes. *Bioessays* 27, 17-28.
- Cuthbert, G. L., Daujat, S., Snowden, A. W., Erdjument-Bromage, H., Hagiwara, T., Yamada, M., Schneider, R., Gregory, P. D., Tempst, P., Bannister, A. J., and Kouzarides, T. (2004). Histone deimination antagonizes arginine methylation. *Cell* 118, 545-553.
- Dahlsveen, I. K., Gilfillan, G. D., Shelest, V. I., Lamm, R., and Becker, P. B. (2006). Targeting determinants of dosage compensation in *Drosophila*. *PLoS Genet* 2, e5.
- Davey, C. A., Sargent, D. F., Luger, K., Maeder, A. W., and Richmond, T. J. (2002). Solvent mediated interactions in the structure of the nucleosome core particle at 1.9 Å resolution. *J Mol Biol* 319, 1097-1113.
- Deckert, J., and Struhl, K. (2002). Targeted recruitment of Rpd3 histone deacetylase represses transcription by inhibiting recruitment of Swi/Snf, SAGA, and TATA binding protein. *Mol Cell Biol* 22, 6458-6470.
- Dhalluin, C., Carlson, J. E., Zeng, L., He, C., Aggarwal, A. K., and Zhou, M. M. (1999). Structure and ligand of a histone acetyltransferase bromodomain. *Nature* 399, 491-496.
- Dignam, J. D., Lebovitz, R. M., and Roeder, R. G. (1983). Accurate transcription initiation by RNA polymerase II in a soluble extract from isolated mammalian nuclei. *Nucleic Acids Res* 11, 1475-1489.
- Dimitri, P., Corradini, N., Rossi, F., and Verni, F. (2005). The paradox of functional heterochromatin. *Bioessays* 27, 29-41.
- Dorigo, B., Schalch, T., Kulangara, A., Duda, S., Schroeder, R. R., and Richmond, T. J. (2004). Nucleosome arrays reveal the two-start organization of the chromatin fiber. *Science* 306, 1571-1573.
- Dou, Y., and Gorovsky, M. A. (2000). Phosphorylation of linker histone H1 regulates gene expression in vivo by creating a charge patch. *Mol Cell* 6, 225-231.
- Dou, Y., Milne, T. A., Tackett, A. J., Smith, E. R., Fukuda, A., Wysocka, J., Allis, C. D., Chait, B. T., Hess, J. L., and Roeder, R. G. (2005). Physical association and coordinate function of the H3 K4 methyltransferase MLL1 and the H4 K16 acetyltransferase MOF. *Cell* 121, 873-885.
- Durocher, D., Henckel, J., Fersht, A. R., and Jackson, S. P. (1999). The FHA domain is a modular phosphopeptide recognition motif. *Mol Cell* 4, 387-394.
- Durocher, D., and Jackson, S. P. (2002). The FHA domain. *FEBS Lett* 513, 58-66.
- Dutnall, R. N., Tafrov, S. T., Sternglanz, R., and Ramakrishnan, V. (1998). Structure of the histone acetyltransferase Hat1: a paradigm for the GCN5-related N-acetyltransferase superfamily. *Cell* 94, 427-438.
- Earnshaw, W. C., and Migeon, B. R. (1985). Three related centromere proteins are absent from the inactive centromere of a stable isodicentric chromosome. *Chromosoma* 92, 290-296.
- Edgar, R. C. (2004). MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res* 32, 1792-1797.
- Eggert, H., Gortchakov, A., and Saumweber, H. (2004). Identification of the *Drosophila* interband-specific protein Z4 as a DNA-binding zinc-finger protein determining chromosomal structure. *J Cell Sci* 117, 4253-4264.

