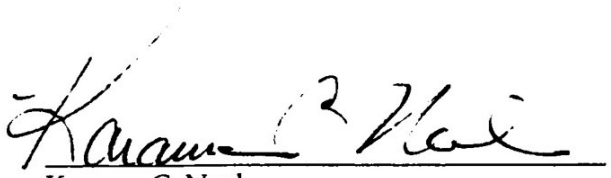
In presenting this dissertation as a partial fulfillment of the requirements for an advanced degree from Emory University, I agree that the Library of the University shall make it available for inspection and circulation in accordance with its regulations, governing materials of this type. I agree that permission to copy from, or to publish, this dissertation may be granted by the professor under whose direction it was written, or, in his absence, by the Dean of the Graduate School when such copying or publication is solely for scholarly purposes and does not involve potential financial gain. It is understood that any copying from or publication of this dissertation which involves potential financial gain will not be allowed without written permission.



Karama C. Neal

Characterization of Human and Drosophila Homologues
of Two Drosophila Dosage Compensation Genes

By

Karama Carrol Neal
Doctor of Philosophy

Program in Genetics and Molecular Biology
Graduate Division of Biological and Biomedical Sciences

John C. Lucchesi
Advisor

Jeremy Boss
Committee Member

Victoria Finnerty
Committee Member

Sue Jinks-Robertson
Committee Member

Barry Yedvobnick
Committee Member

Accepted:

Dean of the Graduate School

August 5, 2000
June 19, 2000

Characterization of Human and Drosophila Homologues of Two Drosophila Dosage Compensation Genes

By

Karama Carrol Neal
B.A., Swarthmore College, 1993

Advisor: John C. Lucchesi, Ph.D.

An abstract of
a dissertation submitted to the faculty of the
Graduate School of Emory University
in partial fulfillment of the requirements
for the degree of
Doctor of Philosophy

Program in Genetics and Molecular Biology
Graduate Division of Biological and Biomedical Sciences

2000

ABSTRACT

Dosage compensation is the mechanism by which the amount of X-linked gene product is equalized between the males and females. Drosophila accomplish this by a two fold hypertranscription of the X chromosome in males. Five proteins are known to regulate this process and functional absence of any one of the five causes male specific lethality. Recently, the genes known to be involved in dosage compensation, collectively known as male specific lethals or MSLs, have been cloned and characterized. Biochemical analyses of the Drosophila dosage compensation machinery have linked this process to more widely conserved processes of chromatin modification and remodeling. The MSLs are associated in a multiprotein complex that binds hundreds of sites on the X chromosome of male flies. The male X chromatin also has a specific isoform of histone H4, acetylated at lysine 16. The MSL protein MOF, a histone acetyltransferase (HAT), has been shown to be responsible for this modification. It is thought that this modification, in combination with the activities of the other MSLs including an RNA helicase, results in the hypertranscription of X-linked genes in males.

Several of these MSL genes are evolutionarily conserved. Homologues of the RNA helicase MLE are found in mammals. Additionally, homologues of MOF, a MYST family histone acetyltransferase, and MSL3, a chromo and chromoshadow domain-containing protein, have been found in yeasts, mammals and flies. Because of the conserved nature of these proteins, I hypothesize that MYST family HATs associate specifically with MSL3-like proteins, as is the case in the Drosophila dosage compensation complex. This work explores this

hypothesis by examining activity, localization and interaction partners of MOF and MSL3 homologues in humans and Drosophila. Specifically, I present evidence that the Drosophila homologues are present in multiprotein complexes and characterize the HAT activity of hMOF, a human MYST HAT and homologue of Drosophila MOF. Possible cellular roles for the human and Drosophila proteins are discussed.

Characterization of Human and Drosophila Homologues of Two Drosophila Dosage Compensation Genes

By

Karama Carrol Neal
B.A., Swarthmore College, 1993

Advisor: John C. Lucchesi, Ph.D.

A dissertation submitted to the faculty of the
Graduate School of Emory University
in partial fulfillment of the requirements
for the degree of
Doctor of Philosophy

Program in Genetics and Molecular Biology
Graduate Division of Biological and Biomedical Sciences

2000

ACKNOWLEDGEMENTS

This Ph.D. has only been possible through the support and encouragement of my friends, relatives, and colleagues.

Thanks to my advisor, John Lucchesi, for his constant support and for providing me with the tools necessary to complete this work. Thanks also to the members of the Lucchesi lab, past and present, for their encouragement and for making the lab such a pleasant place to work.

I am thankful to Annice Steadman, who taught me biology and taught me to teach biology. I am most grateful to Helen Benes, for training me to do science and to love the lab and to Helen Benes, Laurie Tompkins and Marius Sudol for treating me with respect and always assuming that I would succeed.

I thank my mother, Janet Cobb, for her undying love and belief in my abilities. I thank her for having the good sense to give me a microscope so that I could start my scientific career at the age of six. I am grateful to my grandmother, Fanilla Cobb, who always already loves me, even in her absence. Thanks to my father, Olly Neal, who never doubted that I would succeed and taught me never to give up.

Obrigada Jim Wallace and gracias Althea Grant, both of whom are always there for me. Thanks to my friends Althea Grant, Carolyn Monteilh, Chris Wilson, Triscia Wharton Hendrickson, Ollie Kelley Appleberry and Michelle Johnson who not only sympathized but empathized with my every struggle and success.

Thanks to the dancers and drummers of the Afrikinetic West African Dance Class and the Uhuru West African Dance Workshop for keeping me healthy in body and soul.

A warm and hearty "thank you!" to all my friends and family for their continuous love and confidence in me.

I am, of course, most grateful to the Goddess who continues to guide my life in ever more wonderful directions.

Asante sana.

DEDICATION

This work is dedicated to the ~60 million people who died in transit from Africa to the Americas and to four children of those who survived:

Willie Beatrice Jones Neal (1912 – 1968)

Rev. Ollie Neal (1896 – 1978)

Fanilla Suttles Cobb (1910 – 1997)

Marshall Cobb (1902 – 1957)

TABLE OF CONTENTS

LIST OF FIGURES AND TABLES

CHAPTER 1

Introduction

Nature has devised several mechanisms of determining sex in organisms that undergo sexual reproduction. In addition to environmental mechanisms that depend on temperature (in turtles and crocodilians) or physical location of the embryo (echinuroid worms and slipper snails) (Gilbert, 2000), there are mechanisms of sex determination that depend on the Mendelian segregation of certain genes which are sex determining factors (Wilson, 1905). Often, as in the cases of mammals, insects, nematodes, and birds among others, these genes are localized to a particular chromosome. Because the region of the chromosome bearing the sex-determining factor is isolated in a particular sex, over evolutionary time this chromosome can become genetically and morphologically distinct from its homologue (Ohno, 1967). When this happens a problem arises: since one sex has two copies of the non-differentiated chromosome and the other has only one, there can exist an inequality in the amount of gene product between the two sexes. This is particularly relevant if there are genes on the chromosome that are equally important in both sexes specifically genes, *i.e.* genes that do not have only sex-specific functions. If, for example, males have a single X chromosome and females have two, there must exist some mechanism to equalize the amount of X-linked gene product between the sexes. This equalization process is called dosage compensation.

A. Dosage compensation in Drosophila

Discovery and initial characterization: Dosage compensation was first noted in Drosophila by Muller (1932) when he observed that females with two copies of *white apricot* (w^a), a hypomorphic allele of the gene *white* causing an eye color defect, have identical eye color to males with one copy of the gene. Additionally, he saw that males with a duplication of this gene had much darker eyes than did wild type females with their normal two copies (Figure 1-1). This work showed that even though there is a dosage response in each sex, one dose of the gene in males is equal to two doses in females thus providing the first evidence of the existence of a dosage compensation mechanism.

Dosage compensation could occur either in males or females. In females, it could be achieved by a mechanism similar to that of mammalian dosage compensation. In mammals, dosage compensation occurs by inactivation of one of the two X chromosomes in each cell of the female body (for a review see Heard *et al.*, 1997). Because this inactivation is random (in eutherian mammals), heterozygosity for X-linked genes with cell autonomous products results in a mosaic phenotype. Analyses of X-linked genes with cell-autonomous expression, such as *yellow* and *forked* in Drosophila females revealed no evidence of this type of mosaicism. This

suggested that X inactivation is not the dosage compensation mechanism that is operative in Drosophila.

An additional set of observations supported this conclusion. The X-linked enzyme 6-phosphogluconate dehydrogenase (6PGD) has multiple electrophoretic variants and exists in the cell as a dimer. If a female fly is heterozygous at the 6PGD locus and has one of her X-chromosomes inactivated, there would only be two types of 6PGD dimer formed. This however, is not what is observed. Instead, three types of dimers are formed (Kazazian *et al.*, 1965; Young, 1966). The simplest explanation of the data is that both copies of the gene are active within each cell and the two gene products pair randomly. Taken together these data strongly suggested that both X chromosomes in the female fly are active.

Evidence that dosage compensation in Drosophila proceeds by a transcription-based mechanism was provided by Mukerjee and Beermann (1965), who measured levels of tritiated uridine incorporation by salivary gland polytene chromosomes and showed that the amount of incorporation on the single X chromosome in males is equivalent to the combined level of incorporation observed on the X chromosomes in the female.

Geneticists later identified four autosomal loci that result in male specific lethality. These are *maleless* (Golubowsky and Ivanov, 1972; Fukanaga *et al.* 1975; Tanaka *et al.*, 1976), *male specific lethals 1* and 2 (Belote and Lucchesi, 1980a, b) and *maleless on the third* (Uchida,1981),. These are abbreviated *mle*, *msl1*, *msl2* and *msl3* respectively and are collectively called

the msls. A fifth msl gene, called *mof* for *males absent on the first*, was identified in 1997 (Hilfiker *et al.*, 1997). Loss of function mutations of each of these genes result in the death of all males before or at the early pupal stage. Measurement of tritiated uridine incorporation into RNA transcribed from salivary gland chromosomes and X-linked enzyme levels in msl mutant larvae (Belote and Lucchesi, 1980a) and measurement of steady state levels of specific X-linked gene transcripts (Breen and Lucchesi, 1986) showed levels (of RNA or enzyme activity) reduced by 50 – 60% in mutant male larvae. It was then proposed that the products of the msl genes regulate the dosage compensation mechanism (Belote and Lucchesi, 1980a; Lucchesi, 1983). The msl genes were subsequently cloned and characterized and their protein products were shown to associate with hundreds of sites along the X chromosome of male flies (Kuroda *et al.*, 1991; Palmer *et al.*, 1993; Gorman *et al.*, 1995; Zhou *et al.*, 1995; Kelley *et al.*, 1995; Bashaw and Baker, 1995 and Hilfiker *et al.*, 1997) supporting their involvement in dosage compensation. All five MSL proteins exhibit the same distribution along the X chromosome and association of any one of them with the X-chromatin depends on the presence and functional integrity of the others. These observations suggested that the MSLs form a multiprotein complex , a conclusion that was recently demonstrated to be true (Smith *et al.*, 2000).

The MSL proteins: The MSL1 protein has an acidic N terminus and a coiled-coil domain (Palmer *et al.*, 1993). Both of these are features often found

in proteins involved in transcriptional regulation. The MSL2 protein (Zhou *et al.*, 1995; Kelley *et al.*, 1995; Bashaw and Baker, 1995) has a metallothionein domain and a RING finger domain, the latter of which may be involved in protein/protein interaction. To date, there have been no homologues of MSL1 or MSL2 identified in Drosophila or other organisms.

The MSL3 protein is characterized by the presence of a chromo domain and a chromoshadow domain (Koonin *et al.*, 1995). The evolutionarily conserved *chr*omatin *o*rganization *mo*difier domain was first identified based on homology between the Drosophila proteins Polycomb (Pc) and heterochromatin protein 1 (HP1) (Paro and Hogness, 1991). Pc is required for appropriate silencing of the homeotic genes (Paro, 1990) while HP1 is necessary for heterochromatic silencing (Eissenberg *et al.*, 1990). Both of these proteins function in large multiprotein complexes and are involved in the silencing of large genetic regions. The chromodomain is thought to be important for protein-protein interactions; *i.e.* it may "be a vehicle that delivers both positive and negative transcription regulators to the sites of their action on chromatin" (Koonin *et al.*, 1995). It may be in this role which MSL3 functions. Some chromodomain containing proteins, such as HP1 and MSL3, also have a so-called chromoshadow domain found C terminal to the chromodomain. The chromoshadow domain is loosely related to the chromodomain by sequence and appears to be specific to proteins involved in the establishment or maintenance of the chromatin state (Aasland and Stewart, 1995; Koonin *et al.*, 1995).

Several homologues of MSL3 have been identified in a variety of organisms, including yeasts and mammals. Additionally, there are proteins very similar to MSL3, but which represent a distinct group (the MRG family) which are found in *C. elegans*, Drosophila and mammals (Bertram *et al.*, 1999). The only one among these related proteins to which a function has been assigned is MSL3.

The MLE protein is a nucleic acid helicase with associated ATPase activity (Lee *et al.*, 1997; Kuroda *et al.*, 1991). MLE is a member of the DEAD/H box family of RNA helicases. These proteins have been shown to be involved in many stages of RNA metabolism. Like other helicases, the DEAD/H box helicases unwind double stranded nucleic acid (RNA:RNA, RNA:DNA, or DNA:DNA) by binding to one of the strands and using the energy released from hydrolysis of ATP to translocate along the strand and unwind the helix (Gibson and Thompson, 1994). These proteins are found in all organisms including viruses, archaebacteria, bacteria, yeast, flies, and vertebrates. They have been implicated in DNA repair, transcription, translation, ribosome assembly, and RNA splicing (reviewed in Eisen and Lucchesi, 1998). Recently, mutations in genes that encode helicases or helicase-like proteins have been shown to cause diseases such as Werner syndrome (Gray *et al.*, 1997) and Bloom syndrome (Ellis *et al.*, 1995). Another Drosophila DEAD/H box helicase is described in Appendix B of this volume (Eisen, 1998).

Three mammalian homologues of MLE have been identified: bovine nuclear DNA helicase II (NDHII) (Zhang *et al.*, 1995), human RNA helicase A (RHA) (Lee and Hurwitz, 1993) and murine RNA helicase A (Lee *et al.*, 1998b). Human RHA and MLE have been shown to be biochemically equivalent (Lee *et al.*, 1997). Targeted disruption of the murine RHA results in animals that have problems completing gastrulation and die by embryonic day 9 (Lee *et al.*, 1998a).

In contrast to Drosophila MLE, which has neither a chromoshadow or chromodomain, the conserved CHD (*c*hromatin and *h*elicase *d*omain) proteins have both a chromodomain and an active helicase domain combined in the same molecule (Stokes and Perry, 1995; Woodage *et al.*, 1997). Members of the CHD family are found in the yeasts *S. pombe* and *S. cerevisiae*, in *C. elegans* and *D. melanogaster* and in birds and mammals (Woodage *et al.*, 1997; Ellergen, 1996; Griffiths and Korn, 1996). The murine CHD1 protein is associated with chromatin and its intracellular localization is cell cycle regulated (Woodage *et al.*, 1997) but the function of this class of proteins is still not well understood.

The 1992 observation by Turner and colleagues that the X chromosome of Drosophila males is specifically enriched in an isoform of histone H4 acetylated at lysine 16 (H4AcK16), suggested that histone acetyltransferases might play a role in the dosage compensation mechanism. Evidence that this is the case was produced in our laboratory by Andres Hilfiker and colleagues (1997), who cloned the fifth male specific lethal gene, *mof*, and showed it to

encode a protein homologous to histone acetyltransferases. The *mof* gene maps to the 5C region of the X chromosome which made it somewhat difficult to isolate genetically: without certain genetic manipulations, it is difficult to determine if an X-linked lethal is a male *specific* lethal or simply and more commonly a general lethal affecting both males and females. Recently, Smith *et al.* (2000) have shown that MOF is directly responsible for the presence of H4AcK16 in the X-chromosome chromatin of Drosophila males.

MOF is a member of the MYST family of histone acetyltransferases (HATs). The family is named for the initially described members: human *MOZ*, yeast *YBF2/SAS3*, yeast *SAS2* and human *Tip60*. The MYST proteins that have been shown to have HAT activity tend to prefer to acetylate histone H4 with a lesser affinity for histones H3 and H2A (Neal *et al.*, 2000; Smith *et al.*, 2000; Yamamoto and Horikoshi, 1997; Smith *et al.*, 1998). The number of known members of the MYST family is increasing with additional MYST HATs identified in *S. cerevisiae* (ESA1), *S. pombe* (GenBank accession number Z69795), *C. elegans* (GenBank accession number Z752512) humans (HBO1) and, intriguingly, the carrot *D. carota* (GenBank accession number BAA32822) (Neal *et al.*, 2000; Hilfiker *et al.*, 1997; Iizuka and Stillman, 1999). All of these proteins are similar over a large region that includes the putative acetyl Co-enzyme A binding site (Lu, 1996). The mutation that allowed the identification of the *mof* gene occurs in a conserved glycine residue in the region thought to bind acetyl Co-enzyme A. Several of the MYST proteins

including MOF and Tip60 have a single zinc finger just N-terminal to the acetyl CoA binding site. The zinc finger in SAS3 was recently shown to be required for HAT activity (Takechi and Nakayama, 1999). MOF, Tip60 and ESA1 also have a second smaller region of homology N-terminal to the first, which encompasses a chromodomain not found in some of the other MYST HATs.

Untranslated RNA components: Richter and colleagues reported that following treatment with RNase, MLE no longer associated with the X chromosome in male salivary gland preparations (Richter *et al.*, 1996). This observation was followed by the discovery of two X-linked untranslated RNAs that bind specifically to the X chromosome in males (Amrein and Axel, 1997; Meller *et al.*, 1997). The genes that code for these untranslated RNAs are termed *roX1* and *roX2*, for *R*NA *o*n the *X*. Expression and stability of the roXs are under genetic control of the dosage compensation genes (Amrein and Axel, 1997). These two RNAs are found along the X chromosome and coat it in a pattern indistinguishable from that of the MSL proteins. These RNAs are indeed components of the MSL complex (Smith *et al.*, 2000). Homozygous deletion of *roX1* has no obvious phenotype in males or females (Meller *et al.*, 1997) suggesting that the two roX RNAs are redundant. In support of this hypothesis, Franke and Baker (1999) have shown that *rox1/rox2* double mutants have an altered MSL binding pattern.