- Ehrenhofer-Murray, A. E., Rivier, D. H., and Rine, J. (1997). The role of Sas2, an acetyltransferase homologue of *Saccharomyces cerevisiae*, in silencing and ORC function. *Genetics* *145*, 923-934.
- Elbashir, S. M., Harborth, J., Lendeckel, W., Yalcin, A., Weber, K., and Tuschl, T. (2001). Duplexes of 21-nucleotide RNAs mediate RNA interference in cultured mammalian cells. *Nature* *411*, 494-498.
- Fagegaltier, D., and Baker, B. S. (2004). X chromosome sites autonomously recruit the dosage compensation complex in *Drosophila* males. *PLoS Biol* *2*, e341.
- Fischle, W., Tseng, B. S., Dormann, H. L., Ueberheide, B. M., Garcia, B. A., Shabanowitz, J., Hunt, D. F., Funabiki, H., and Allis, C. D. (2005). Regulation of HP1-chromatin binding by histone H3 methylation and phosphorylation. *Nature* *438*, 1116-1122.
- Flanagan, J. F., Mi, L. Z., Chruszcz, M., Cymborowski, M., Clines, K. L., Kim, Y., Minor, W., Rastinejad, F., and Khorasanizadeh, S. (2005). Double chromodomains cooperate to recognize the methylated histone H3 tail. *Nature* *438*, 1181-1185.
- Flaus, A., and Owen-Hughes, T. (2003). Dynamic properties of nucleosomes during thermal and ATP-driven mobilization. *Mol Cell Biol* *23*, 7767-7779.
- Forler, D., Kocher, T., Rode, M., Gentzel, M., Izaurralde, E., and Wilm, M. (2003). An efficient protein complex purification method for functional proteomics in higher eukaryotes. *Nat Biotechnol* *21*, 89-92.
- Franke, A., and Baker, B. S. (1999). The rox1 and rox2 RNAs are essential components of the compensasome, which mediates dosage compensation in *Drosophila*. *Mol Cell* *4*, 117-122.
- Gilbert, N., Boyle, S., Fiegler, H., Woodfine, K., Carter, N. P., and Bickmore, W. A. (2004). Chromatin architecture of the human genome: gene-rich domains are enriched in open chromatin fibers. *Cell* *118*, 555-566.
- Gilfillan, G. D., Straub, T., de Wit, E., Greil, F., Lamm, R., van Steensel, B., and Becker, P. B. (2006). Chromosome-wide gene-specific targeting of the *Drosophila* dosage compensation complex. *Genes Dev* *20*, 858-870.
- Glozak, M. A., Sengupta, N., Zhang, X., and Seto, E. (2005). Acetylation and deacetylation of non-histone proteins. *Gene* *363*, 15-23.
- Gorman, M., Franke, A., and Baker, B. S. (1995). Molecular characterization of the male-specific lethal-3 gene and investigations of the regulation of dosage compensation in *Drosophila*. *Development* *121*, 463-475.
- Gortchakov, A. A., Eggert, H., Gan, M., Mattow, J., Zhimulev, I. F., and Saumweber, H. (2005). Chriz, a chromodomain protein specific for the interbands of *Drosophila melanogaster* polytene chromosomes. *Chromosoma* *114*, 54-66.
- Gottesfeld, J. M., Belitsky, J. M., Melander, C., Dervan, P. B., and Luger, K. (2002). Blocking transcription through a nucleosome with synthetic DNA ligands. *J Mol Biol* *321*, 249-263.
- Grant, P. A., Duggan, L., Cote, J., Roberts, S. M., Brownell, J. E., Candau, R., Ohba, R., Owen-Hughes, T., Allis, C. D., Winston, F., *et al.* (1997). Yeast Gcn5 functions in two multisubunit complexes to acetylate nucleosomal histones: characterization of an Ada complex and the SAGA (Spt/Ada) complex. *Genes Dev* *11*, 1640-1650.
- Grozinger, C. M., and Schreiber, S. L. (2002). Deacetylase enzymes: biological functions and the use of small-molecule inhibitors. *Chem Biol* *9*, 3-16.
- Grunstein, M. (1997). Histone acetylation in chromatin structure and transcription. *Nature* *389*, 349-352.
- Gu, W., and Roeder, R. G. (1997). Activation of p53 sequence-specific DNA binding by acetylation of the p53 C-terminal domain. *Cell* *90*, 595-606.
- Gu, W., Szauter, P., and Lucchesi, J. C. (1998). Targeting of MOF, a putative histone acetyl transferase, to the X chromosome of *Drosophila melanogaster*. *Dev Genet* *22*, 56-64.
- Gu, W., Wei, X., Pannuti, A., and Lucchesi, J. C. (2000). Targeting the chromatin-remodeling MSL complex of *Drosophila* to its sites of action on the X chromosome requires both acetyl transferase and ATPase activities. *Embo J* *19*, 5202-5211.

- Hamiche, A., Sandaltzopoulos, R., Gdula, D. A., and Wu, C. (1999). ATP-dependent histone octamer sliding mediated by the chromatin remodeling complex NURF. *Cell* **97**, 833-842.
- Hassan, A. H., Prochasson, P., Neely, K. E., Galasinski, S. C., Chandy, M., Carrozza, M. J., and Workman, J. L. (2002). Function and selectivity of bromodomains in anchoring chromatin-modifying complexes to promoter nucleosomes. *Cell* **111**, 369-379.
- Haushalter, K. A., and Kadonaga, J. T. (2003). Chromatin assembly by DNA-translocating motors. *Nat Rev Mol Cell Biol* **4**, 613-620.
- Heitz, E. (1928). Das Heterochromatin der Moose. *Jahrb Wiss Bot* **1**, 762-818.
- Hilfiker, A., Hilfiker-Kleiner, D., Pannuti, A., and Lucchesi, J. C. (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**, 2054-2060.
- Hirota, T., Lipp, J. J., Toh, B. H., and Peters, J. M. (2005). Histone H3 serine 10 phosphorylation by Aurora B causes HP1 dissociation from heterochromatin. *Nature* **438**, 1176-1180.
- Holbert, M. A., and Marmorstein, R. (2005). Structure and activity of enzymes that remove histone modifications. *Curr Opin Struct Biol* **15**, 673-680.
- Hollmann, M., Simmerl, E., Schafer, U., and Schafer, M. A. (2002). The essential *Drosophila melanogaster* gene *wds* (will die slowly) codes for a WD-repeat protein with seven repeats. *Mol Genet Genomics* **268**, 425-433.
- Horn, P. J., and Peterson, C. L. (2002). Molecular biology. Chromatin higher order folding-wrapping up transcription. *Science* **297**, 1824-1827.
- Hsu, J. Y., Sun, Z. W., Li, X., Reuben, M., Tatchell, K., Bishop, D. K., Grushcow, J. M., Brame, C. J., Caldwell, J. A., Hunt, D. F., *et al.* (2000). Mitotic phosphorylation of histone H3 is governed by *Ipl1/aurora* kinase and *Glc7/PP1* phosphatase in budding yeast and nematodes. *Cell* **102**, 279-291.
- Huang, Y., Fang, J., Bedford, M. T., Zhang, Y., and Xu, R. M. (2006). Recognition of histone H3 lysine-4 methylation by the double tudor domain of JMJD2A. *Science* **312**, 748-751.
- Ito, T., Bulger, M., Pazin, M. J., Kobayashi, R., and Kadonaga, J. T. (1997). ACF, an ISWI-containing and ATP-utilizing chromatin assembly and remodeling factor. *Cell* **90**, 145-155.
- Jenuwein, T., and Allis, C. D. (2001). Translating the histone code. *Science* **293**, 1074-1080.
- Kadosh, D., and Struhl, K. (1997). Repression by Ume6 involves recruitment of a complex containing Sin3 corepressor and Rpd3 histone deacetylase to target promoters. *Cell* **89**, 365-371.
- Kageyama, Y., Mengus, G., Gilfillan, G., Kennedy, H. G., Stuckenholz, C., Kelley, R. L., Becker, P. B., and Kuroda, M. I. (2001). Association and spreading of the *Drosophila* dosage compensation complex from a discrete roX1 chromatin entry site. *Embo J* **20**, 2236-2245.
- Kamakaka, R. T., and Biggins, S. (2005). Histone variants: deviants? *Genes Dev* **19**, 295-310.
- Katan-Khaykovich, Y., and Struhl, K. (2002). Dynamics of global histone acetylation and deacetylation in vivo: rapid restoration of normal histone acetylation status upon removal of activators and repressors. *Genes Dev* **16**, 743-752.
- Kelley, R. L., Meller, V. H., Gordadze, P. R., Roman, G., Davis, R. L., and Kuroda, M. I. (1999). Epigenetic spreading of the *Drosophila* dosage compensation complex from roX RNA genes into flanking chromatin. *Cell* **98**, 513-522.
- Kelley, R. L., Wang, J., Bell, L., and Kuroda, M. I. (1997). Sex lethal controls dosage compensation in *Drosophila* by a non-splicing mechanism. *Nature* **387**, 195-199.
- Kimura, A., Matsubara, K., and Horikoshi, M. (2005). A decade of histone acetylation: marking eukaryotic chromosomes with specific codes. *J Biochem (Tokyo)* **138**, 647-662.

- Kind, J., and Akhtar, A. (2007). Cotranscriptional recruitment of the dosage compensation complex to X-linked target genes. *Genes Dev* 21, 2030-2040.
- Kleff, S., Andrusis, E. D., Anderson, C. W., and Sternglanz, R. (1995). Identification of a gene encoding a yeast histone H4 acetyltransferase. *J Biol Chem* 270, 24674-24677.
- Kornberg, R. D. (1974). Chromatin structure: a repeating unit of histones and DNA. *Science* 184, 868-871.
- Kouzarides, T. (2007). Chromatin modifications and their function. *Cell* 128, 693-705.
- Kuo, M. H., Brownell, J. E., Sobel, R. E., Ranalli, T. A., Cook, R. G., Edmondson, D. G., Roth, S. Y., and Allis, C. D. (1996). Transcription-linked acetylation by Gcn5p of histones H3 and H4 at specific lysines. *Nature* 383, 269-272.
- Kuo, M. H., vom Baur, E., Struhl, K., and Allis, C. D. (2000). Gcn4 activator targets Gcn5 histone acetyltransferase to specific promoters independently of transcription. *Mol Cell* 6, 1309-1320.
- Kurdistani, S. K., and Grunstein, M. (2003). Histone acetylation and deacetylation in yeast. *Nat Rev Mol Cell Biol* 4, 276-284.
- Kuroda, M. I., Kernan, M. J., Kreber, R., Ganetzky, B., and Baker, B. S. (1991). The maleless protein associates with the X chromosome to regulate dosage compensation in *Drosophila*. *Cell* 66, 935-947.
- Kusch, T., Florens, L., Macdonald, W. H., Swanson, S. K., Glaser, R. L., Yates, J. R., 3rd, Abmayr, S. M., Washburn, M. P., and Workman, J. L. (2004). Acetylation by Tip60 is required for selective histone variant exchange at DNA lesions. *Science* 306, 2084-2087.
- Lachner, M., O'Carroll, D., Rea, S., Mechtler, K., and Jenuwein, T. (2001). Methylation of histone H3 lysine 9 creates a binding site for HP1 proteins. *Nature* 410, 116-120.
- Ladurner, A. G. (2003). Inactivating chromosomes: a macro domain that minimizes transcription. *Mol Cell* 12, 1-3.
- Langst, G., and Becker, P. B. (2001). Nucleosome mobilization and positioning by ISWI-containing chromatin-remodeling factors. *J Cell Sci* 114, 2561-2568.
- Lau, O. D., Courtney, A. D., Vassilev, A., Marzilli, L. A., Cotter, R. J., Nakatani, Y., and Cole, P. A. (2000). p300/CBP-associated factor histone acetyltransferase processing of a peptide substrate. Kinetic analysis of the catalytic mechanism. *J Biol Chem* 275, 21953-21959.
- Legube, G., McWeeney, S. K., Lercher, M. J., and Akhtar, A. (2006). X-chromosome-wide profiling of MSL-1 distribution and dosage compensation in *Drosophila*. *Genes Dev* 20, 871-883.
- Lenart, P., Bacher, C. P., Daigle, N., Hand, A. R., Eils, R., Terasaki, M., and Ellenberg, J. (2005). A contractile nuclear actin network drives chromosome congression in oocytes. *Nature* 436, 812-818.
- Li, J., Smith, G. P., and Walker, J. C. (1999). Kinase interaction domain of kinase-associated protein phosphatase, a phosphoprotein-binding domain. *Proc Natl Acad Sci U S A* 96, 7821-7826.
- Lieb, J. D., Capowski, E. E., Meneely, P., and Meyer, B. J. (1996). DPY-26, a link between dosage compensation and meiotic chromosome segregation in the nematode. *Science* 274, 1732-1736.
- Lin, Y., Fletcher, C. M., Zhou, J., Allis, C. D., and Wagner, G. (1999). Solution structure of the catalytic domain of GCN5 histone acetyltransferase bound to coenzyme A. *Nature* 400, 86-89.
- Litt, M. D., Simpson, M., Gaszner, M., Allis, C. D., and Felsenfeld, G. (2001). Correlation between histone lysine methylation and developmental changes at the chicken beta-globin locus. *Science* 293, 2453-2455.
- Luger, K., Rechsteiner, T. J., Flaus, A. J., Wayne, M. M., and Richmond, T. J. (1997). Characterization of nucleosome core particles containing histone proteins made in bacteria. *J Mol Biol* 272, 301-311.
- Lyko, F., Ramsahoye, B. H., and Jaenisch, R. (2000). DNA methylation in *Drosophila melanogaster*. *Nature* 408, 538-540.