There is at least one other example of an untranslated RNA being associated with a histone acetyltransferase in a multiprotein complex. The steroid receptor RNA activator (SRA) gene produces an untranslated RNA that is associated with the proteins steroid receptor coactivator (SRC1) and activation function 2 (AF2) forming the steroid receptor complex (Lanz *et al.*, 1999). SRA is a transcriptional coactivator for steroid hormone receptors and functions in the presence of the translation inhibitor cyclohexamide (Lanz *et al*, 1999) and it had previously been shown that SRC1 has HAT activity (Jenster *et al.*, 1997).

Targeting the X chromosome: Since the MSL proteins were shown to associate with the X chromosome in males, it has been of interest to determine the mechanism by which they target the X chromosome. There must be some feature (likely specified by sequence) of the X chromosome that makes it distinct from the autosomes and a target for MSL complex assembly or binding. The nature of such a feature has yet to be determined. In our laboratory, Weigang Gu and colleagues (1998) have addressed the problem of MSL complex assembly and have produced experimental evidence to support a model in which MSL1 and MSL2 bind the chromosome initially followed then by MLE and later MOF and MSL3. Additional work suggests that the roX RNAs' incorporation into the complex requires the presence of MLE (Meller *et al.*, 2000; Richter *et al.*, 1996). Recently, Gu (pers. comm.) has shown that the enzymatic activities of MOF and MLE are necessary for the

"spreading" of the MSL complex to the hundreds of sites it normally occupies on the X chromosome.

Work in our lab and by others is focusing on the identification of additional components of the MSL complex. Recently, Jin and colleagues (1999) have identified a kinase, Jil1, that is present throughout the chromosomes in both sexes but is enriched to a two-fold level on the X chromosome of males. Interestingly, the Jil1 kinase is capable of autophosphorylation and of phosphorylation of histone H3 *in vitro*. The *in vivo* targets of the kinase are presently unknown. Work is in progress to determine if Jil1 kinase is associated with the MSL complex.

B. Chromatin and transcription

In spite of the analyses of the enzyme activities of at least two components of the MSL complex, we still do not understand how it accomplishes the two-fold increase in transcriptional activity. We do not know if or how the complex interacts with RNA polymerase or whether it acts at transcription initiation or during elongation. We also do not understand the role of H4AcK16 in the hypertranscription process. In addition to continuing the direct biochemical analysis of the MSL complex itself, some insights into its function may be gained by the study of homologous proteins and complexes responsible for chromatin modification and remodeling in Drosophila and other organisms.

If stretched out and laid end to end, the DNA that makes up the human genome would be approximately 200 cm long. By contrast, the diameter of the human nucleus, which houses the genetic material, is only 3-10 microns (Alberts *et al.*, 1994). Even in eukaryotic organisms with smaller genomes, orderly packaging of the DNA into the nucleus is a major molecular engineering challenge. This packing is accomplished with the use of small basic proteins known as histones. The two classes of histones are the H1 or linker histones and the nucleosomal histones H4, H3, H2A and H2B. These proteins, particularly histone H3 and histone H4 are among the most well conserved proteins known.

Two of each of the four nucleosomal histone proteins associate with each other in an octamer around which the DNA is wrapped to form a nucleosome. This structure is repeated along the length of the DNA molecule yielding a "beaded string" that is coiled to form the chromatin fiber. The compact chromatin fiber is highly repressive to transcription and must be modified or remodeled in order for RNA polymerase to gain physical access to the DNA template so that genes can be actively transcribed. Modifications may chemically alter the nucleosome such that the DNA is less tightly bound to it. Remodeling processes alter the placement of nucleosomes so that the transcriptional machinery can have access to the region of the DNA template necessary to activate transcription.

Recently a large number of transcriptional coactivator complexes have been identified and characterized as chromatin remodeling or chromatin

modification complexes. These complexes do not activate silent genes but instead increase the rate of transcription of previously activated genes by several fold. Most complexes operate on a subset of often functionally unrelated genes. These multiprotein complexes can be classed into two major types: those with ATPase activity and those with HAT or histone deacetylase activity. The first class uses the ATPase/helicase activities to displace particular nucleosomes that may block promoter sequences while the second class modifies nucleosomes by acetylating (or deacetylating) the histone tails (Mizzen and Allis, 1998; Cairns *et al.*, 1996).

Complexes that use ATP hydrolysis: The *S. cerevisiae* SWI/SNF complex was the first chromatin remodeling complex to be isolated. The eleven proteins that make up the 2 MDa complex affect the mating switch process and sucrose fermentation (see Winston and Carlson, 1992). The SWI2 protein component is a DNA-stimulated ATPase (Coté *et al.*, 1994) and has known homologues in Drosophila (e.g. MLE) and mammals.

SWI2 is also a component of the *S. cerevisiae* complex RSC, an essential chromatin remodeling complex (Cairns *et al.*, 1996). Various complexes from other organisms, including Drosophila NURF (Tsukiyama and Wu, 1995) ACF (Ito *et al.*, 1999), CHRAC (Varga-Weisz *et al.*, 1997) and human FACT (Orphanides *et al.*, 1998) also have ATP-dependent nucleosome disruptive properties. The mammalian SWI2 homologue is involved in a human SWI/SNF complex that uses ATP to convert the normal nucleosome

to a stable altered state and back again (Schnitzler *et al.*, 1998). All of these complexes have an ATPase/helicase as one of the core components of the machinery and some of them share some other components. Thus this type of complex can be assembled in more than one way, using some common and some unique components.

Complexes with HAT activity: The second class of coactivator complex are those that modify histones. There are many types of modifications that are possible, including acetylation, phosphorylation, methylation and ribosylation. Generally these modifications are thought to decrease the affinity of the histone for DNA by neutralization of positive charge (acetylation (Brownell and Allis, 1996)), increase of negative charge (phosphorylation) or by changes in nucleosome structure, perhaps by introducing steric hindrance. Several chromatin modification complexes that acetylate histones have been identified in various organisms including yeast SAGA, (Grant *et al.*, 1997) and ADA (Eberharter *et al.*, 1999), yeast NuA4 (Allard *et al.*, 1999) and NuA3 (Eberharter *et al.*, 1998), human TFTC (Brand *et al.*, 1999). In addition, proteins such as the CREB binding protein (CBP) and p300 (Ogryzko *et al.*, 1996; Bannister and Kouzarides, 1996), the nuclear hormone receptor ACTR (Chen *et al.*, 1997) the steroid receptor coactivator SRC1 (Spencer *et al.*, 1997) and BRCA 2 (Siddique *et al.*, 1998), have been shown to have HAT activity.

The yeast SAGA, ADA, NuA3 and NuA4 complexes target histones H3 and H4 *in vitro* (Allard *et al.*, 1999; Grant *et al.*, 1997; Eberharter *et al.*, 1998). The SAGA and ADA complex both have the HAT GCN5 as their catalytic subunit (Grant *et al.*, 1997) and acetylate nucelosomal histone H3 (Grant *et al.*, 1997). This histone is also the target of the NuA3 complex (Eberharter *et al.*, 1998) which has the HAT SAS3 as its catalytic subunit (John *et al.*, 2000).

Perhaps the most relevant of these yeast complexes for the research reported in this thesis is the *S. cerevisiae* complex NuA4. The NuA4 complex acetylates histone H4 when tested using oligonucleosomes as the substrate (Allard *et al.*, 1999). Esa1, the catalytic subunit of NuA4 (Allard *et al.*, 1999), is a MYST HAT and is closely related to Drosophila MOF. Recombinant ESA1 primarily targets histone H4 where it acetylates lysines 5, 8, 12 and 16 in *in vitro* assays using free histones (Smith *et al.*, 1998).

The MSL complex: The MSL complex is special among chromatin remodeling and modification complexes in at least two ways. First, it is the only known complex that exhibits both histone acetyltransferase activity (MOF) and helicase/ATPase activity (MLE). Thus, it has the potential to both modify and remodel chromatin. Second, the level of transcriptional modulation is only two-fold which is much lower than that produced by other known modification or remodeling complexes.

Recently Cho and colleagues (1998) isolated a human RNA polymerase II complex and showed that both HAT activity and ATPase activity contribute

to its chromatin modification and remodeling function. The HAT activity is ascribable to CBP/p300 and PCAF (p300/CBP associated factor) while the human SWI/SNF complex provides its ATPase activity. These workers showed that p300 is associated with the initiation-competent, nonphosphorylated form of polymerase while PCAF is associated with the elongation-competent, phosphorylated form of this enzyme. These observations may provide a model for understanding the Drosophila dosage compensation complex, which also has HAT and helicase activities.

C. Experimental rationale

Because so many of the proteins in the chromatin modification and remodeling complexes have been identified and cloned by homology clues, one might begin to find a pattern by which the proteins associate with each other. It might then be possible to develop a paradigm by which identification of one component would suggest its possible protein partners and infer the function of the complex to which they belong. The experiments discussed in this thesis are an attempt to begin realization of this goal. Extension of this work could not only lead to a better understanding of Drosophila dosage compensation but also of general mechanisms of eukaryotic transcriptional regulation.


♂
♀
w^a
w^a
w^a
w^a
w^a
w^a
w^a
w^a
w^a
expression level
1
2
3
4
WILD TYPE

Figure 1-1: Muller's experiment demonstrates dose dependent expression levels in males and females with one dose in males equal to two doses in females (Muller, 1932). See text for details.

CHAPTER 2

The human homologues

To date, no homologues of MSL1 or MSL2 have been identified. As previously mentioned RNA helicase A (RHA) is the human homologue of MLE (Lee and Hurwitz, 1993) and additional homologues have been identified in the mouse (Lee *et al.*, 1998b), cow (Zhang *et al.*, 1995) and in organisms as distantly related to flies as Arabidopsis (Wei *et al.*, 1997). The mammalian versions of RHA seem to have a more general role than that of MLE in Drosophila. RHA has been shown to mediate the association of RNA polymerase II with the CREB Binding Protein, CBP (Nakajima *et al.*, 1997) and with the breast cancer tumor suppressor protein, BRCA1 (Anderson *et al.*, 1998). The protein has also been shown to interact with the alpha chain of the interleukin-9 receptor (Sliva *et al.*, 1999). Additionally, RHA has been implicated in the post-transcriptional regulation of HIV-1 genes (Li *et al.*, 1999). Targeted disruption of RHA in mice results in early embryonic lethality showing that RHA is essential for normal gastrulation (Lee *et al.*, 1998a). It seems then that, unlike Drosophila MLE which is only required in males, mammalian RHA has a more general role in development and gene regulation. It should be noted that there has been no suggestion to date that RHA is involved in mammalian mechanisms of dosage compensation.

Prior to my work on this project, Antonio Pannuti in the Lucchesi lab, conducted a search of the human EST database and identified several ESTs (e.g. GenBank accession numbers: AA460000, H15179, and N95731) with 52% identity to Drosophila MOF. A human EST from an MSL3 related protein (hMRG15 - MORF4 related gene, chromosome 15, GenBank accession number AF100615) was also found. We hypothesized that these two proteins may exist in a multiprotein complex with chromatin modification and remodeling properties analogous to those of the Drosophila dosage compensation complex. Experimental support for this hypothesis would begin to describe another system of transcriptional regulation in mammals. The experiments reported in this chapter were designed to characterize human MOF and MRG15 and to investigate possible biochemical interactions between these proteins.

A. Characterization of MRG15

The MRG15 cDNA clone was obtained from Otsuka GEN Pharamaceutical Company (Tokushima, Japan) and completely sequenced. The predicted protein has a chromodomain and chromoshadow domain like its homologue MSL3 (Figure 2-1). A transcript of approximately 1.9 kb was detected by northern analysis of Raji cell poly A^+ RNA (Figure 2-2). This size was confirmed by others who also reported that GFP tagged MRG15 localizes to the nucleus when transiently transfected into HeLa cells (Bertram *et al.*,

1999). Antisera were raised against two synthetic peptides and the full-length protein with an N-terminal 10-histidine tag generated by cloning the cDNA into the pET19b vector (Figure 2-3). Sera raised against the full-length protein identified a protein of 40 kDa in western analysis (Figure 2-4).

B. Characterization of hMOF

Several hMOF cDNA clones were obtained from the IMAGE consortium and completely sequenced. The predicted protein has a zinc finger, a chromodomain and an acetyl CoA binding site and exhibits similarity to Drosophila MOF, its homologue (Figure 2-5). The longest clone contained 1583 nt but appeared to be missing the 5' end of the cDNA as northern analysis showed a transcript of approximately 1.8 kb (Figure 2-6). Additionally, the available sequence does not encode a start methionine. The available cDNA was cloned into the pET 19b vector (Figure 2-3) and expressed in bacteria producing recombinant protein with an N-terminal 10-histidine tag. This protein (hMOF C), along with two synthetic peptides, was used to generate several antisera. Several of the sera raised to hMOF C were able to identify a protein of approximately 50 kDa by western analysis (Figure 2-7).

I was able to map hMOF to the short arm of human chromosome 16 in region 11.2 using the following information: 1) some hMOF sequence is present in the 3' flanking region of the PRSS8 gene which was mapped to the

region by fluorescence *in situ* hybridization (Yu *et al.*, 1996); 2) a sequence tagged site (STS SHGC-15904) is contained within the known hMOF sequence.

C. Attempts to clone a full-length hMOF cDNA

I first screened (Sambrook *et al.*, 1989) a human heart cDNA library (nondirectionally cloned into Lambda ZAP II vector) and identified two positive clones, neither of which extended the previously known sequence. I then screened a brain cDNA library (directionally cloned into pCMV Sport) using a PCR-based method. Forward and reverse primers were designed to amplify a small region at the extreme known 5' end of the existing cDNA. The cDNA library to be screened was amplified, aliquoted and used as template for the PCR reactions. Aliquots producing a band of the expected size upon electrophoretic analysis were reamplified, divided into aliquots and assayed by PCR, using a set of primers internal to the initial set. This process was repeated three times at which point individual colonies were assayed by PCR. This screening process identified one positive clone that did not extend the known cDNA sequence.

Next, I tried a second PCR-based method to analyze brain, heart, spleen (directionally cloned into the UniZAP II XR vector) and thymocyte (nondirectionally cloned into Lambda ZAP II vector, kindly donated by Harish Joshi) cDNA libraries. In this procedure, I used a vector primer and a gene-specific primer to amplify product from a single aliquot of the cDNA

library. A small aliquot of the PCR reaction was then used with a vector primer and one of two nested gene-specific primers. These two PCR reactions were run in parallel and the products analyzed by electrophoresis. Bands which appeared in both reactions (suggesting a true and not false positive) were excised, the DNA extracted, subcloned into the pTA vector (Invitrogen TA cloning system) and sequenced. Despite controls and precautions taken, this method only generated false positives. Brief attempts were made to use inverse PCR and asymmetric PCR, but neither of these methods generated any additional hMOF sequence information.

I then focused on 5' rapid amplification of cDNA ends to try to isolate the remainder of the 5' end. I began by using the GibcoBRL protocol and materials with both total and poly A^+ RNA from Raji cells, a B cell lymphoma line known to express hMOF (Figure 2-6). I used a series of gene-specific primers with Superscript reverse transcriptase to generate the first strand cDNA. The Gibco BRL protocol then uses terminal deoxytransferase to produce a known 3' end on the cDNA. An anchor primer and nested gene-specific primer are subsequently used in a PCR reaction to amplify the product. I was unable to generate any specific product while using this protocol.

I considered the possibility that if the efficiency of the terminal deoxytransferase addition of the 3' sequences were low, then the remainder of the protocol would not be productive. That would explain the lack of any specific product from the GibcoBRL system. I then began using the Clontech

protocol and materials. The Clontech protocol uses a ligase-based reaction to add known sequence the 5' end of the cDNA. Initially, the "libraries" were generated using a modified oligo-dT primer with AMV reverse transcriptase as suggested by the manufacturer. Subsequent work using this protocol resulted only in false positives (among them β-catenin, ribosomal protein L3 and elongation factor K). I made several improvements to the protocol by using gene-specific primers (alone or in pairs) to prime the reverse transcription and by using Superscript II at 50 degrees to make the first strand cDNA. These changes eliminated the occurrence of false positives, but I still was unable to extend the hMOF sequence.

Aliquots of libraries previously made for use with the Clontech system were obtained from Elizabeth Stillwell (Harish Joshi lab). These libraries used Superscript II at 50 degrees with the modified oligo-dT primer, random primers, or random primers with a gene specific primer (not related to hMOF). PCR reactions with these libraries using hMOF specific primers (Table 2) and the provided reverse primer did not produce hMOF-related product.

Examination of the longest hMOF ESTs that existed at the time showed that they all terminated within 5 nucleotides of each other (Figure 2-8). This suggests that there is some sequence or structural feature that makes it difficult for the RNA-dependent DNA polymerase to proceed through the region. The sequence may encode a hard pause or result in a structural anomaly that makes the area refractory to *in vitro* reverse transcription.

Later analysis of the human EST database showed a sequencing error in the five-prime-most known region of the cDNA. The corrected sequence is given in the text of this document, though the Appendix A (Neal *et al.*, 2000) has the incorrect sequence. The correct sequence can be found using the GenBank accession number AF260665. The Human Genome Project has since sequenced through the 16p11.2 region. Using the genomic sequence, I have been able to predict the full-length protein sequence of hMOF.

D. hMOF is a histone acetyltransferase

The His-tagged recombinant protein was used in a histone acetyl transferase liquid assay. This C-terminal portion of hMOF (Figure 2-5) was shown to have HAT activity directed primarily toward histone H4 with some activity directed toward H3 and H2A (Figure 2-9). This specificity is similar to that of other MYST family HATs. The enzymatic characterization of hMOF is summarized in Neal *et al.*, 2000 (Appendix A).

E. Immunoprecipitation of hMOF and MRG15

To determine if MRG15 and hMOF interact, antibodies generated and described above were used in immunoprecipitation experiments with Raji cell or HeLa cell nuclear extracts (Santa Cruz). Despite being able to pull down hMRG15 and hMOF separately, western analysis of the immunoprecipitates

did not provide any evidence for association between hMRG15 and hMOF (Figure 2-10).

At this point in time, the EST database revealed another cDNA (hMSL3L1, GenBank accession number AC0004554) more similar to Drosophila MSL3 than is hMRG15. This human MSL3 homologue maps to Xp22.3 and is characterized in Prakash *et al.*, 1999. Given its higher degree of similarity to MSL3, it seems more likely that hMOF interacts with hMSL3 than with hMRG15.

In addition, several human MRG proteins were described by Bertram *et al.* (1999). These proteins are related to MORF4, a protein that maps to 4q1.2 and had been incorrectly thought to be a mortality factor (Bertram *et al.*, 1999). MRG15, which maps to 15q24, is the only one of the MORF4 related genes with a chromodomain. The others, MRGX (Xq22), MRG1 (1q4.1-2), MRG5 (5p14-15.1), MRG11 (11p telomere) and MORF4, have a region of homology that encompasses only the chromoshadow domain (Bertram *et al.*, 1999). Interestingly, the hMSL3L1 gene has a second translation start site that produces a protein that lacks the chromodomain (Prakash *et al.*, 1999) making it similar to the MRG proteins (MRG5, MRG11, MRGX, MORF 4) that lack the initial chromodomain.

These considerations provided a plausible explanation for the negative co-immunoprecipitation results. They also indicated that testing the hypothesis of specific association between MYST HATs and MRG15/MSL3 family members would be much more difficult that originally expected.

Given that a significant percentage of the Drosophila genome had been sequenced and that the EST database appeared to be very extensive, the likelihood that additional MOF or MSL3 homologues existed could be expected to be small. Therefore, it appeared more promising to test the hypothesis in Drosophila.

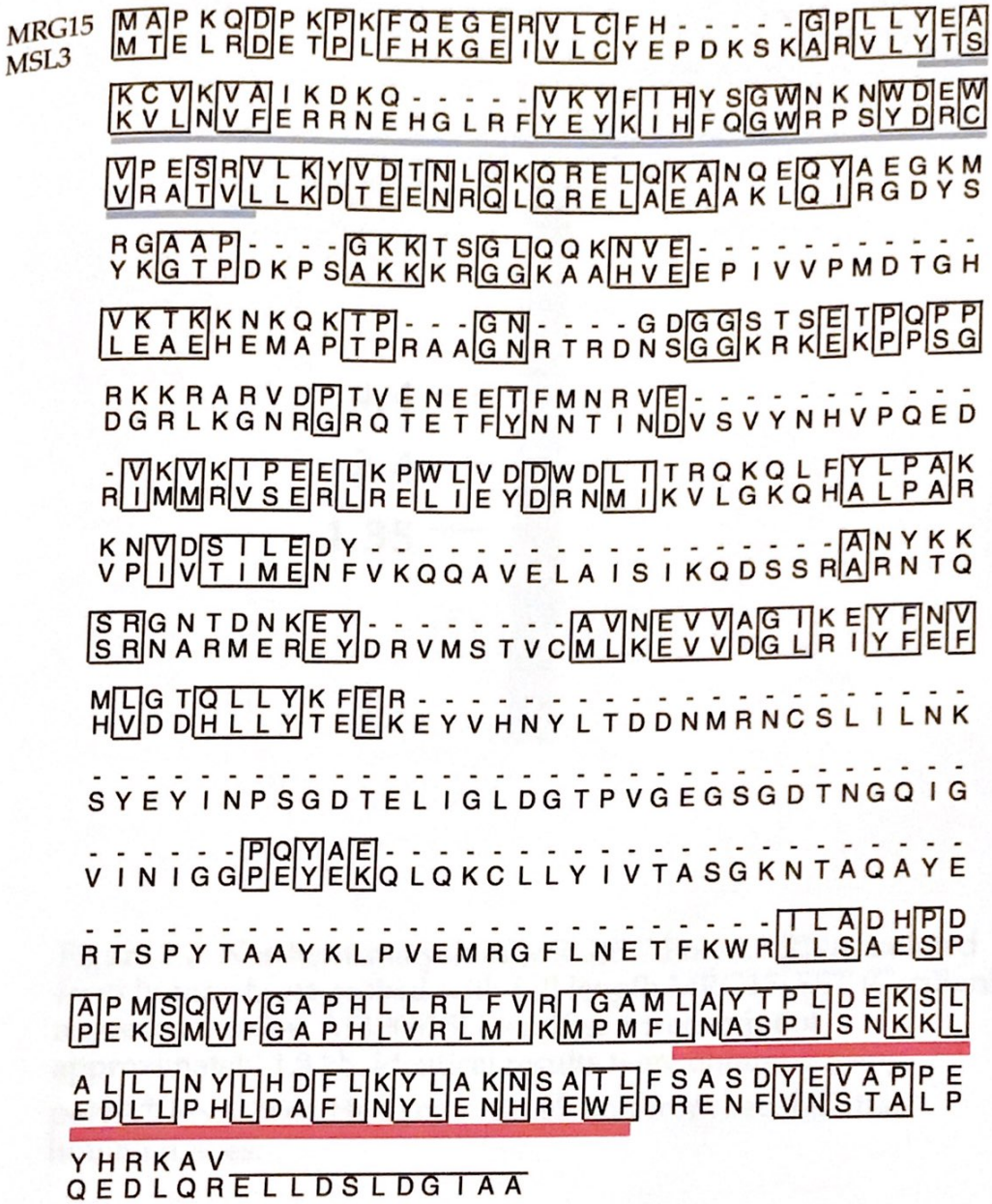


Figure 2-1: The predicted sequence of MRG15 as compared to Drosophila MSL3. Homologies as determined by BLOSUM 62 are boxed. The chromo domain is underlined in blue and the chromoshadow domain is underlined in red.

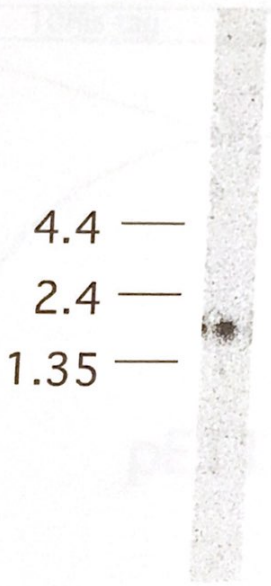


Figure 2-2: Northern analysis of MRG15. PolyA$^+$ RNA isolated from human heart probed with full length MRG15 EST (GenBank accession number AF100615) identifies a transcript of approximately 1.9 kb. Identical results were obtained using polyA$^+$ RNA from HeLa or RAji cell and from several other human tissues.

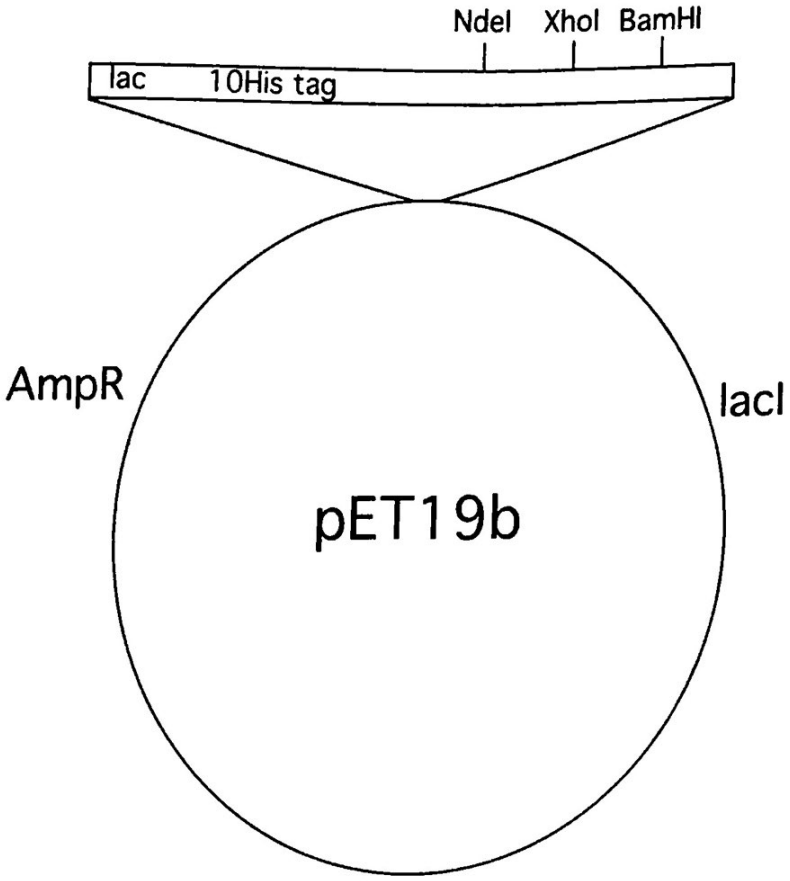


Figure 2-4: The pET 19b vector (Novagen). hMOFC and hMRG15 were subcloned into pET19b using the restriction sites shown.

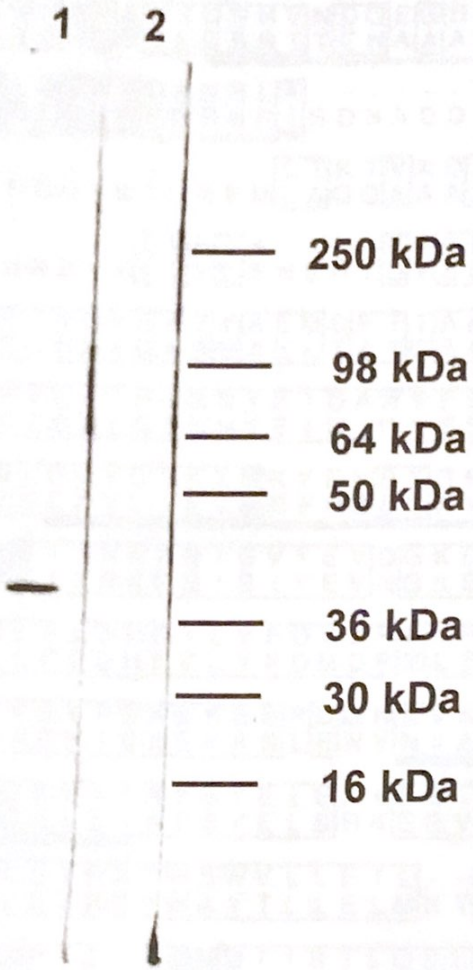


Figure 2-5: Western analysis of hMRG15. Rat polyclonal immune serum identifies a 40 kDa band (lane 1) that is not recognized by preimmune serum (lane 2). 30 ug of Raji cell nuclear extract was run in each lane.

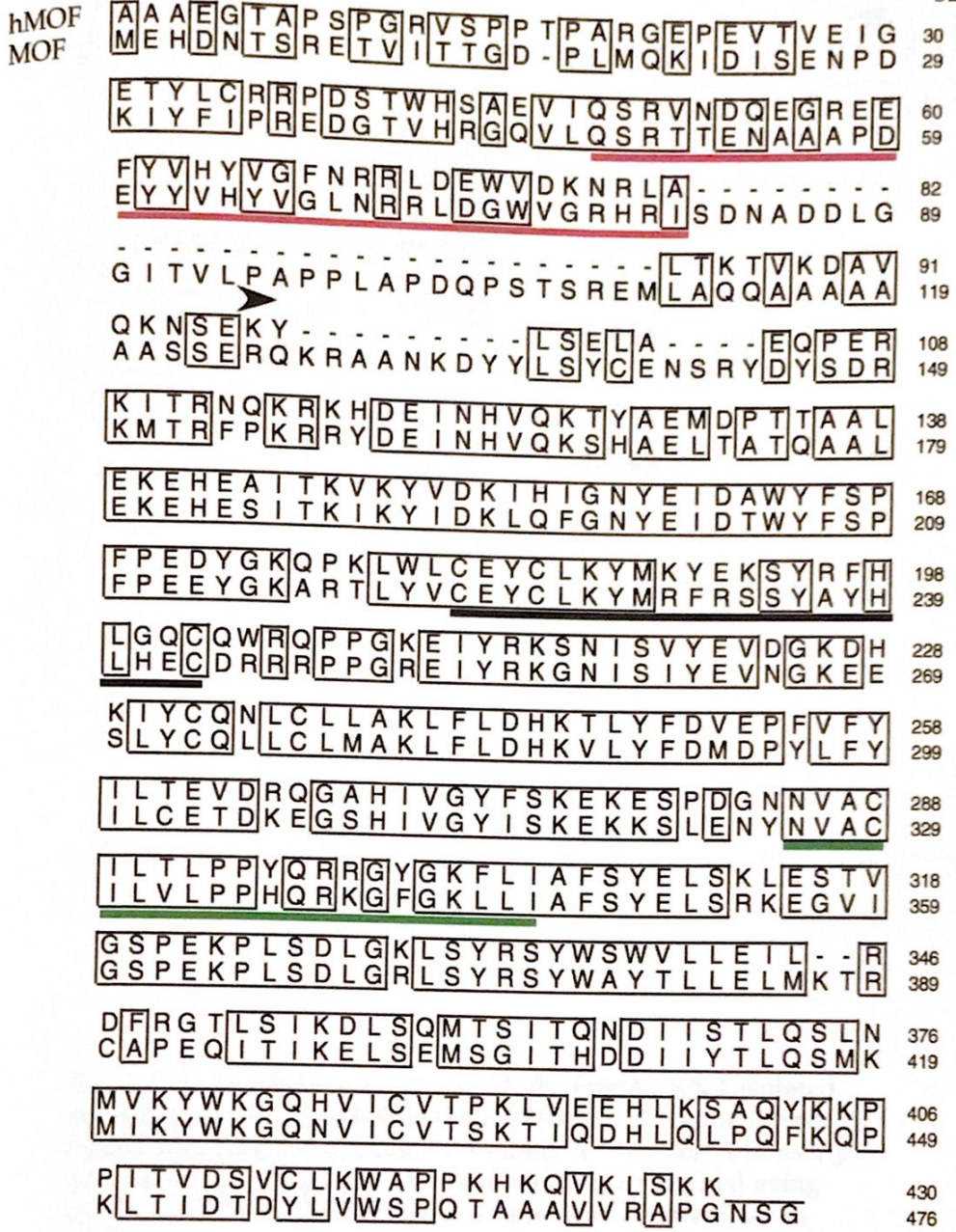


Figure 2-5: Predicted sequence of hMOF as compared to Drosophila MOF. Homologies as determined by BLOSUM 62 are boxed. The chromo domain is underlined in red, the zinc finger in black and the AcCoA binding site in green. The arrowhead indicates the initial amino acid of hMOFC.

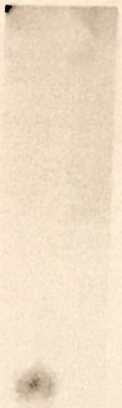


Figure 2-6: Northern analysis of hMOF. PolyA+ RNA isolated from Raji cells was probed witha a 470 bp AvaI fragment from an hMOF EST (GenBank accession number AA460000). A transcript of 1.8 kb is identified and identical results are obtained using polyA+ RNA from HeLa cells and from several other human tissues.

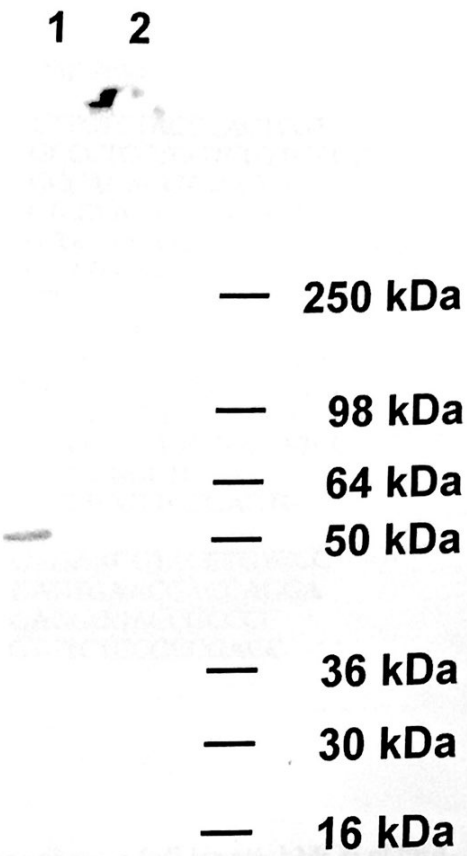


Figure 2-7: Western analysis of hMOF. Rabbit polyclonal immune serum identifies a 50 kDa band (lane 1) that is not identified by the preimmune serum (lane 2). 30µg of Raji cell nuclear extract was run in each lane.

name	direction	sequence
hMOFGSP1	reverse	CTTGTCTACCCACTCGT
hMOFGSP2	reverse	GCCCTCCTGGTCGTTGACTC
hMOFGSP3	reverse	GGCACAGGTACGTTCTC
HMO1	reverse	GTCGTTCACTCGAGACTGGATCACTTCAGC
HMO2	reverse	GTGCTTACCGGTCGCCGGCACAGGTACGTTTCT
HMO3	reverse	GAATGCCAGGTGCTATCCGGTC
HMO4	reverse	TTCTCTGAGTTCTTCTGTACAGCATCCTTCACTG
HMO5	reverse	GCTCAGGCTGCTCTGCGAG
hMOFPCR1	reverse	TCATGCTCCTTCTCCAAG
hMOFPCR2	reverse	GTCCATCTCTGCATAAGTCT
hMOFPCR3	reverse	CCGGCACAGGTAGCTTTCT
hMOFPCR4	reverse	GAATGCCAGGTGCTATCC
hMOFPCR5	reverse	GCTCAGGCTGCTC
hMOFPCR6	reverse	CTTTTGGTTGCGAGTG
hMOFGSP3r	direct	GAGAACGTACCTGTGCC
hMOFGSP2r	direct	GAGTGAACGACCAGGA
hMOF5	direct	GAGGGGACCGCCCC
hMOF6	direct	GTCTCTCCGCCGACC

Table 2: Primers used in attempts to clone a full length hMOF cDNA. See text for details of RACE analyses and other PCR based screening methods used.

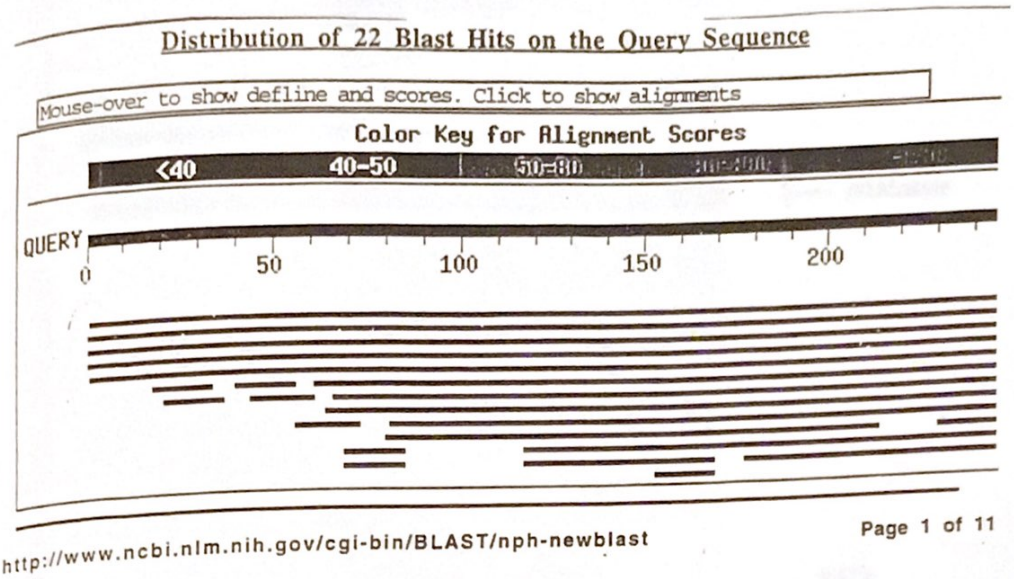


Figure 2-8: BLAST search using the available hMOF sequence showed that the five longest ESTs all terminated within 2-4 nucelotides of each other.