- Lyman, L. M., Copps, K., Rastelli, L., Kelley, R. L., and Kuroda, M. I. (1997). *Drosophila* male-specific lethal-2 protein: structure/function analysis and dependence on MSL-1 for chromosome association. *Genetics* *147*, 1743-1753.
- Lyon, M. F. (1961). Gene action in the X-chromosome of the mouse (*Mus musculus* L.). *Nature* *190*, 372-373.
- Marin, I. (2003). Evolution of chromatin-remodeling complexes: comparative genomics reveals the ancient origin of "novel" compensasome genes. *J Mol Evol* *56*, 527-539.
- Martens, J. A., and Winston, F. (2003). Recent advances in understanding chromatin remodeling by Swi/Snf complexes. *Curr Opin Genet Dev* *13*, 136-142.
- Mata, X., Taourit, S., and Le Provost, F. (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*, 137-144.
- Matangkasombut, O., and Buratowski, S. (2003). Different sensitivities of bromodomain factors 1 and 2 to histone H4 acetylation. *Mol Cell* *11*, 353-363.
- Mayer, B. J., and Eck, M. J. (1995). SH3 domains. Minding your p's and q's. *Curr Biol* *5*, 364-367.
- Meller, V. H., Gordadze, P. R., Park, Y., Chu, X., Stuckenholtz, C., Kelley, R. L., and Kuroda, M. I. (2000). Ordered assembly of roX RNAs into MSL complexes on the dosage-compensated X chromosome in *Drosophila*. *Curr Biol* *10*, 136-143.
- Meller, V. H., Wu, K. H., Roman, G., Kuroda, M. I., and Davis, R. L. (1997). roX1 RNA paints the X chromosome of male *Drosophila* and is regulated by the dosage compensation system. *Cell* *88*, 445-457.
- Mendjan, S., Taipale, M., Kind, J., Holz, H., Gebhardt, P., Schelder, M., Vermeulen, M., Buscaino, A., Duncan, K., Mueller, J., *et al.* (2006). Nuclear pore components are involved in the transcriptional regulation of dosage compensation in *Drosophila*. *Mol Cell* *21*, 811-823.
- Mersfelder, E. L., and Parthun, M. R. (2006). The tale beyond the tail: histone core domain modifications and the regulation of chromatin structure. *Nucleic Acids Res* *34*, 2653-2662.
- Meyer, B. J. (2000). Sex in the wormcounting and compensating X-chromosome dose. *Trends Genet* *16*, 247-253.
- Min, J., Zhang, Y., and Xu, R. M. (2003). Structural basis for specific binding of Polycomb chromodomain to histone H3 methylated at Lys 27. *Genes Dev* *17*, 1823-1828.
- Mizzen, C. A., Yang, X. J., Kokubo, T., Brownell, J. E., Bannister, A. J., Owen-Hughes, T., Workman, J., Wang, L., Berger, S. L., Kouzarides, T., *et al.* (1996). The TAF(II)250 subunit of TFIID has histone acetyltransferase activity. *Cell* *87*, 1261-1270.
- Mora-Bermudez, F., and Ellenberg, J. (2007). Measuring structural dynamics of chromosomes in living cells by fluorescence microscopy. *Methods* *41*, 158-167.
- Morales, V., Straub, T., Neumann, M. F., Mengus, G., Akhtar, A., and Becker, P. B. (2004). Functional integration of the histone acetyltransferase MOF into the dosage compensation complex. *Embo J* *23*, 2258-2268.
- Morton, C. J., and Campbell, I. D. (1994). SH3 domains. Molecular 'Velcro'. *Curr Biol* *4*, 615-617.
- Murnion, M. E., Adams, R. R., Callister, D. M., Allis, C. D., Earnshaw, W. C., and Swedlow, J. R. (2001). Chromatin-associated protein phosphatase 1 regulates aurora-B and histone H3 phosphorylation. *J Biol Chem* *276*, 26656-26665.
- Narlikar, G. J., Fan, H. Y., and Kingston, R. E. (2002). Cooperation between complexes that regulate chromatin structure and transcription. *Cell* *108*, 475-487.
- Neer, E. J., Schmidt, C. J., Nambudripad, R., and Smith, T. F. (1994). The ancient regulatory-protein family of WD-repeat proteins. *Nature* *371*, 297-300.
- Neigeborn, L., and Carlson, M. (1984). Genes affecting the regulation of SUC2 gene expression by glucose repression in *Saccharomyces cerevisiae*. *Genetics* *108*, 845-858.

- Neuwald, A. F., and Landsman, D. (1997). GCN5-related histone N-acetyltransferases belong to a diverse superfamily that includes the yeast SPT10 protein. *Trends Biochem Sci* 22, 154-155.
- Nielsen, P. R., Nietlispach, D., Buscaino, A., Warner, R. J., Akhtar, A., Murzin, A. G., Murzina, N. V., and Laue, E. D. (2005). Structure of the chromo barrel domain from the MOF acetyltransferase. *J Biol Chem* 280, 32326-32331.
- Nowak, S. J., and Corces, V. G. (2004). Phosphorylation of histone H3: a balancing act between chromosome condensation and transcriptional activation. *Trends Genet* 20, 214-220.
- Nybakken, K., Vokes, S. A., Lin, T. Y., McMahon, A. P., and Perrimon, N. (2005). A genome-wide RNA interference screen in *Drosophila melanogaster* cells for new components of the Hh signaling pathway. *Nat Genet* 37, 1323-1332.
- Ogryzko, V. V., Schiltz, R. L., Russanova, V., Howard, B. H., and Nakatani, Y. (1996). The transcriptional coactivators p300 and CBP are histone acetyltransferases. *Cell* 87, 953-959.
- Oh, H., Bone, J. R., and Kuroda, M. I. (2004). Multiple classes of MSL binding sites target dosage compensation to the X chromosome of *Drosophila*. *Curr Biol* 14, 481-487.
- Ohno, S. (1972). So much "junk" DNA in our genome. *Brookhaven Symp Biol* 23, 366-370.
- Olins, A. L., and Olins, D. E. (1974). Spheroid chromatin units (v bodies). *Science* 183, 330-332.
- Ollis, D. L., Cheah, E., Cygler, M., Dijkstra, B., Frolow, F., Franken, S. M., Harel, M., Remington, S. J., Silman, I., Schrag, J., and et al. (1992). The alpha/beta hydrolase fold. *Protein Eng* 5, 197-211.
- Owen-Hughes, T., Utley, R. T., Cote, J., Peterson, C. L., and Workman, J. L. (1996). Persistent site-specific remodeling of a nucleosome array by transient action of the SWI/SNF complex. *Science* 273, 513-516.
- Palmer, M. J., Mergner, V. A., Richman, R., Manning, J. E., Kuroda, M. I., and Lucchesi, J. C. (1993). The male-specific lethal-one (*msl-1*) gene of *Drosophila melanogaster* encodes a novel protein that associates with the X chromosome in males. *Genetics* 134, 545-557.
- Pelham, R. J., and Chang, F. (2002). Actin dynamics in the contractile ring during cytokinesis in fission yeast. *Nature* 419, 82-86.
- Peters, A. H., Kubicek, S., Mechtler, K., O'Sullivan, R. J., Derijck, A. A., Perez-Burgos, L., Kohlmaier, A., Opravil, S., Tachibana, M., Shinkai, Y., et al. (2003). Partitioning and plasticity of repressive histone methylation states in mammalian chromatin. *Mol Cell* 12, 1577-1589.
- Peterson, C. L., Dingwall, A., and Scott, M. P. (1994). Five SWI/SNF gene products are components of a large multisubunit complex required for transcriptional enhancement. *Proc Natl Acad Sci U S A* 91, 2905-2908.
- Peterson, C. L., and Herskowitz, I. (1992). Characterization of the yeast SWI1, SWI2, and SWI3 genes, which encode a global activator of transcription. *Cell* 68, 573-583.
- Pollard, T. D., and Borisy, G. G. (2003). Cellular motility driven by assembly and disassembly of actin filaments. *Cell* 112, 453-465.
- Ponting, C. P. (1997). Tudor domains in proteins that interact with RNA. *Trends Biochem Sci* 22, 51-52.
- Rath, U., Wang, D., Ding, Y., Xu, Y. Z., Qi, H., Blacketer, M. J., Girton, J., Johansen, J., and Johansen, K. M. (2004). Chromator, a novel and essential chromodomain protein interacts directly with the putative spindle matrix protein skeleton. *J Cell Biochem* 93, 1033-1047.
- Rea, S., Xouri, G., and Akhtar, A. (2007). Males absent on the first (MOF): from flies to humans. *Oncogene* 26, 5385-5394.
- Reenan, R. A., Hanrahan, C. J., and Barry, G. (2000). The *mle*(*napts*) RNA helicase mutation in *drosophila* results in a splicing catastrophe of the para Na⁺ channel transcript in a region of RNA editing. *Neuron* 25, 139-149.