Figure 2-9: Histone acetyltransferase activity of hMOF. Lane 1: no protein extract. Lane 2: pET19b vector only. Lane 3: GCN5. Lane 4: ESA1. Lane 5: hMOFC. Panel A: Coomassie stained gel. Panel B: Fluorogram of labeled histones in Panel A gel.

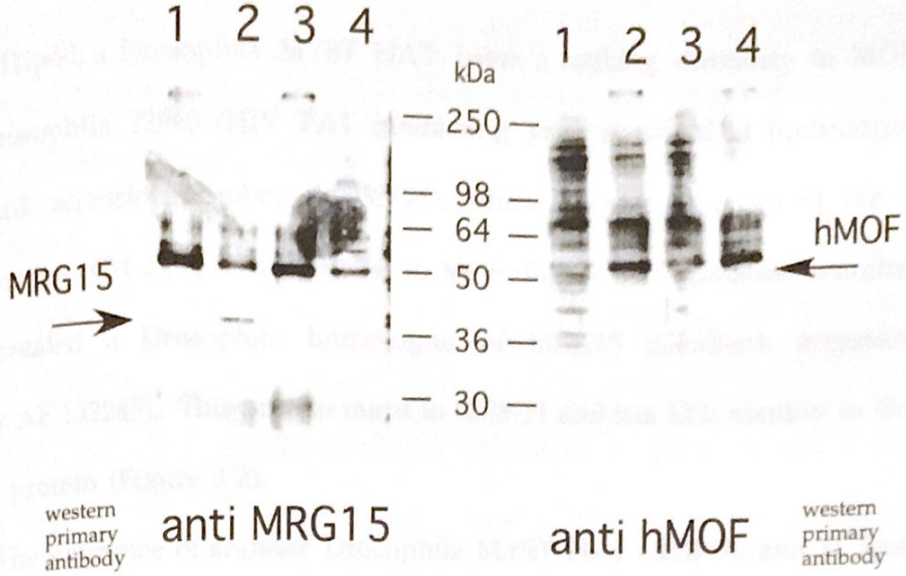


Figure 2-10: Immunoprecipitation of MRG15 and hMOF. Proteins were immunopreciptitated from Raji cell nuclear extract using the polyclonal rabbit antiserum against hMOF and the polyclonal rat antiserum against MRG15. Precipitated proteins were then analyzed by western blot for the presence of hMOF or MRG15 using the same sera. Identical results were obtained using HeLa cell nuclear extract. Lane 1: MRG15 preimmune IP. Lane 2: MRG15 immune IP. Lane 3: hMOF preimmune IP. Lane 4: hMOF immune IP. Left panel: anti-MRG15 western blot. Right panel: anti-hMOF western blot.

CHAPTER 3

The Drosophila homologues

dTip60, a Drosophila MYST HAT, bears a striking similarity to MOF. The Drosophila TIP60 (HIV *TAT interacting protein, 60* kDa) homologue (GenBank accession number AL033125), which maps to the tip of the X chromosome (4B1-2), is 53% identical to MOF (Figure 3-1). Database searches also revealed a Drosophila homologue of MRG15 (GenBank accession number AF 152245). This protein maps to 88E9-11 and has 42% identity to the human protein (Figure 3-2).

The existence of another Drosophila MYST HAT (dTIP60) and at least one other MSL3 homologue (dMRG15) suggested the existence of a second MSL-like complex in Drosophila. Given the observation that the yeast homologues ESA1 (yMOF) and yMSL3 interact via their chromodomains (Coté, pers.comm.), it is reasonable to propose that MYST HATs have specific MRG/MSL3-like partners with which they interact (Neal *et al.*, 2000), and more specifically that dTIP60 interacts with dMRG15. The experiments described in this chapter were designed to test this hypothesis.

A. Establishing the cell lines that express tagged proteins

In order to assay for an interaction between the proteins, I generated two stable Schneider 2 Drosophila cell lines (MOF-HA/MSL3-Flag and dTIP-

HA/dMRG15-Flag) each transfected with two constructs designed to produce differently tagged proteins upon copper sulfate induction of a metallothionien promoter. The MOF and dTIP60 cDNAs were cloned into the pMTHA vector (Figure 3-3) while the pMK33cFLAG vector (Figure 3-4) was used with MSL3 and dMRG15. These vectors were derived from the pMK33/pMtHy vector by Weigang Gu and Antonio Pannuti in our lab. The expressed proteins were tagged at their C termini. Commercially available antisera anti-HA and anti-Flag were used for immunoprecipitation and western analysis.

The plasmids were transfected into Schneider 2 cells, a Drosophila cell line shown to be of male origin, based on the lack of expression of *Sxl*, the presence of the histone H4 isoform acetylated at lysine 16, and the binding of the MSL proteins to the X chromosome (Gu and Pannuti, pers. comm.). After transfection, stable lines with integrated exogenous DNA were selected using hygromycin.

Having shown by western analysis that both proteins were expressed in each cell line (Figure 3-5), I next used immunofluorescence to establish that the vast majority, if not all, of the cells expressed each of the transfected cDNAs (Figure 3-6a, b).

B. Intracellular localization of dTIP60 and dMRG15

Cells were treated with copper sulfate to induce the metallothionein promoter in the absence of the selective agent hygromycin. Stable lines remain stable for many generations in the absence of hygromycin and hygromycin inhibits induction and leads to greater variability in expression. Induction with 50 uM $CuSO_4$ of cloned lines expressing MOF-HA only (produced by others in the lab) showed a localization pattern that mimicked the wild type pattern. Therefore, I proceeded on the assumption that any reproducible patterns observed at low levels of induction of dTIP60 and dMRG15 would represent the normal pattern of intracellular localization of these proteins in wild-type cells.

MOF-HA/MSL3-Flag cells treated with 50 µM $CuSO_4$ had tagged protein localized in their nuclei; the two proteins colocalized on the putative X chromosome as expected (Figure 3-7). In the other cell line, dTIP60-HA and dMRG15 were generally dispersed throughout the nucleus and no further sublocalization could be determined. The pattern of dTIP60-HA localization was similar to that observed for GFP-tagged human TIP60 (Yamamoto and Horikoshi, 1997) while the pattern of dMRG15 localization was similar to that of GFP-tagged human MRG15 (Bertram *et al.*, 1999).

Analysis of several dividing cells showed no obvious association of dTIP60 or dMRG15 with mitotic chromosomes (Figures 3-8 and 3-9). Consistent with this observation, there are no published data that show chromosomal association of the human homologues of these proteins. In contrast, MOF-HA has been shown in this and other work to maintain its

association with the X chromosome throughout mitosis (Figure 3-10, Lavender *et al.*, 1994; Gu, pers.comm.). Likewise, MSL3 is also associated with the mitotic X chromosome in these cells (data not shown).

C. Immunoprecipitation analysis of dTIP60 and MRG15

Nuclear extracts from the MOF-HA/MSL3-Flag cell line were incubated with anti-Flag agarose (Sigma) and the precipitated complexes were analyzed by western blot. As shown in figure 3-7, MOF-HA and MSL3-Flag are colocalized as was expected (Smith *et al.*, 2000). Similar treatment of nuclear extracts from the dTIP60-HA/dMRG15-Flag cell line yielded dMRG-Flag protein in the immunoprecipitate but failed to identify any associated dTIP60-HA protein (Figure 3-11).

D. Preliminary identification of proteins associated with dMRG15

The absence of evidence to support an interaction between dMRG15 and dTIP60 did not exclude the possibility that these two proteins are present in separate multiprotein complexes. Nuclear extracts were prepared from cells metabolically labeled with ^{35}S – methionine. Proteins precipitated with anti-Flag antibody (M2 antibody sepharose, Sigma) were separated on a denaturing gel and analyzed by fluorography. The results show that dMRG15 is associated with distinct proteins of various sizes suggesting its presence in

multiprotein complex (Figure 3-12). This same immunoprecipitate did not have HAT activity in the liquid assay. Using 12CA5 anti-HA agarose (Roche), anti-HA sepharose (Santa Cruz) or HA.11b anti-HA agarose (Berkeley Antibody Company), I was unable to precipitate dTIP60-HA protein (Figure 3-13). These results suggest that there is no HAT associated with dMRG15 in this putative complex.

E. Search for other Drosophila homologues of MSL3 using sequence analysis

A BLAST search of the HTGS (high throughput genome sequence) database using the full sequence of MSL3 or dMRG15 returns a long list of sequences many of which share only short regions of similarity with a small region of the query protein. Additionally, use of the entire gene sequence does not give greater priority to any particular region of the protein (*e.g.* the chromodomain) that may be important for identifying possible homologues on the basis of sequence similarity only.

Various algorithms have been designed to collapse similarities between proteins into a single consensus sequence. One of these algorithms, COnsensus Biasing By Locally Embedding Residues (COBBLER) determines the consensus sequence by considering not only how often a particular amino acid residue appears in a particular location but also its general frequency in a particular protein context (Henikoff *et al.*, 1995). This in effect "weights" particular amino acids which occur in certain contexts. The COBBBLER

algorithm has been used to identify a chromodomain consensus sequence (EYMIKWKGWNEMHNTWEPEENL) which is published on the World Wide Web at the site www.block.fhcrc.org. This sequence is derived from HP1 and Pc-related proteins whose chromodomains are distantly related to the chromodomains of MSL3 and MRG. A BLAST search using this chromo COBBLER block fails to identify either dMRG15 or MSL3 (Figure 3-14). Additionally, there is no published COBBLER block for the chromoshadow domain.

In light of these considerations, I designed new COBBLER blocks using the chromo and chromoshadow domain sequences from the known MRG/MSL3 proteins. I chose 71 amino acid-long regions (Figure 3-15) which encompass either the chromo or the chromoshadow domain of 6 proteins: *S. pombe* MSL3 (GenBank accession number Z98977.4) *S. cerevisiae* MSL3 (GenBank accession number Z71255), *D. melanogaster* MSL3 (GenBank accession number X81321), *D. melanogaster* MRG15 (GenBank accession number AF152245), *H. sapiens* MSL3 (GenBank accession number AF117065) and *H. sapiens* MRG15 (GenBank accession number AF152245). These sequences were loaded into the COBBLER block program at http://dot.imgen.bcm.tmc.edu and consensus sequence COBBLER blocks were generated (Figure 3-15). When used to search the nr data base, inclusive of all non-redundant coding sequences in GenBank, both of the newly generated consensus sequences separately identified all six proteins (in addition to many other related sequences) used to generate the consensus. When the newly

generated COBBLER blocks were used to search the high throughput genome sequence (HTGS) database (restricted to Drosophila sequences), a number of sequence contigs were identified. A comparison of the lists generated from the chromo COBBLER block and the chromoshadow COBBLER block revealed only two contigs (GenBank accession numbers AC019950 ad AC018039) present on both lists (Figure3-16). Analysis of the sequence of these two contigs shows that one of them is MSL3 (AC018039) and the other is dMRG15 (AC019950). This suggests that at this time MSL3 and dMRG15 are the only two proteins of this type in the Drosophila database. Although this analysis was initiated before the Drosophila genome sequencing was completed, the same results are obtained by searching the entire Drosophila genome. These data are summarized in Table 3.

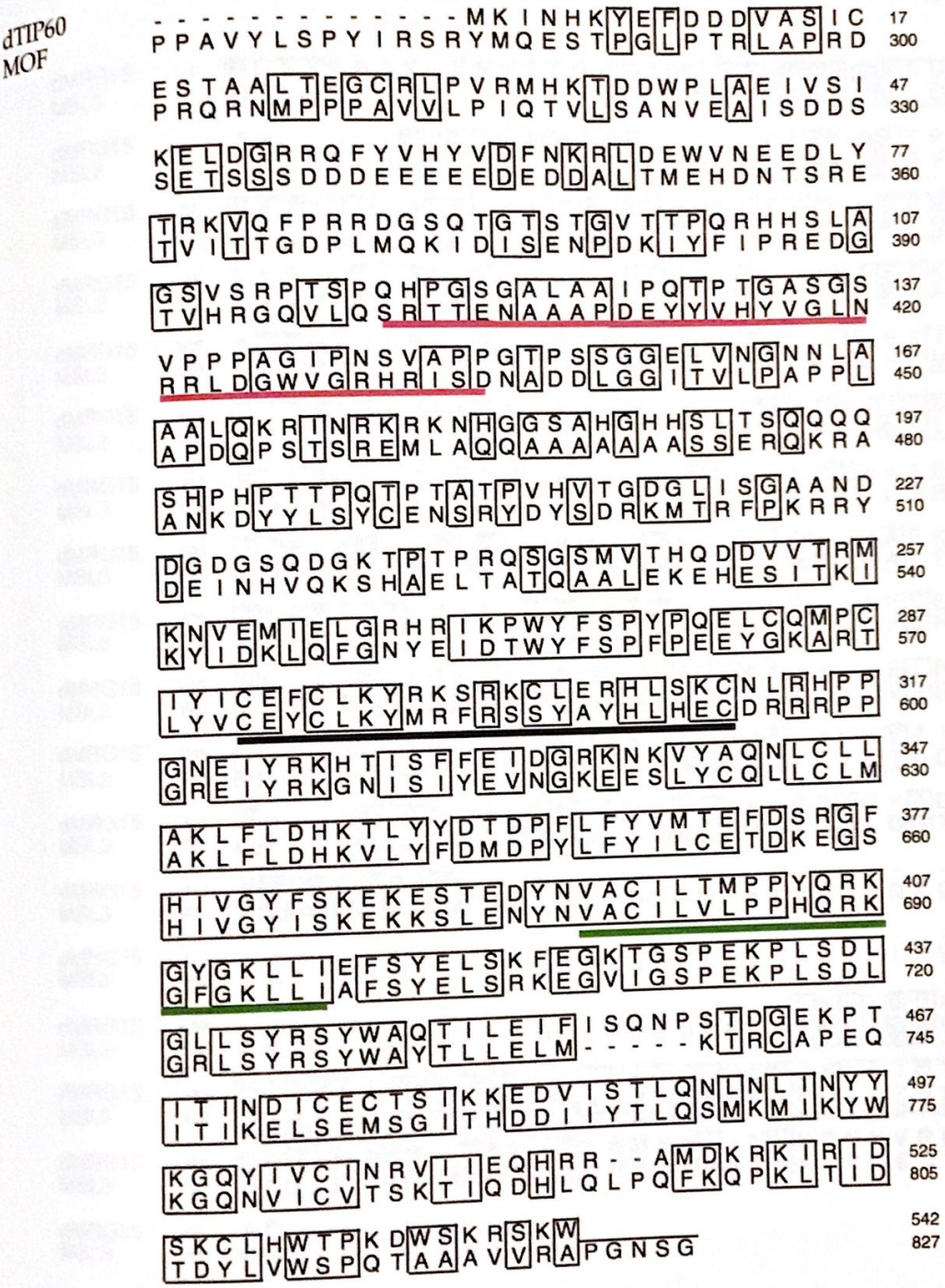


Figure 3-1: Predicted sequence of dTIP60 as compared to MOF. Homologies as predicted by BLOSUM 62 are boxed. The chromodomain is underlined in red, the zinc finger in black and the AcCoA binding site in green.

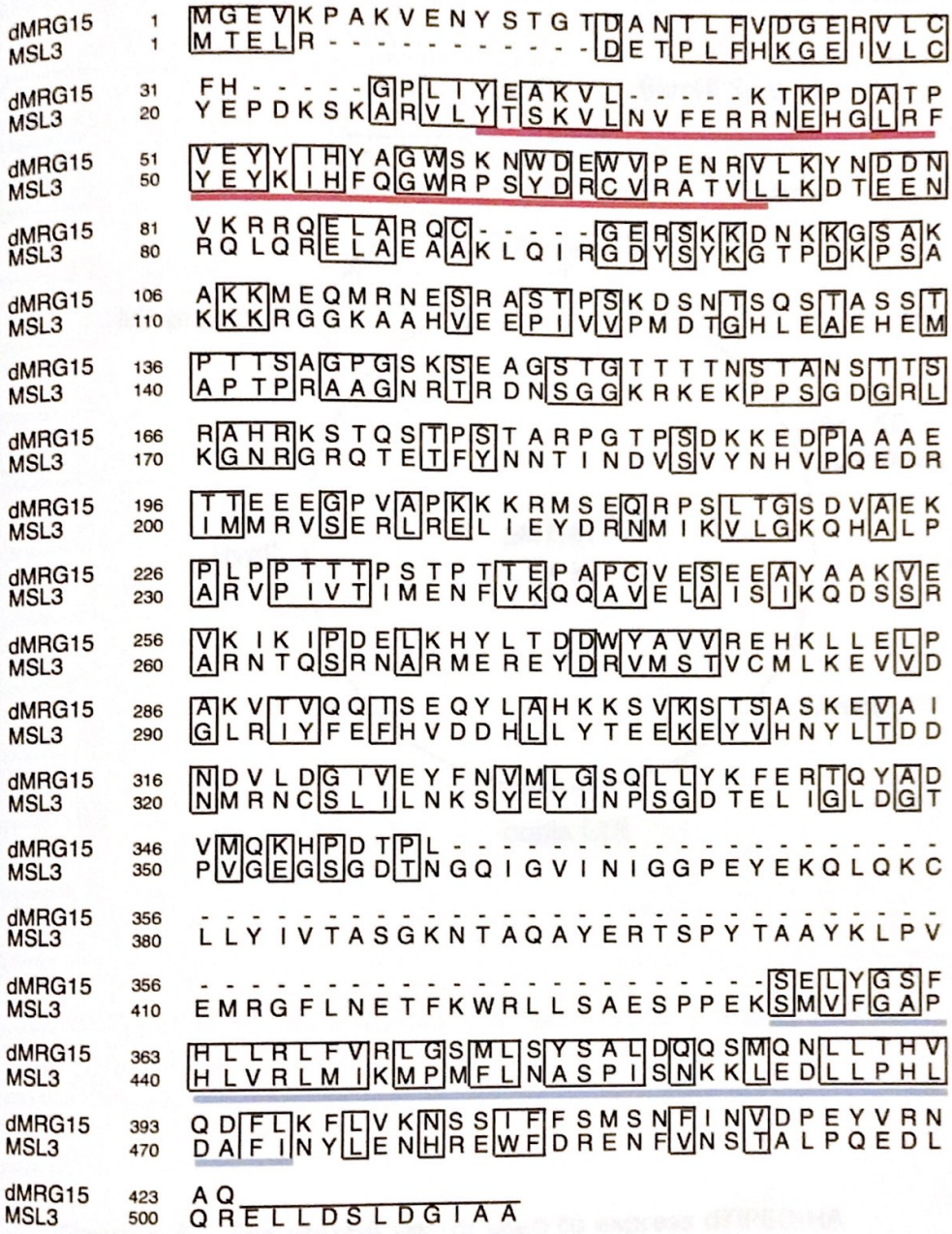


Figure 3-2: The predicted sequence of dMRG15 as compared to MSL3. Homologies determined by BLOSUM 62 are boxed. The chromodomains are underlined in red and the chromoshadow domains are underlined in blue.

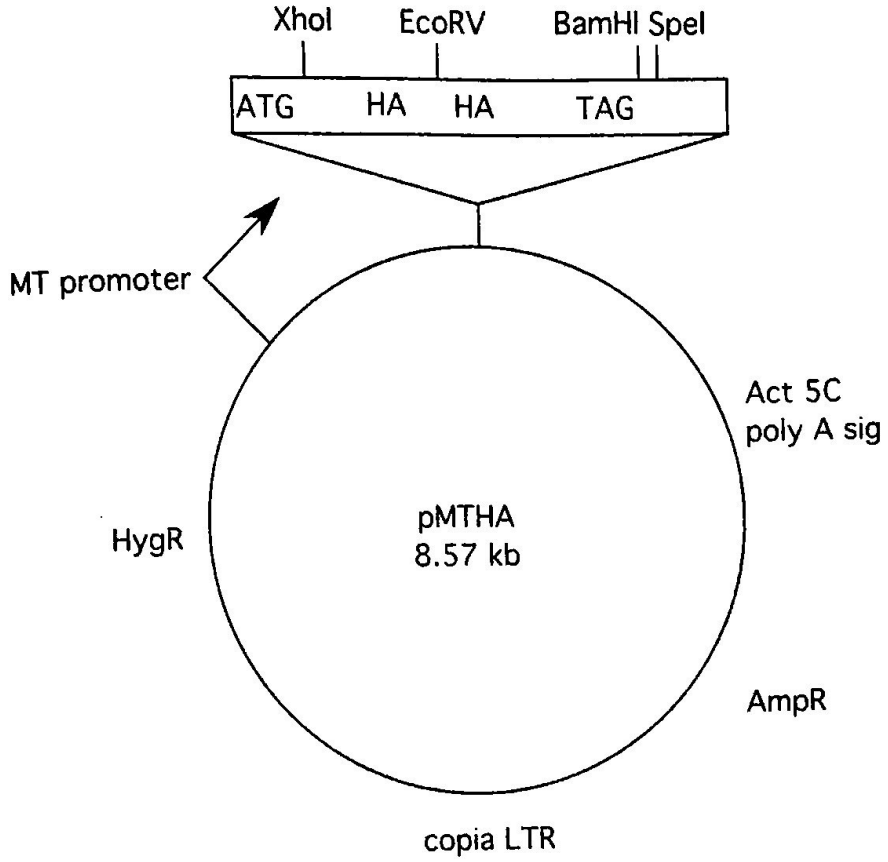


Figure 3-3: The pMTHA vector used to express dTIP60-HA and MOF-HA. Vector design by A. Pannuti and W. Gu.

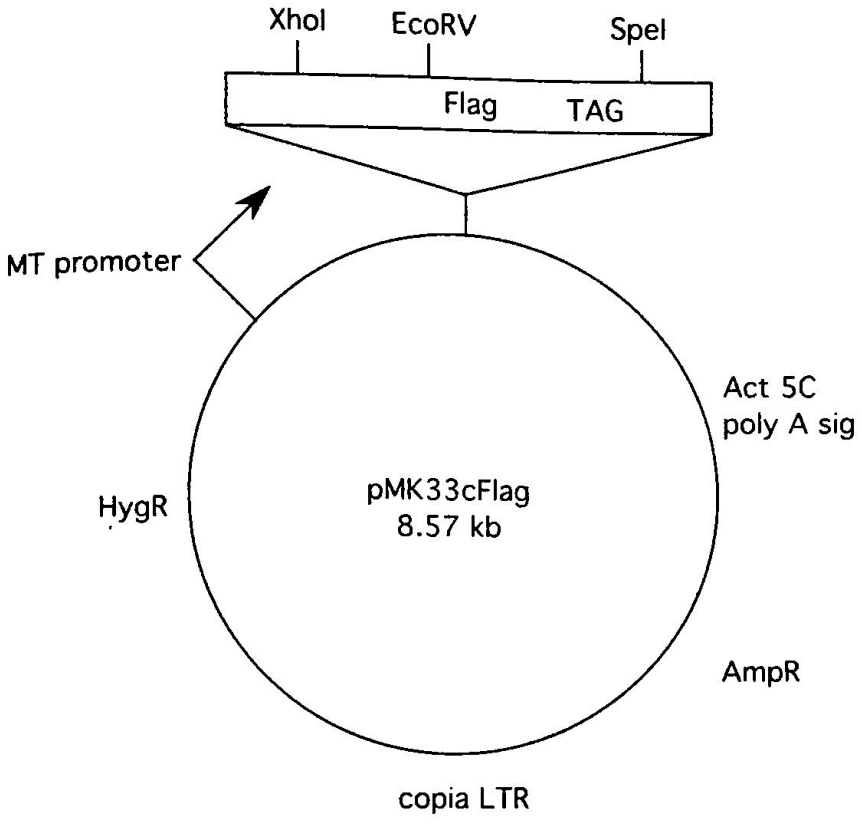


Figure 3-4: The pMK33cFLAG vector used to express MSL3-Flag and dMRG-Flag. Vector construction by A. Pannuti and W. Gu.

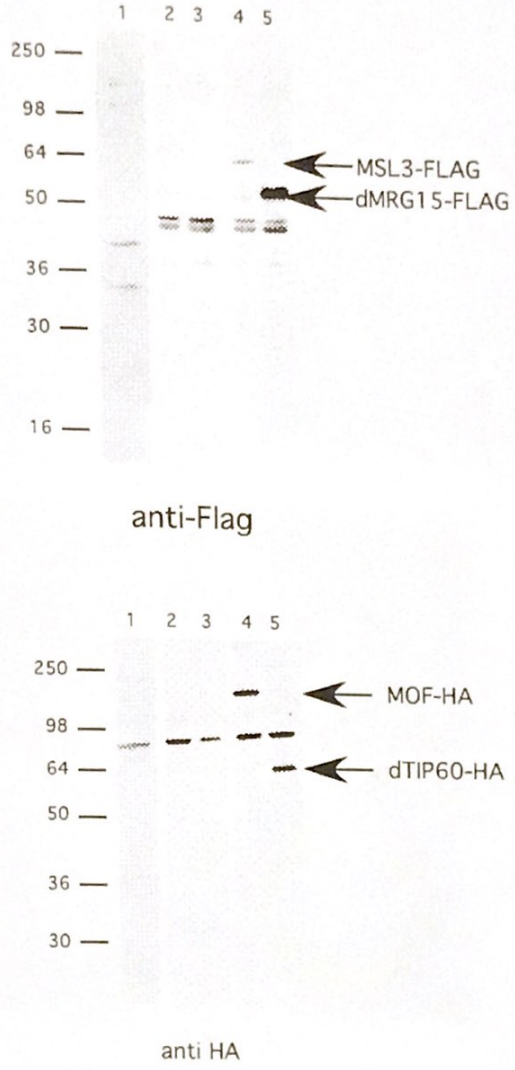


Figure 3-5: Western analysis of the stable cell lines used. Each lane contains 35mg of nuclear extract from the transfected cell lines indicated. Lane 1: untransfected Schneider 2 cells. Lane 2: uninduced MOF-HA/MSL3-Flag cells. Lane 3: uninduced dTIP60-HA/dMRG15-Flag cells Lane 4: induced MOF-HA/MSL3-Flag cells (200µM CuSO4). Lane 5: induced dTIP60-HA/dMRG15-Flag cells (200 µM CuSO4). Upper panel: Anti-Flag western blot with mouse monoclonal anti-Flag M2 ant.ibodies (Sigma). Lower panel: Anti-HA western blot with mouse monoclonal anti-HA 12CA5 antibodies (Roche).

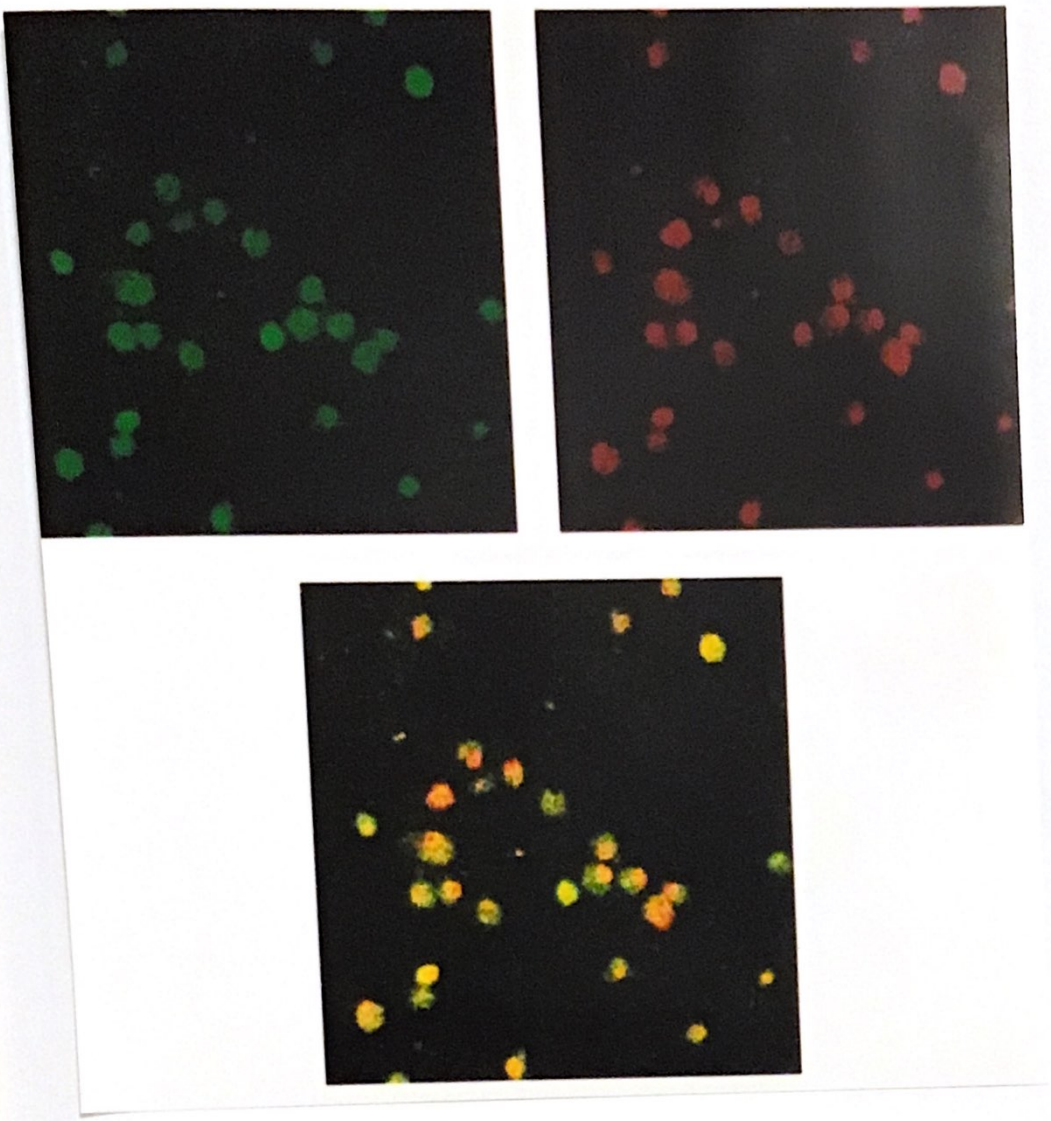


Figure 3-6a: MOF-HA/MSL3-Flag cells express both tagged proteins. Rabbit anti-HA monoclonal antibodies (Santa Cruz) were used to localize MOF-HA while mouse anti-Flag M2 monoclonal antibodies (Sigma) were used to localize MSL3-Flag. Secondary antibodies (Jackson Labs) labeled with either Cy5 (red-MOF-HA) or FITC (green-MSL3-Flag) were used to visualize the proteins using confocal microscopy. The merged image (yellow) is showed in the lower a panel.

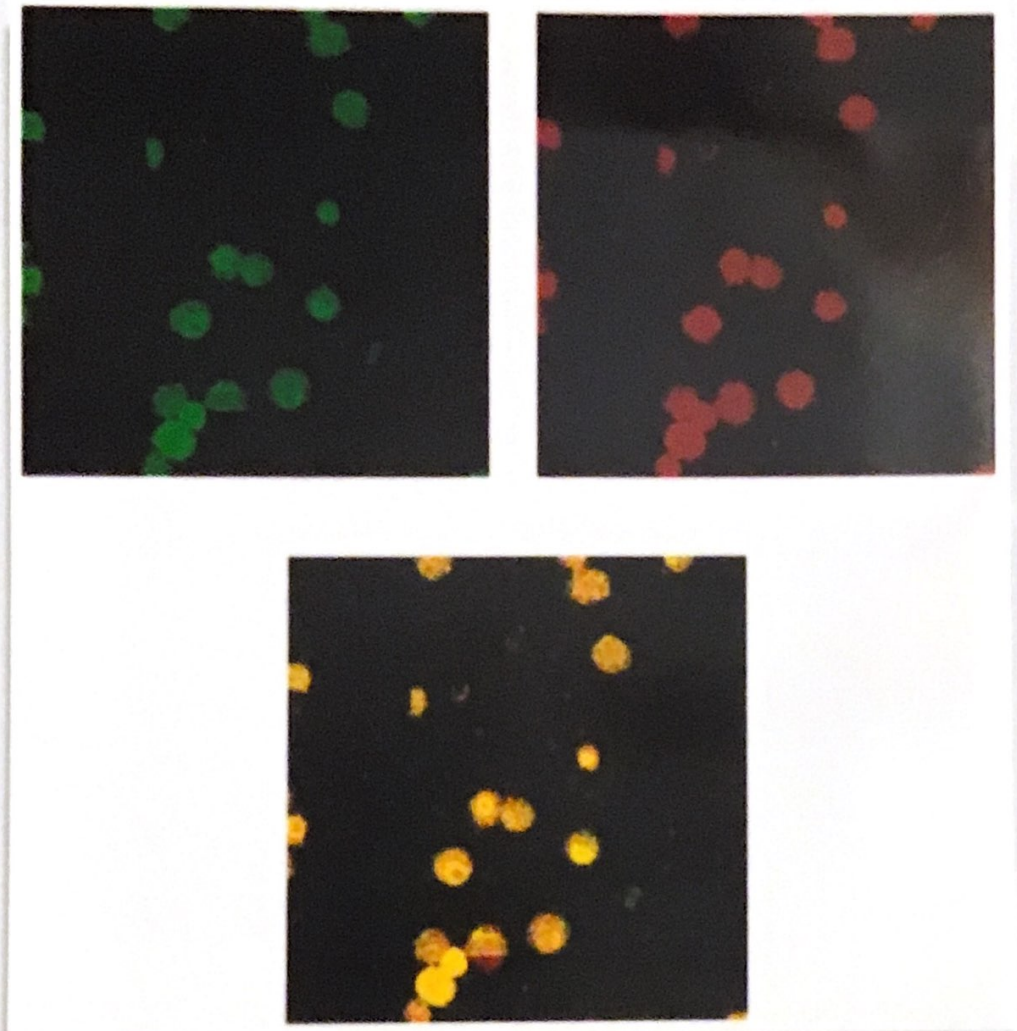


Figure 3-6b: dTIP60-HA/dMRG15-Flag cells express both tagged proteins. Rabbit anti-HA monoclonal antibodies (Santa Cruz) were used to localize dTIP60-HA while mouse anti-Flag M2 monoclonal antibodies (Sigma) were used to localize dMRG15-Flag. Secondary antibodies (Jackson Labs) labeled with either Cy5 (red, dTIP60-HA) or FITC (green, dMRG15-Flag) were used to visualize the proteins using confocal microscopy. The merged image (yellow) is showed in the lower a panel.