- Reinhart, B. J., and Bartel, D. P. (2002). Small RNAs correspond to centromere heterochromatic repeats. *Science* **297**, 1831.
- Richardson, J. S. (1981). The anatomy and taxonomy of protein structure. *Adv Protein Chem* **34**, 167-339.
- Rigaut, G., Shevchenko, A., Rutz, B., Wilm, M., Mann, M., and Seraphin, B. (1999). A generic protein purification method for protein complex characterization and proteome exploration. *Nat Biotechnol* **17**, 1030-1032.
- Ringrose, L., Ehret, H., and Paro, R. (2004). Distinct contributions of histone H3 lysine 9 and 27 methylation to locus-specific stability of polycomb complexes. *Mol Cell* **16**, 641-653.
- Robinson, P. J., Fairall, L., Huynh, V. A., and Rhodes, D. (2006). EM measurements define the dimensions of the "30-nm" chromatin fiber: evidence for a compact, interdigitated structure. *Proc Natl Acad Sci U S A* **103**, 6506-6511.
- Rogakou, E. P., Pilch, D. R., Orr, A. H., Ivanova, V. S., and Bonner, W. M. (1998). DNA double-stranded breaks induce histone H2AX phosphorylation on serine 139. *J Biol Chem* **273**, 5858-5868.
- Rojas, J. R., Trievel, R. C., Zhou, J., Mo, Y., Li, X., Berger, S. L., Allis, C. D., and Marmorstein, R. (1999). Structure of Tetrahymena GCN5 bound to coenzyme A and a histone H3 peptide. *Nature* **401**, 93-98.
- Roloff, T. C., Ropers, H. H., and Nuber, U. A. (2003). Comparative study of methyl-CpG-binding domain proteins. *BMC Genomics* **4**, 1.
- Saffery, R., Sumer, H., Hassan, S., Wong, L. H., Craig, J. M., Todokoro, K., Anderson, M., Stafford, A., and Choo, K. H. (2003). Transcription within a functional human centromere. *Mol Cell* **12**, 509-516.
- Saha, A., Wittmeyer, J., and Cairns, B. R. (2006). Chromatin remodelling: the industrial revolution of DNA around histones. *Nat Rev Mol Cell Biol* **7**, 437-447.
- Santisteban, M. S., Kalashnikova, T., and Smith, M. M. (2000). Histone H2A.Z regulates transcription and is partially redundant with nucleosome remodeling complexes. *Cell* **103**, 411-422.
- Santos-Rosa, H., Schneider, R., Bannister, A. J., Sherriff, J., Bernstein, B. E., Emre, N. C., Schreiber, S. L., Mellor, J., and Kouzarides, T. (2002). Active genes are trimethylated at K4 of histone H3. *Nature* **419**, 407-411.
- Schreiber, S. L., and Bernstein, B. E. (2002). Signaling network model of chromatin. *Cell* **111**, 771-778.
- Schwartz, B. E., and Ahmad, K. (2005). Transcriptional activation triggers deposition and removal of the histone variant H3.3. *Genes Dev* **19**, 804-814.
- Shevchenko, A., Wilm, M., Vorm, O., and Mann, M. (1996). Mass spectrometric sequencing of proteins silver-stained polyacrylamide gels. *Anal Chem* **68**, 850-858.
- Shi, Y., Lan, F., Matson, C., Mulligan, P., Whetstine, J. R., Cole, P. A., and Casero, R. A. (2004). Histone demethylation mediated by the nuclear amine oxidase homolog LSD1. *Cell* **119**, 941-953.
- Shim, S., Yoon, C. S., and Han, J. K. (2000). A novel gene family with a developmentally regulated expression in *Xenopus laevis*. *Biochem Biophys Res Commun* **267**, 558-564.
- Shogren-Knaak, M., Ishii, H., Sun, J. M., Pazin, M. J., Davie, J. R., and Peterson, C. L. (2006). Histone H4-K16 acetylation controls chromatin structure and protein interactions. *Science* **311**, 844-847.
- Smith, C. L., and Peterson, C. L. (2005). ATP-dependent chromatin remodeling. *Curr Top Dev Biol* **65**, 115-148.
- Smith, E. R., Pannuti, A., Gu, W., Steurnagel, A., Cook, R. G., Allis, C. D., and Lucchesi, J. C. (2000). The drosophila MSL complex acetylates histone H4 at lysine 16, a chromatin modification linked to dosage compensation. *Mol Cell Biol* **20**, 312-318.
- Smith, M. M. (2002). Centromeres and variant histones: what, where, when and why? *Curr Opin Cell Biol* **14**, 279-285.

- Smith, S., and Stillman, B. (1989). Purification and characterization of CAF-I, a human cell factor required for chromatin assembly during DNA replication in vitro. *Cell* 58, 15-25.
- Sobel, R. E., Cook, R. G., Perry, C. A., Annunziato, A. T., and Allis, C. D. (1995). Conservation of deposition-related acetylation sites in newly synthesized histones H3 and H4. *Proc Natl Acad Sci U S A* 92, 1237-1241.
- Sondek, J., Bohm, A., Lambright, D. G., Hamm, H. E., and Sigler, P. B. (1996). Crystal structure of a G-protein beta gamma dimer at 2.1A resolution. *Nature* 379, 369-374.
- Song, H., Li, Y., Chen, G., Xing, Z., Zhao, J., Yokoyama, K. K., Li, T., and Zhao, M. (2004). Human MCRS2, a cell-cycle-dependent protein, associates with LPTS/PinX1 and reduces the telomere length. *Biochem Biophys Res Commun* 316, 1116-1123.
- Stern, M., Jensen, R., and Herskowitz, I. (1984). Five SWI genes are required for expression of the HO gene in yeast. *J Mol Biol* 178, 853-868.
- Strahl, B. D., and Allis, C. D. (2000). The language of covalent histone modifications. *Nature* 403, 41-45.
- Strahl, B. D., Briggs, S. D., Brame, C. J., Caldwell, J. A., Koh, S. S., Ma, H., Cook, R. G., Shabanowitz, J., Hunt, D. F., Stallcup, M. R., and Allis, C. D. (2001). Methylation of histone H4 at arginine 3 occurs in vivo and is mediated by the nuclear receptor coactivator PRMT1. *Curr Biol* 11, 996-1000.
- Strunnikov, A. V., and Jessberger, R. (1999). Structural maintenance of chromosomes (SMC) proteins: conserved molecular properties for multiple biological functions. *Eur J Biochem* 263, 6-13.
- Taipale, M. (2005). On the histone acetyltransferase hMOF. E-Thesis, University of Helsinki.
- Taipale, M., and Akhtar, A. (2005a). Chromatin mechanisms in Drosophila dosage compensation. *Prog Mol Subcell Biol* 38, 123-149.
- Taipale, M., Rea, S., Richter, K., Vilar, A., Lichter, P., Imhof, A., and Akhtar, A. (2005b). hMOF histone acetyltransferase is required for histone H4 lysine 16 acetylation in mammalian cells. *Mol Cell Biol* 25, 6798-6810.
- Takechi, S., and Nakayama, T. (1999). Sas3 is a histone acetyltransferase and requires a zinc finger motif. *Biochem Biophys Res Commun* 266, 405-410.
- Tan, S. (2001). One HAT size fits all? *Nat Struct Biol* 8, 8-10.
- Tanner, K. G., Langer, M. R., Kim, Y., and Denu, J. M. (2000). Kinetic mechanism of the histone acetyltransferase GCN5 from yeast. *J Biol Chem* 275, 22048-22055.
- Taunton, J., Hassig, C. A., and Schreiber, S. L. (1996). A mammalian histone deacetylase related to the yeast transcriptional regulator Rpd3p. *Science* 272, 408-411.
- Thompson, P. R., Kurooka, H., Nakatani, Y., and Cole, P. A. (2001). Transcriptional coactivator protein p300. Kinetic characterization of its histone acetyltransferase activity. *J Biol Chem* 276, 33721-33729.
- Trievel, R. C., Rojas, J. R., Sterner, D. E., Venkataramani, R. N., Wang, L., Zhou, J., Allis, C. D., Berger, S. L., and Marmorstein, R. (1999). Crystal structure and mechanism of histone acetylation of the yeast GCN5 transcriptional coactivator. *Proc Natl Acad Sci U S A* 96, 8931-8936.
- Tsukada, Y., Fang, J., Erdjument-Bromage, H., Warren, M. E., Borchers, C. H., Tempst, P., and Zhang, Y. (2006). Histone demethylation by a family of JmjC domain-containing proteins. *Nature* 439, 811-816.
- Turner, B. M. (1993). Decoding the nucleosome. *Cell* 75, 5-8.
- Turner, B. M. (2000). Histone acetylation and an epigenetic code. *Bioessays* 22, 836-845.
- Turner, B. M., Birley, A. J., and Lavender, J. (1992). Histone H4 isoforms acetylated at specific lysine residues define individual chromosomes and chromatin domains in Drosophila polytene nuclei. *Cell* 69, 375-384.
- Uhlmann, F. (2004). The mechanism of sister chromatid cohesion. *Exp Cell Res* 296, 80-85.