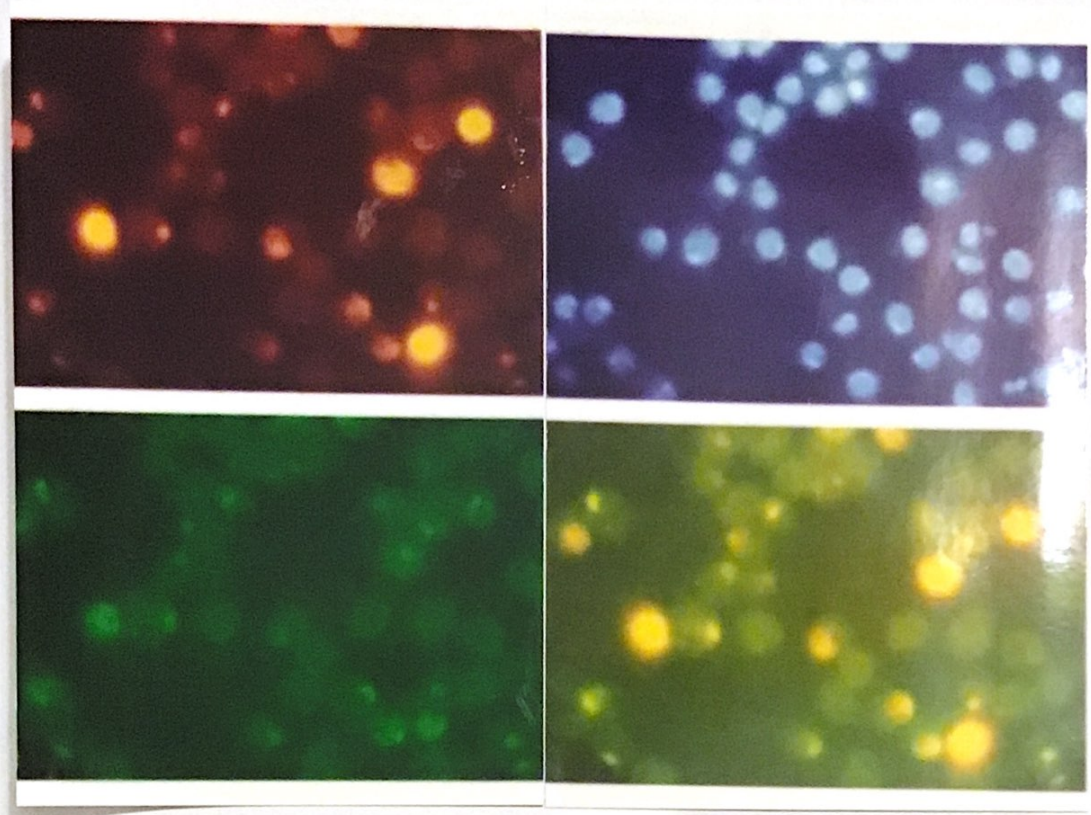


Figure 3-7: MOF-HA and MSL3-Flag colocalize. Rabbit anti-HA monoclonal antibodies (Santa Cruz) were used to localize MOF-HA while mouse anti-Flag M2 monoclonal antibodies (Sigma) were used to localize MSL3-Flag. Secondary antibodies (Jackson Labs) labeled with either Cy3 (red-MOF-HA) or FITC (green-MSL3-Flag) were used to visualize the proteins using epifluorescence microscopy. DNA is labeled with DAPI (blue). CY3 and FITC are merged in the lower right panel.

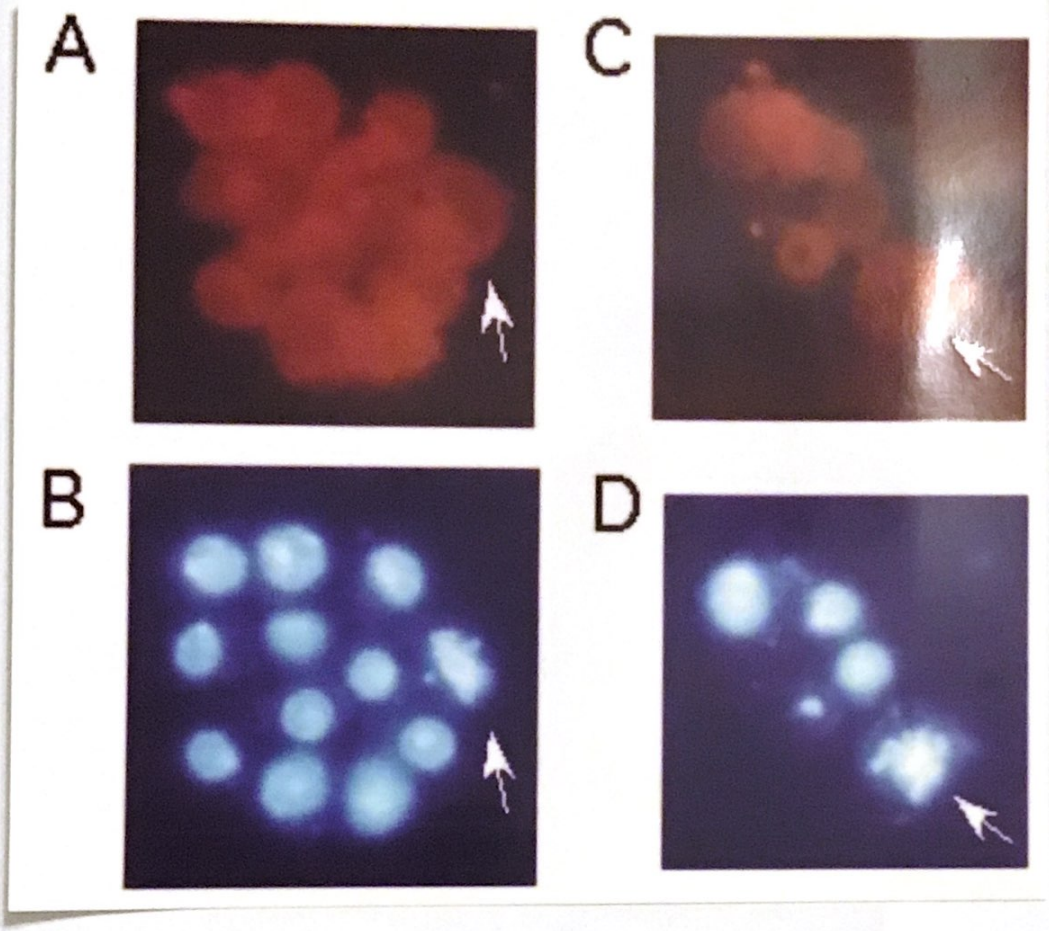


Figure 3-8: dTIP60-HA is not associated with mitotic chromosomes. Rabbit anti-HA monoclonal antibodies (Santa Cruz) were used to localize dTIP60-HA. Secondary antibodies (Jackson Labs) labeled with Cy5 (red, dTIP60-HA, A and C) were used to visualize the protein using epifluorescence microscopy. DNA is labeled with DAPI (blue,B and D). Arrows indicate mitotic cells.

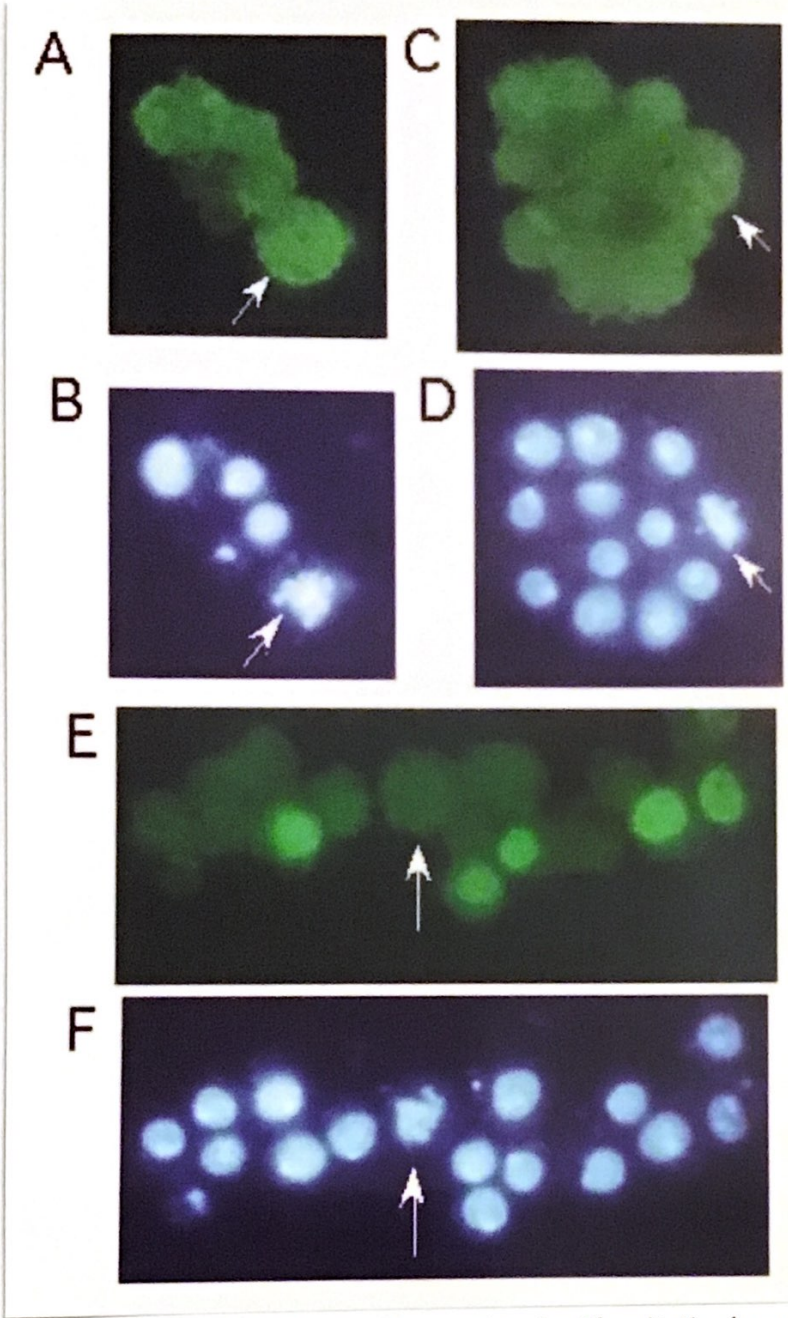


Figure 3-9: dMRG15-Flag is not associated with mitotic chromosomes. Mouse anti-Flag M2 monoclonal antibodies (Sigma) were used to localize dMRG15-Flag. Secondary antibodies (Jackson Labs) labeled with FITC (red, dMRG15-Flag, A, C and E) were used to visualize the protein using epifluorescence microscopy. DNA is labeled with DAPI (blue, B, D and F). Arrows indicate mitotic cells.

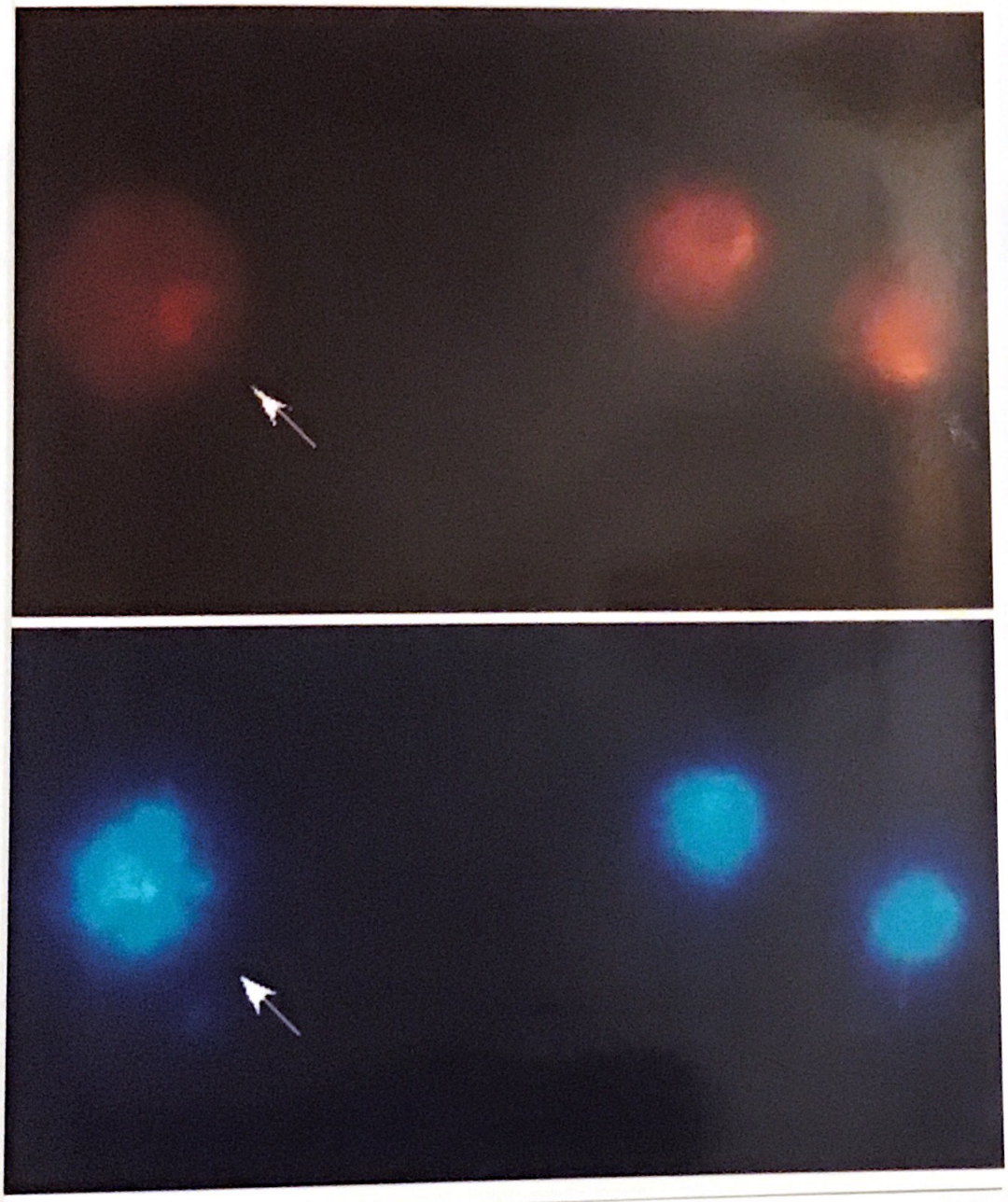


Figure 3-10: MOF-HA is associated with mitotic chromosomes. Rabbit anti-HA monoclonal antibodies (Santa Cruz) were used to localize MOF-HA. Secondary antibodies (Jackson Labs) labeled with Cy5 (red, dTIP60-HA) were used to visualize the protein using epifluorescence microscopy. DNA is labeled with DAPI (blue). Arrows indicate the mitotic cell. MOF-HA staining is prevalent in a specific region of the premetaphase cell, presumably the X chromosome.

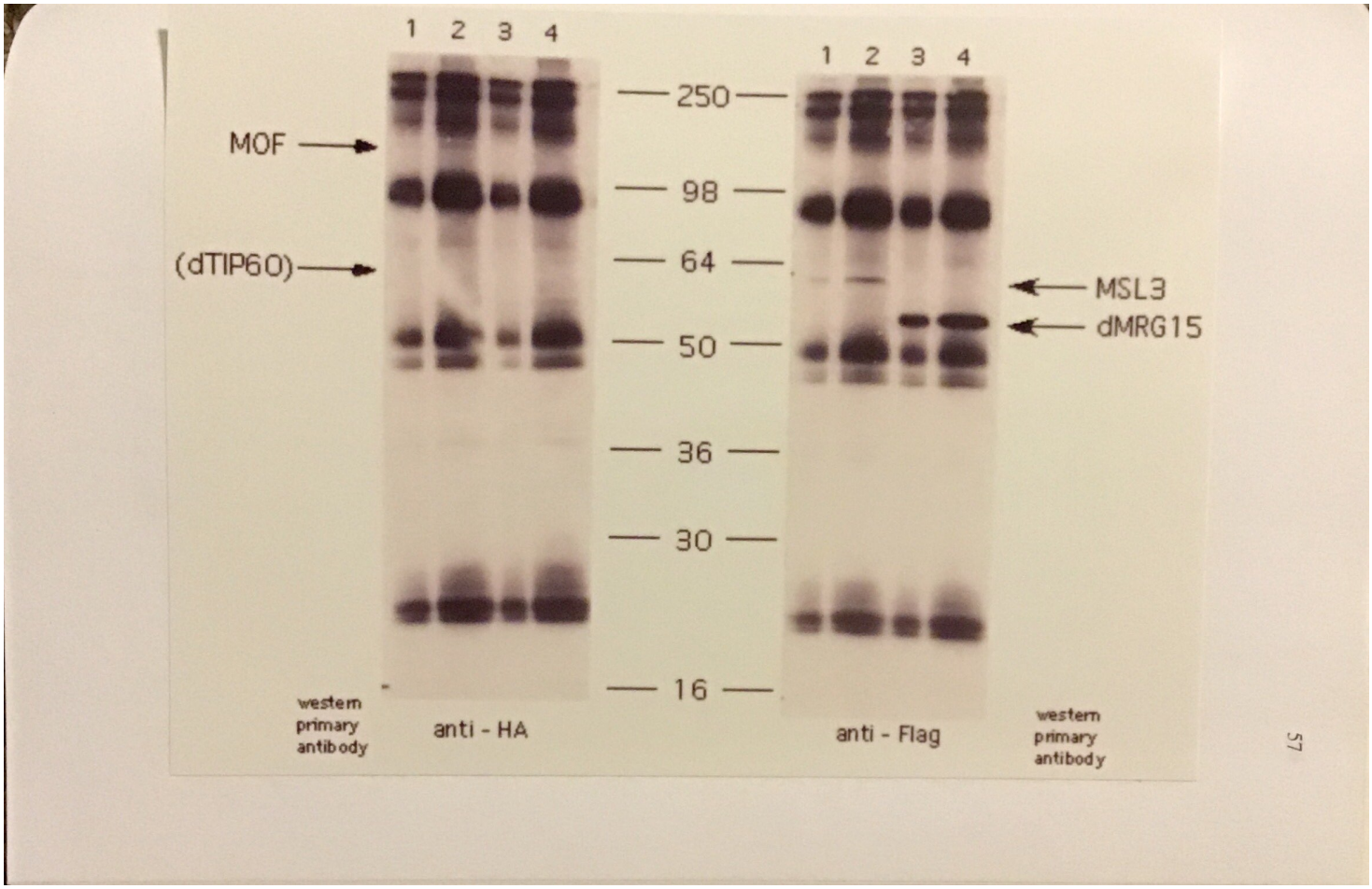

1 2 3 4
MOF
(dTIP60)
250
98
64
50
36
30
16
western primary antibody
anti - HA
1 2 3 4
anti - Flag
MSL3
dMRG15
western primary antibody

Figure 3-11: Immunoprecipitation of the Drosophila proteins. Proteins were immunoprecipitated from nuclear extract from induced (200μM $CuSO_4$) stable cell lines indicated. MSL3-Flag and dMRG15-Flag and associated proteins were precipitated using the using mouse monoclonal anti-Flag M2 antibodies (Sigma). Anti-Flag IP of MOF-HA/ MSL3-Flag nuclear extract (lanes 1 and 2) and of dTIP60-HA/dMRG15-Flag (lanes 3 and 4). Lanes 1 and 3: 25 μg (from nuclear extract) Lanes 2 and 4: 75 μg (from nuclear extract. Right panel: Anti-Flag western blot with mouse monoclonal anti-Flag M2 antibodies (Sigma). Left panel: Anti-HA western blot with mouse monoclonal anti-HA 12CA5 antibodies (Roche). The similarity in background bands is attributed to the secondary antibody.

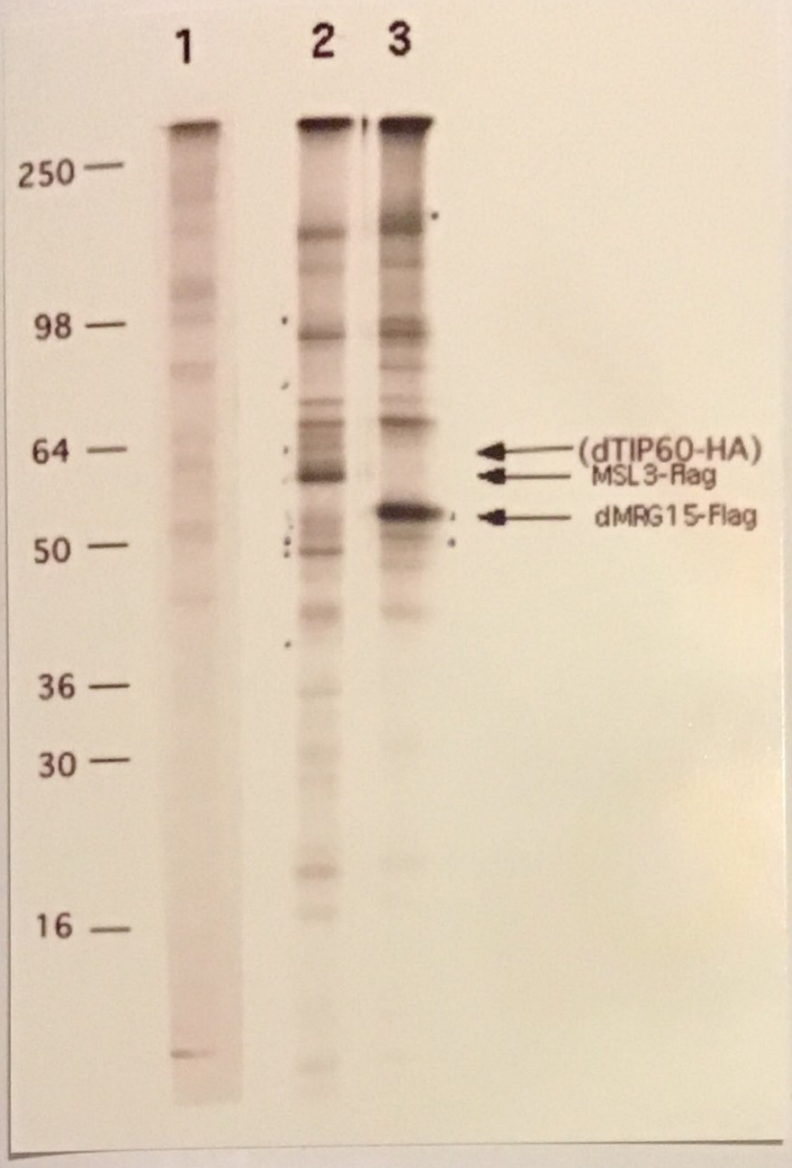


Figure 3-12: dMRG-Flag is associated with other proteins. Radiolabeled nuclera proteins wer uesd inand anti-Flag IP with mouse monoclonal anti-Flag M2 antibodies (Sigma). 10^8 cell equivalants used in each lane. Lane 1: untransfected Schneider 2 cells. Lane 2: MOF-HA/MSL3-Flag cells. Lane3: dTIP60-Ha/dMRG15 Flag cells.

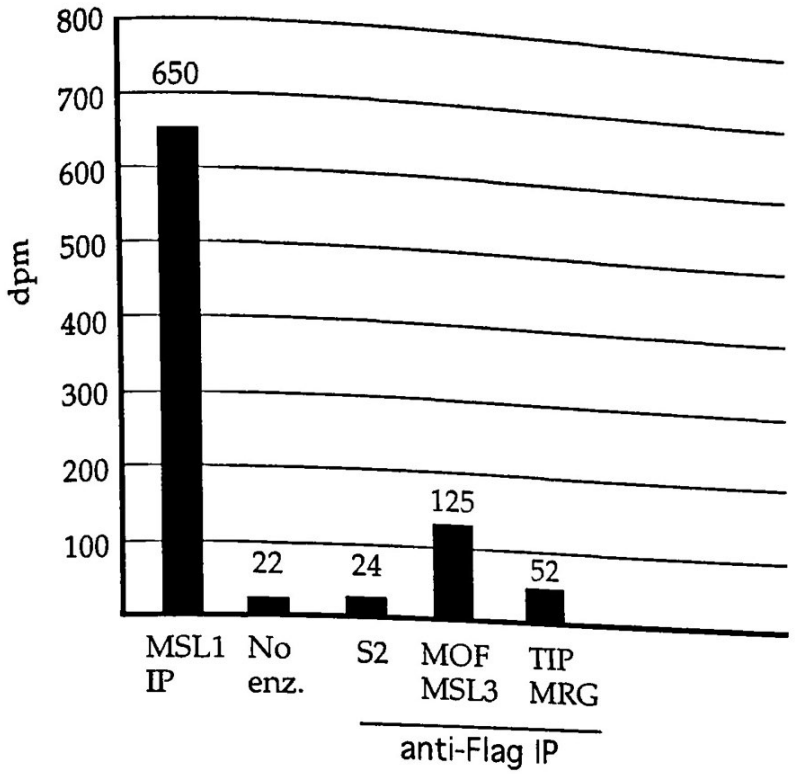


Figure 3-13: Histogram of HAT activity analysis of proteins associated with dMRG15-Flag, MSL3-Flag or MSL1. The MSL1 IP represents purified MSL complex with MOF HAT activity. HAT activity (MOF) is also found in the MSL3-Flag IP. The number of associated counts is much lower because MSL3-Flag is overexpressed and much of the free protein is likely not associated with MOF. There does not appear to be any HAT activity associated with dMRG15. MSL1 is precipitated with a rabbit polyclonal antibody to the protein. The other samples had proteins precipitated by mouse monoclonalanit-Flag M2 antibodies (Sigma).

Chromodomain

```
hMSL3     LYDAKIVVVIVGKDEKGRKIPEYLIHFNGWNRSWDRWAAEDHVLRDTDENRRLQRKLARKAVARLRSTGRK
MRG15     FHGPLLYEAKCVKVAIKDKQVKYFIHYSGWNKNWDEWVPESRVLKYVDTNLQKQRELQKANQEQYAEGKMR
MSL3      LYTSKVLNVFERRNEHGLRFYEYKIHFQGWRPSYDRCVRATVLLKDTEENRQLQRELAEAAKLQIRGDYSY
dMRG15    CFHGPLIYEAKVLKTKPDATPVEYYIHYAGWSKNWDEWVPENRVLKYNDDNVKRRQELARQCGERSKKDNK
ScMSL3    GGRCLAFHGPLMYEAKILKIWDPSSKMYTSIPNDKPGGSSQATKEIKPQKLGEDESIPEEIINGKCFIHYQ
SpMSL3    RVLCFHGPLLYEAKIVDTEMKGDVTTYLIHYKGWKNSWDEWVEQDRILQWTEENLKTQKELKNAAISTRQK
```

CHROMO COBBLER BLOCK

YLIHYKGWNPSWDEWVPEDRVLKWTDENAKKQRELK

Chromoshadow domain

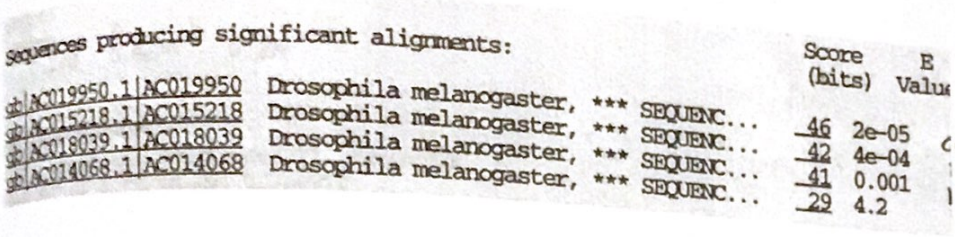
```
hMSL3     PGDQPPPPSYIYGAQHLLRLFVKLPEILGKMSFSEKNLKALLKHFDLFLRFLAEYHDDFFPESAYVAACEA
MRG15     PDAPMSQVYGAPHLLRLFVRIGAMLAYTPLDEKSLALLLNYLHDFLKYLAKNSATLFSASDYEVAPPEYHR
MSL3      PEKSMVFGAPHLVRLMIKMPMFLNASPISNKKLEDLLPHLDAFINYLENHREWFDRENFVNSTALPQEDLQ
dMRG15    HPDTPLSELYGSFHLLRLFVRLGSMLSYSALDQQSMQNLLTHVQDFLKFLVKNSSIFFSMSNFINVDPEYV
ScMSL3    LVPIRIYGAIHLLRLISVLPELISSTTMDLQSCQLLIKQTEDFLVWLLMHVDEYFNDKDPNRSDDALYVNT
SpMSL3    RQQYPDTEMCDLYGVEHLIRLFVSLPELIDRTNMDSQSIECLLNYIEEFLKYLVLHKDEYFIKEYQNAPPN
```

CHROMOSHADOW COBBLER BLOCK

SQIYGAPHLLRLFVKLPEMLSYTPMDEKSLECLLNHLDDFLKYLVKHKDEFF

Figure 3-14: Sequences used in the COBBLER analyses and the resulting COBBLER blocks.

A



B

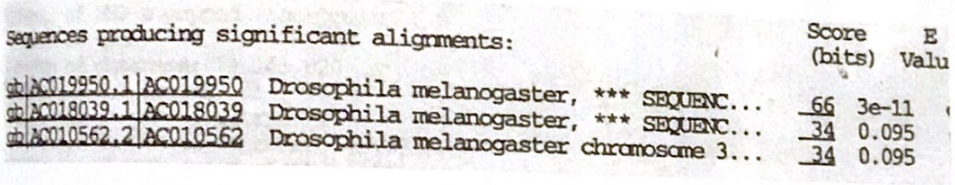


Figure 3-15 : There appear to be no additional MSL3/dMRG15 like proteins in the Drosophila genome. Panel A: the COBBLER chromo block used in BLAST analysis of the Drosophila high throughput genome sequence. Panel B: the COBBLER chromoshadow block used in BLAST analysis of the Drosophila high throughput genome sequence. See text for an explanation of the generation of the blocks and analysis of the results.

Summary of BLAST and COBBLER analyses

Query	Results
dMRG15	>100 proteins
MSL3	>100 proteins
published Chromo block	HP1, Pc and others (no MSL3, no dMRG15)
MSL3 chromo	MSL3 (no dMRG15)
MSL3 chromoshadow	MSL3 (no dMRG15)
dMRG15 chromo	dMRG15 and MSL3
dMRG15 chromoshadow	dMRG15 (no MSL3)
Chomo COBBLER block	dMRG15 and MSL3 (and 5 others)
Chromoshadow COBBLER block	dMRG15 and MSL3 (and 1 other)

Table 3: Summary of BLAST and COBBLER analyses. All searches used advanced BLAST default parameters, limited to the HTGS informationand limited to Drosophila. The first two sets used the protein database.

CHAPTER 4

Discussion

A. Conclusions

Two major conclusions can be drawn from the work described herein. First, given its HAT activity, hMOF has the potential to be involved in chromatin modification and perhaps chromatin remodeling. Second, the association between MYST HATs and MSL3/MRG15-like proteins is not random, that is, these proteins do not associate indiscriminately. The lack of coimmunoprecipitation evidence for an association between any of the studied proteins other than MOF with MSL3 suggests that there must be some factors or constraints that govern the pairing of these proteins.

Based on the known association of MOF with MSL3, I developed and tested the hypothesis that MYST HATs generally pair with MSL3/MRG15-like proteins. Given the level of the known human sequence information available at the time, it is, in retrospect, clear that the experimental approach chosen to test the hypothesis was too simplistic. Clearly, additional, as yet uncharacterized, human MSL3/MRG15-like proteins could exist and one of these may interact with hMOF. In contrast, the situation in Drosophila appears more amenable to experimental investigation. Using COBBLER blocks to search the entire Drosophila genome, I was unable to find any other proteins similar to MSL3 or dMRG15. Furthermore, MOF and MSL3 were

known to interact by their presence in a specific multiprotein complex. In spite of these promising parameters, I was unable to demonstrate an association of dTIP60 with dMRG15. This suggests that dTIP60 may interact with some other protein partner that has only one or no chromodomains and, therefore, is not a member of the MSL3/MRG15 family. It is also possible that its interaction with dMRG15 was not detected by my experimental approach.

B. Evolutionary considerations

It is perhaps not surprising that the proteins involved in chromatin remodeling and modification have been so well conserved during evolution. The problems of overcoming chromatin repression of transcription occur in all eukaryotic species, though these problems may become more severe with increased organismal complexity. This is evidenced by phylogenetic analysis of the MSL3/MRG15 family. There is only one member of the family present in the yeast, *S. cerevisiae*. The open reading frame YPR023C (yMSL3) encodes a protein that interacts with the essential MYST HAT ESA1 via their chromodomains (Eisen, Coté, Lucchesi, in prep). Though the specific genes that are under the control of the NuA4 complex have yet to be identified, it is likely that this complex is important for modulation of transcriptional rate of some, perhaps essential genes, in the *S. cerevisiae* genome. This is suggested by the observation that a kanamycin-mediated disruption of the ESA 1 gene leads to arrested growth (Smith *et al.*, 1998).

The yeast *S. pombe* and the nematode C. *elegans* also have only one known member of the MSL3/MRG15 family. In contrast, Drosophila have two members of the family, MSL3 and dMRG15. MSL3 may have arisen through a duplication of dMRG15 and been recruited to the male-specific dosage compensation machinery allowing dMRG15 to retain its function in both sexes. In this respect, it will be of interest to determine the mutant phenotype of dMRG15.

I should note here that there appear to be no homologues of either *msl1* or *msl2* in any other organisms. It could be that these two proteins serve as dosage compensation specific adapters for the more general subcomplex found in Drosophila males as well as female, yeast and mammals. Evidence to support this idea was provided by Gu and colleagues who have shown that, of the five known MSL proteins, MSL1 and MSL2 bind polytene chromosomes first and that neither will bind without the other (Gu *et al.*, 1998). In addition, Copps et al. (Copps *et al.*, 1998) have shown that the MSL1 and MSL2 proteins interact in vitro (each can be immunoprecipitated with antibody to the other) and in the yeast-two-hybrid system via the MSL2 RING finger.

The situation in mammals appears to be more complex. There are at least two MSL3/MRG15 genes in mammalian species, one of which (MSL3L1) is predicted to produce multiple splice variants. Analysis of the genomic sequence and several cDNAs predicts two protein products, one of which lacks an N-terminal chromodomain (Prakash *et al.*, 1999). Similar results

were found for the Drosophila homologue, though western analysis of MSL3 protein products suggests that only one protein, with both chromo and chromoshadow domains, is synthesized (Gorman *et al.*, 1995). The protein products of the MSL3L1 gene have not been analyzed and so it is not known if multiple protein species are present in vivo.

Because Prakash *et al.* (1999), who characterized the MRG proteins, did not recognize the presence of the chromoshadow domain, there has been no discussion of the significance of a chromoshadow domain in the absence of a chromodomain. Recent evidence has shown that chromoshadow domains form homodimers (Brasher *et al.*, 2000; Cowieson *et al.*, 2000). It is possible, therefore, that the chromoshadow domain can function in the absence of the chromodomain.

It is worth noting that the prediction of leucine zippers at the C-terminal end of MORF4, MRGX and MRG15 (Bertram *et al.*, 1999) and MSL3L1 (Prakash *et al.*, 1999) are likely to be in error. The leucine zipper motif requires a leucine at every seventh residue and a coiled-coil domain. Though the MRG proteins fulfill the first requirement, domain prediction programs (e.g. http://dot.imgen.bcm.tmc.edu:9331/seq-search/struc-predict.html and http://www.rockefeller.edu/rucs/toolkit/structure2.html) do not identify a coiled-coil domain for the MRG proteins.

There are multiple MYST HATs in *S. cerevisiae*, *C. elegans*, *D. melanogaster* and *H. sapiens*. Within this family, the MOF protein has been well conserved with homologues in yeast, worms, plants and humans (Neal

et al., 2000). Tip60 is also well conserved and a C-terminal region of human TIP60 can acetylate free histones in vitro (Yamamoto and Horikoshi, 1997). Later work showed that full-length Tip60 acetylates synthetic peptides from the amino-terminal ends of histones H3, H4 and H2A (Kimura and Horikoshi, 1998). The enzyme had preference for lys-5 of histone H2A, lys-14 of histone H3 and lys-5, 8, 12, and 16 of histone H4. Tip60 had been initially identified on the basis of interaction with the HIV TAT transactivator and *in vitro* assays suggested that the protein may be involved in the regulation of HIV genes (Kamine *et al.*, 1996). Later work showed that the HAT activity of Tip60 can be inhibited by interaction with HIV TAT (Creaven *et al.*, 1999). These workers also identify Mn-SOD as a gene with Tip60-dependent transcriptional activity and suggested that the TAT-mediated inhibition of Tip60 HAT activity decreases the expression of genes (such as Mn-SOD) that would normally interfere with viral propagation. Several groups have reported that the transcriptional activity of TAT is regulated by TAT acetylation, notably by the histone acetyltransferase p300 (Ott *et al.*, 1999; Kiernan *et al.*, 1999). Recently, Gavaravarapu and Kamine (2000) showed that Tip60 inhibits the activation of CREB protein by protein kinase A. This occurs via an interaction between Tip60 and CREB and is not dependent on HAT activity of Tip60 (Gavaravarapu and Kamine, 2000)

In addition to the interactions just discussed, a number of papers have shown that Tip60 interacts with various other proteins including the alpha chain of the interleukin-9 receptor (Sliva *et al.*, 1999), NFκB p50 (Dechend *et*

al., 1999) and the androgen receptor (Brady *et al.*, 1999). Tip60 also induces transactivation through the estrogen receptor and progesterone receptor in a ligand dependent manner. This work showed that Tip60 is a receptor coactivator and enhances transactivation at similar levels as the steroid receptor coactivator 1 (SRC 1) (Brady *et al.*, 1999). As discussed in the Introduction, SRC1 is a histone acetyltransferase as well (Spencer *et al.*, 1997), and is associated with an untranslated RNA that is a coactivator (Lanz *et al.*, 1999).

The phylogenetic increase in complexity seen with the MSL3/MRG15 proteins is not observed with the MYST HATs. If, however, chromodomains function as "a vehicle that delivers both positive and negative transcription regulators to the sites of their action on chromatin" (Koonin *et al.*, 1995), then it may not be surprising that there are multiple MYST HATs found in all of these organisms, since the MSL3/MRG15 proteins could target the HATs to the appropriate genes. This increase in MSL3/MRG15 family complexity may reflect the increase in the complexity of transcriptional regulation in higher organisms.