- Uhlmann, F., Lottspeich, F., and Nasmyth, K. (1999). Sister-chromatid separation at anaphase onset is promoted by cleavage of the cohesin subunit Scc1. *Nature* *400*, 37-42.
- Utlej, R. T., and Cote, J. (2003). The MYST family of histone acetyltransferases. *Curr Top Microbiol Immunol* *274*, 203-236.
- van Leeuwen, F., and Gottschling, D. E. (2002). Genome-wide histone modifications: gaining specificity by preventing promiscuity. *Curr Opin Cell Biol* *14*, 756-762.
- Varga-Weisz, P. D., Wilm, M., Bonte, E., Dumas, K., Mann, M., and Becker, P. B. (1997). Chromatin-remodelling factor CHRAC contains the ATPases ISWI and topoisomerase II. *Nature* *388*, 598-602.
- Vignali, M., Hassan, A. H., Neely, K. E., and Workman, J. L. (2000). ATP-dependent chromatin-remodeling complexes. *Mol Cell Biol* *20*, 1899-1910.
- Vogelauer, M., Wu, J., Suka, N., and Grunstein, M. (2000). Global histone acetylation and deacetylation in yeast. *Nature* *408*, 495-498.
- Volpe, T. A., Kidner, C., Hall, I. M., Teng, G., Grewal, S. I., and Martienssen, R. A. (2002). Regulation of heterochromatic silencing and histone H3 lysine-9 methylation by RNAi. *Science* *297*, 1833-1837.
- Wang, H., Huang, Z. Q., Xia, L., Feng, Q., Erdjument-Bromage, H., Strahl, B. D., Briggs, S. D., Allis, C. D., Wong, J., Tempst, P., and Zhang, Y. (2001). Methylation of histone H4 at arginine 3 facilitating transcriptional activation by nuclear hormone receptor. *Science* *293*, 853-857.
- Waterborg, J. H. (2002). Dynamics of histone acetylation in vivo. A function for acetylation turnover? *Biochem Cell Biol* *80*, 363-378.
- Wetlaufer, D. B. (1973). Nucleation, rapid folding, and globular intrachain regions in proteins. *Proc Natl Acad Sci U S A* *70*, 697-701.
- Whetstine, J. R., Nottke, A., Lan, F., Huarte, M., Smolikov, S., Chen, Z., Spooner, E., Li, E., Zhang, G., Colaiacovo, M., and Shi, Y. (2006). Reversal of histone lysine trimethylation by the JMJD2 family of histone demethylases. *Cell* *125*, 467-481.
- Wolf, E., Vassilev, A., Makino, Y., Sali, A., Nakatani, Y., and Burley, S. K. (1998). Crystal structure of a GCN5-related N-acetyltransferase: *Serratia marcescens* aminoglycoside 3-N-acetyltransferase. *Cell* *94*, 439-449.
- Wysocka, J., Swigut, T., Milne, T. A., Dou, Y., Zhang, X., Burlingame, A. L., Roeder, R. G., Brivanlou, A. H., and Allis, C. D. (2005). WDR5 associates with histone H3 methylated at K4 and is essential for H3 K4 methylation and vertebrate development. *Cell* *121*, 859-872.
- Yamamoto, T., and Horikoshi, M. (1997). Novel substrate specificity of the histone acetyltransferase activity of HIV-1-Tat interactive protein Tip60. *J Biol Chem* *272*, 30595-30598.
- Yang, X. J. (2004). The diverse superfamily of lysine acetyltransferases and their roles in leukemia and other diseases. *Nucleic Acids Res* *32*, 959-976.
- Yang, X. J., Ogryzko, V. V., Nishikawa, J., Howard, B. H., and Nakatani, Y. (1996). A p300/CBP-associated factor that competes with the adenoviral oncoprotein E1A. *Nature* *382*, 319-324.
- Zeng, L., and Zhou, M. M. (2002). Bromodomain: an acetyl-lysine binding domain. *FEBS Lett* *513*, 124-128.
- Zhang, L., Eugeni, E. E., Parthun, M. R., and Freitas, M. A. (2003). Identification of novel histone post-translational modifications by peptide mass fingerprinting. *Chromosoma* *112*, 77-86.
- Zhang, Y., and Reinberg, D. (2001). Transcription regulation by histone methylation: interplay between different covalent modifications of the core histone tails. *Genes Dev* *15*, 2343-2360.

ABBREVIATIONS

ACF	ATP-utilizing chromatin assembly and remodeling factor
ATP	adenosine triphosphate
CAF-1	chromatin assembly factor 1
CBD	chromo-barrel domain
CBP	calmodulin-binding peptide
cDNA	complementary DNA
CENP-A	centromere protein A
ChIP	chromatin immunoprecipitation
Chr	chromator
CHRAC	chromatin accessibility complex
CoA	co-enzyme A
CoIP	coimmunoprecipitation
DCC	dosage compensation complex
DNA	deoxyribonucleic acid
DREF	DNA replication element factor
dsRNA	double-stranded RNA
EGFP	enhanced green fluorescent protein
eIF4e	eukaryotic translation initiation factor 1
EM	electron microscope
FACS	fluorescence activated cell sorting
FLAG	protein tag with the peptide sequence DYKDDDDK
GAPDH	glyceraldehyde 3-phosphate dehydrogenase
GCN5	general control non-derepressible 5
GNAT	Gcn5-related acetyltransferase
GST	glutathione-S-transferase
HA	influenza hemagglutinin
HAT	histone acetyltransferase
Hat1p	histone acetyltransferase 1
HDAC	histone deacetylase
HP1	hetero-chromatin protein 1
HPLC	high-performance liquid chromatography
HSP70	heat shock protein 70
IgG	immunoglobulin G
IRR	ion intensity ratio
IPTG	isopropyl- β -D-thiogalactopyranosid
ISWI	imitation switch
kDa	kilodalton
LC	liquid chromatography
LC-MS/MS	liquid chromatography tandem mass spectrometry
LSD1	lysine-specific demethylase 1
MBD	methyl CpG binding domain
MCRS2	microspherule protein 2
MLE	maleless
MOF	males absent on the first
mRNA	messenger RNA
MS	mass spectrometry
MSL	male-specific lethal
MT	methyltransferase
MTOR	megator
MYST	MOZ/YBF2/SAS2/TIP60
NSL	non-specific lethal
NUP153	nucleoporin 153
NURF	nucleosome-remodeling factor
o/n	over night
PADI 4	peptidyl arginine deiminase 4
PC	polycomp
PCAF	p300/CBP associated factor
PCR	polymerase chain reaction

qRT-PCR	quantitative real-time PCR
RNA	ribonucleic acid
RNAi	RNA interference
roX	RNA on the X
RT	room temperature
S2	Schneider 2 cell line
SAGA	Spt-Ada-Gcn5-acetyltransferase
SAS2	something about silencing 2
SDS-PAGE	sodium dodecylsulfate polyacrylamide gel electrophoresis
SH3	src homology-3
SIR	silent information regulator
SMC	structural maintenance of chromosomes
SNF	sucrose-non-fermenter
SWI	switch
Sxl	sex-lethal
TAP	tandem affinity purification
TBP	TATA-binding protein
TEV	tobacco etch virus
TFH-NSL1	TAP/FLAG/HA-tagged NSL1 protein
TIP60	Tat-interacting protein 60
UAS	upstream activating sequence
URS	upstream repressive sequence
WDS	will die slowly

PUBLICATIONS

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

A printout of the publication is attached.

I presented parts of my work as a poster with the title 'Purification and biochemical characterization of novel MOF-containing complexes' at the 'EMBO Conference on Chromatin and Epigenetics', EMBL Heidelberg, Germany, May 3 - 6, 2007.

ACKNOWLEDGEMENTS

First of all, I would like to thank Asifa Akhtar for giving me the opportunity of doing a PhD in her laboratory and for introducing me to the exciting world of chromatin. It was an intriguing experience that at almost any time of the day one could push open the doors to the office and participate in your enthusiasm for science.

I am thankful to all the members of my Thesis Advisory Committee, as well as my Thesis Defence Committee for their interest in my work and for their invaluable advice: Prof. Dr. Gabriele Petersen, Prof. Dr. Matthias Hentze, Dr. Michael Sattler and Dr. Klaus Scheffzek.

I would like to thank all former and present members of the Akhtar lab for their help, support and inspiring discussions. I have enjoyed and benefited to work with you! I am especially grateful to Herbert Holz for his constant interest, support and encouragement during the course of this work. He has become a real good friend during my time being in the lab.

This work would have been impossible without the help of our mass spec experts, Marc Gentzel and Sven Fraterman, who dedicated a large part of their time to the successful completion of our joint projects. In addition, I want to thank Gunter Stier who was initially teaching me all of the subtle tricks for efficient protein expression and purification.

Furthermore, I would like to thank Julia Willingale-Theune and Alexandra Manaia who always supported my interest in teaching students and teachers during their unique Learning Labs. These opportunities have been outstanding experiences for me.

I am very grateful that I received a Marie-Curie E-STAR fellowship which provided me with the necessary freedom to go to conferences and get in touch with other science-related activities.

I am most indebted to my little family for their love and support. You, Yvonne and Leonie, deserve all my gratitude! Starting up in Heidelberg with only the two of us was an exciting experience – it has become even livelier since Leonie was joining us;) Without you, I would not have managed to go through this!

And last but not least, I thank my parents, my brother and his family and all of our friends for their understanding, their advice and their continuous encouragement.