C. Future directions

The human homologues: The work with hMOF was initiated with the ultimate goal of determining if there is a human MSL-like complex. Now

that a truehuman homologue of MSL3 has been identified (Prakash *et al.*, 1999), we should next determine if it interacts with hMOF. Additionally the HAT activity of hMOF should be thoroughly characterized. Very little is known about the function of MRG15. Although it was initially thought to be related to a protein (MORF4) involved in cell senescence (Bertram *et al.*, 1999), this work has since been recanted (Bryce *et al.*, 1999) and the function is now unknown. It would also be of interest to determine if MRG15 is associated with a HAT activity.

The Drosophila homologues: I have found no evidence for an interaction between dTIP60 and dMRG15, but there are other proteins associated with dMRG15. These should be identified and characterized. Many of the experimental questions asked of dTIP60 could have been answered, had I had a precipitating dTIP60 antibody. It is likely that the C-terminal Flag tag was masked and therefore not available to the antibody. One solution may be to tag it at the N-terminus. Additionally, polyclonal anitsera generated to dTIP60 would be useful in characterizing the localization and partners of dTIP60. Once dTIP60 is isolated, it should be tested for histone acetyltransferase activity. Any proteins shown to be associated with dTIP60 should also be characterized.

It is still not known how changes in chromatin modulate the transcriptional rate; i.e. how the nature of various forms of chromatin remodeling and modification increase the access of RNA polymerase to the

DNA template. The study of transcription coactivators and chromatin modification and remodeling complexes is informing the study of the Drosophila dosage compensation mechanism. As more data are gathered about the unique nature of this particular Drosophila complex, those who study Drosophila dosage compensation are making invaluable contributions to the growing body of biochemical work on transcriptional coactivator complexes. The work on the MSL complex and the proteins characterized in this thesis is well positioned to help elucidate these mechanisms of transcriptional regulation.

MATERIALS AND METHODS

Isolation of poly A+ RNA: Total RNA was isolated from cells for RACE and Northern analysis using the Qiagen RNeasy Midi Kit. After quantification, the poly A+ RNA was selected using the RNAeasy Oligotex mRNA kit.

Northern analysis: The 12 Lane Multiple Tissue Northern Blot was probed with cDNA fragments as follows: the ~350 bp AvaI fragment from hMOF and the ~750 bp EcoRI/ Bam HI from from hMRG15. The filter was prehybridized with ClonTech Express Hyb solution at 68 degrees for 30 minutes. Following addition of the probe, hybridization continued for one hour. The filter was washed in 2X SSC, 0.05% SDS 3 times for 10 minutes each. A second set of washes were performed with 0.1X SSC, 0.1% SDS 2 times for 20 minutes each. Probe was removed by placing the filter in 0.5% SDS at 95 degrees for 10 minutes and then allowing the 0.5% SDS to cool for 10 minutes. The filter was then exposed to film overnight to ensure removal of the probe.

Western analysis: Proteins were separated on denaturing 10% SDS polyacrylamide gels. After blotting on nitrocellulose, filters were blocked in 5% Carnation non-fat dry milk in TBS with 0.1% Tween-20 (TTBS). After

probing with primary (mouse monoclonal anti-Flag M2 from Sigma, mouse *monoclonal anti-HA* 12CA5 from Roche) and secondary (HRP conjugated *anti-mouse from* Amersham Pharmacia) antibodies, filters were washed and *exposed to film for developing.*

Sequencing: All sequencing was done off-site at the Iowa State *University* (Ames, IA) Molecular Biology Sequencing Core Facility using the *ABI automatic sequencing system.*

Expression of HIS-tagged proteins: The hMOFC and MRG15 constructs *were generated* using mutagenic PCR to introduce necessary restriction sites. *In the first case,* the primers 5′ GTA CAG AAG AAC CAT ATG AAG TAC CTG AGC GAG C 3′ and 5′ AGC AGG GGG GAT CCT GCT CAC TTC 3′were *used with the* yy60e07 EST clone and in the second case, the primers 5′ GAA TCA CTT ACA TAT GGC GCC GAA GCA G 3′ and 5′ GAG TGA GAG GGA TCC TCA CAC AGC TTT C 3′ were used with the cDNA clone (GenBank *accession number* AF 152245) to generate the respective PCR products which *were subcloned* into pCR 2.1 TA (Invitrogen) and subsequently into pET 19b *for tagging* and bacterial expression.

Plasmid constructs were transformed into BL21(DE3) pLysS cells (Stratagene) and grown to OD_{600}=.6 before induction with 0.4 mM IPTG. Cells continued to grow under induction for 2.5 hours at 30° before harvesting.

Expression of Flag tagged proteins: All Flag tagged proteins were generated using mutagenic PCR to introduce necessary restriction sites. PCR products were subcloned into pCR2.1 TA (Invitrogen) and subsequently into pMK33cFlag for tagging and transfection. Primers 5′ TTG TCG ACA TAA AAT GGG AGA AGT AAA AC 3′ and 5′ AAG ATA TCC TGT GCA TTT CGC AC 3′ were used to construct dMRG15-Flag. Primers 5′ GTT GAT ATC CAA AAC ATG AAA ATT AAC CAC 3′ and 5′ CGC GAT ATC TTT GGA GCG CTT GG 3′ were used to construct dTIP60- Flag.

Expression of dTIP-60 HA tagged protein: dTIP60-HA was generated using mutagenic PCR with primers 5′ ATG CTC GAG AAA ATT AAC CAC AAA TAT GAG 3′ and 5′ CTG ATA TCT TTG GAG CGC TTG GAC 3′. The PCR product was subcloned as above.

"traditional' cDNA Library Screening: General manipulations were done as described in Sambrook *et al.*,1989.

HAT Activity Assays: Assays using calf thymus histones (Sigma) as substrates were preformed in a buffer with 50 mM Tris pH 8.0, 10% glycerol, 1 mM PMSF, 1 mM DTT. Reactions were allowed to proceed at 30° C for 20 minutes, after which a portion of the reaction mixture was spotted on

phosphocellulose paper filters which were washed in 50 mM $NaHCO_3$, pH 9.0 and counted in the scintillation counter on the ^3H channel.

Isolation of HeLa/Raji/Schneider 2 cell extracts: Cells were spun down and lysed in 10mM HEPES pH 8.0, 1.5 mM $MgCl_2$, 10mM KCl, 200mM sucrose, 0.05% NP40. Nuclei were incubated on ice for 10 minutes spun down and washed in the initial lysis solution without NP40. Nuclei were lysed in 20mM HEPES pH 7.4, 25% gycerol, 420mM NaCl, 1.5 $MgCl_2$ 0.2 mM EDTA. After incubation on ice for 30 minutes with intermittent vortexing, the mixture was spun a final time and the nuclear proteins extracted in the supernatant.

Cell transfections: Cell were transfected according to the protocol provided by Invitrogen. Briefly, circular plasmid DNA was mixed with .25M $CaCl_2$. HEPES was added (with constant mixing) to a final concentration of 1M. The calcium/DNA precipitate was allowed to form for 30 minutes before being added to the Schneider 2 cells. Cells were incubated in the presence of the precipitate for 16 hours at 23.5°C. Cells were then washed and resuspended in medium. After 48 hours, selection was begun using 200ug/ml hygromycin. Selection continued for approximately three weeks before the selective agent was removed.

Immunofluorescence: Stably transfected Schneider Line 2 cells were allowed to adhere to a slide for 60 hours in the presence of $CuSO_4$. Cells were washed three times in PBS and fixed in 4% formaldehyde in PBS for 15 minutes. After 3 more washes with PBS the cells were permeablized in 0.1% Triton X-100, 1% BSA in PBS for 20 minutes, washed in PBS and blocked in PBT (0.2% Tween 20,1% BSA in PBS) for 30 minutes. The cells were incubated in the primary antibody (mouse monoclonal anti-Flag M2 from Sigma, rabbit polyclonal anti-HA from Clontech) diluted in PBT overnight at 4°. After two more washes in PBS, the slides were blocked in PBT with 0.12% donkey serum for 30 minutes, washed once more in PBS and incubated in the secondary antibody (Jackson Labs) diluted in PBT at room temperature and in the dark. Slides were mounted in VectaStain with DAPI andviewed using epifluoresence or confocal microscopy.

Rapid amplification cDNA ends: RACE was performed according to the protocols provided with the GibCo BRL or Clontech RACE systems. Any changes to these protocols are noted in the text.

Immunoprecipitation: 250 mg of nuclear protein were incubated with 100ul anti-Flag M2 antibody-agarose (Sigma) for 1.5 hour at 4 degrees. Bead-antibody-protein ternary complex was spun down and the unbound fraction removed. Beads were washed six times in PBS with 1mM PMSF. Proteins

were eluted from beads in 2X SDS PAGE gel sample buffer and analysed by electrophoresis.

Maintenence of cells in culture: Transfected and untransfected cells were grown at ambient CO_2 concentrations at 23.5 degrees in either GibCo BRL Schneider 2 cell medium supplemented with 10% FBS and penicillin and streptomycin or in HyClone serum free Schneider 2 medium supplemented with penicillin and streptomycin.

Primer synthesis: All primers were synthesized by GibCo BRL oligo synthesis facility.

Peptide synthesis: All oligopeptides were synthesized by the Emory University Peptide Synthesis facility (Atlanta, GA)

Antisera generation: All polyclonal antisera were generated by Pocono Rabbit Farm and Laboratory (Canadensis, PA) according to their standard protocols for injection of immunogen and collection of sera.

Metabolic labeling: Approximately 10^8 cells were labeled with 0.5 mCi ^{35}S- methionine (Amersham Pharmacia) for 16 hours in media with 10% the usual amount of methionine. Cells were then spun down and resuspended

in complete medium and allowed to grow for 36 hours. Radiolabled nuclear protein was isolated as described above.

LIST OF ABBREVIATIONS

BLAST	Basic Local Alignment Search Tool
CBP	CREB Binding Protein
CHD	Chromo and Helicase Domains
Chromo	CHRomatin Organization and MOdification
COBBLER	COnsensus Biasing By Locally Embedding Residues
DAPI	4'6-diamidino-2 phenylindole
DTT	Dithiotheitol
ESA1	Essential Sas related Acetyltransferase 1
EST	Expressed Sequence Tag
GFP	Green Fluorescent Protein
HAT	Histone AcetylTransferase
HeLa	Henrietta Lacks (1920 - 1951)
HIV	Human Immunodeficiency Virus
IPTG	Isopropyl β-D-thiogalactopyranoside
MLE	Maleless
MOF	Males Absent on the First (chromosome)
MRG	MORF4 Related Gene
MSL	Male Specific Lethal
MYST	MOZ, YBF2/SAS3, SAS2 and Tip60
NuA4	Nucleosomal Acetyltransferase, Histone 4
PAGE	Polyacrylamide Gel Electrophoresis

PBS	Phosphate Buffered Saline
PBT	PBS/BSA/Tween20
PCR	Polymerase chain reaction
PMSF	Phenylmethylsulfonyl flouride
PVDF	Polyvinylidene difloride
RACE	Rapid Amplification of cDNA Ends
RHA	RNA Helicase A
RING	Really Interesting New Gene
RT-PCR	Reverse Transcriptase - PCR
roX	RNA on the X
SDC	Sex determination and Dosage Compensation
SDS	Sodium dodecyl sulfate
SNF	Sucrose non-fermenting
STS	Sequence tagged site
SWI	mating switch defective
TIP60	Tat interacting protein, 60 kDa

LITERATURE CITED

Aasland, R. and Stewart, A.F. (1995) The chromo shadow domain, a second chromo domain in heterochromatin- binding protein 1, HP1. *Nucleic Acids Res*, **23**, 3168-3174.

Alberts, B., Bray, D., Lewis J., Raff M., Roberts K., Watson J.D. (1994) *Molecular Biology of the Cell*. Garland Publishing, New York.

Allard, S., Utley, R.T., Savard, J., Clarke, A., Grant, P., Brandl, C.J., Pillus, L., Workman, J.L. and Coté, J. (1999) NuA4, an essential transcription adaptor/histone H4 acetyltransferase complex containing Esa1p and the ATM-related cofactor Tra1p. *Embo J*, **18**, 5108-5119.

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

Anderson, S.F., Schlegel, B.P., Nakajima, T., Wolpin, E.S. and Parvin, J.D. (1998) BRCA1 protein is linked to the RNA polymerase II holoenzyme complex via RNA helicase A. *Nat Genet*, **19**, 254-256.

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.

Belote, J.M. and Lucchesi, J.C. (1980a) Control of X chromosome transcription by the maleless gene in Drosophila. *Nature*, **285**, 573-575.

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

Bertram, M.J., Berube, N.G., Hang-Swanson, X., Ran, Q., Leung, J.K., Bryce, S., Spurgers, K., Bick, R.J., Baldini, A., Ning, Y., Clark, L.J., Parkinson, E.K., Barrett, J.C., Smith, J.R. and Pereira-Smith, O.M. (1999) Identification of a gene that reverses the immortal phenotype of a subset of cells and is a member of a novel family of transcription factor-like genes. *Mol Cell Biol*, **19**, 1479-1485.

Brady, M.E., Ozanne, D.M., Gaughan, L., Waite, I., Cook, S., Neal, D.E. and Robson, C.N. (1999) Tip60 is a nuclear hormone receptor coactivator. *Journal of Biological Chemistry*, **274**, 17599-17604.

Brand, M., Yamamoto, K., Staub, A. and Tora, L. (1999) Identification of TATA-binding protein-free TAFII-containing complex subunits suggests a role in nucleosome acetylation and signal transduction. *J Biol Chem*, **274**, 18285-18289.

Brasher, S.V., Smith, B.O., Fogh, R.H., Nietlispach, D., Thiru, A., Nielsen, P.R., Broadhurst, R.W., Ball, L.J., Murzina, N.V. and Laue, E.D. (2000) The structure of mouse HP1 suggests a unique mode of single peptide recognition by the shadow chromo domain dimer. *Embo J*, **19**, 1587-1597.

Breen, T.R. and Lucchesi, J.C. (1986) Analysis of the dosage compensation of a specific transcript in *Drosophila melanogaster*. *Genetics*, **112**, 483-491.

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

Bryce, S.D., Forsyth, N.R., Fitzsimmons, S.A., Clark, L.J., Bertram, M.J., Cuthbert, A.P., Newbold, R.F., Pereira-Smith, O.M. and Parkinson, E.K. (1999) Genetic and functional analyses exclude mortality factor 4 (MORF4) as a keratinocyte senescence gene. *Cancer Res*, **59**, 2038-2040.

Cairns, B.R., Lorch, Y., Li, Y., Zhang, M., Lacomis, L., Erdjument-Bromage, H., Tempst, P., Du, J., Laurent, B. and Kornberg, R.D. (1996) RSC, an essential, abundant chromatin-remodeling complex. *Cell*, **87**, 1249-1260.

Chen, H., Lin, R.J., Schiltz, R.L., Chakravarti, D., Nash, A., Nagy, L., Privalsky, M.L., Nakatani, Y. and Evans, R.M. (1997) Nuclear receptor coactivator ACTR is a novel histone acetyltransferase and forms a multimeric activation complex with P/CAF and CBP/p300. *Cell*, **90**, 569-580.

Cho, H., Orphanides, G., Sun, X., Yang, X.J., Ogryzko, V., Lees, E., Nakatani, Y. and Reinberg, D. (1998) A human RNA polymerase II complex containing factors that modify chromatin structure. *Molecular & Cellular Biology*, **18**, 5355-5363.

Copps, K., Richman, R., Lyman, L.M., Chang, K.A., Rampersad-Ammons, J. and Kuroda, M.I. (1998) Complex formation by the Drosophila

MSL proteins: role of the MSL2 RING finger in protein complex assembly. *EMBO Journal*, **17**, 5409-5417.

Coté, J., Quinn, J., Workman, J.L. and Peterson, C.L. (1994) Stimulation of GAL4 derivative binding to nucleosomal DNA by the yeast SWI/SNF complex. *Science*, **265**, 53-60.

Cowieson, N.P., Partridge, J.F., Allshire, R.C. and McLaughlin, P.J. (2000) Dimerisation of a chromo shadow domain and distinctions from the chromodomain as revealed by structural analysis [In Process Citation]. *Curr Biol*, **10**, 517-525.

Creaven, M., Hans, F., Mutskov, V., Col, E., Caron, C., Dimitrov, S. and Khochbin, S. (1999) Control of the histone-acetyltransferase activity of Tip60 by the HIV-1 transactivator protein, Tat. *Biochemistry*, **38**, 8826-8830.

Dechend, R., Hirano, F., Lehmann, K., Heissmeyer, V., Ansieau, S., Wulczyn, F.G., Scheidereit, C. and Leutz, A. (1999) The Bcl-3 oncoprotein acts as a bridging factor between NF-kappaB/Rel and nuclear co-regulators. *Oncogene*, **18**, 3316-3323.

Eberharter, A., John, S., Grant, P.A., Utley, R.T. and Workman, J.L. (1998) Identification and analysis of yeast nucleosomal histone acetyltransferase complexes. *Methods*, **15**, 315-321.

Eberharter, A., Sterner, D.E., Schieltz, D., Hassan, A., Yates, J.R., 3rd, Berger, S.L. and Workman, J.L. (1999) The ADA Complex Is a Distinct Histone Acetyltransferase Complex in Saccharomyces cerevisiae. *Mol Cell Biol*, **19**, 6621-6631.

Eisen, A. and Lucchesi, J.C. (1998) Unraveling the role of helicases in transcription. *Bioessays*, **20**, 634-641.

Eissenberg, J.C., James, T.C., Foster-Hartnett, D.M., Hartnett, T., Ngan, V. and Elgin, S.C. (1990) Mutation in a heterochromatin-specific chromosomal protein is associated with suppression of position-effect variegation in Drosophila melanogaster. *Proc Natl Acad Sci U S A*, **87**, 9923-9927.

Ellis, N.A., Groden, J., Ye, T.Z., Straughen, J., Lennon, D.J., Ciocci, S., Proytcheva, M. and German, J. (1995) The Bloom's syndrome gene product is homologous to RecQ helicases. *Cell*, **83**, 655-666.

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.

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

Gavaravarapu, S. and Kamine, J. (2000) Tip60 inhibits activation of CREB protein by protein kinase A. *Biochem Biophys Res Commun*, **269**, 758-766.

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

Gilbert, S. (2000) *Developmental Biology*. Sinauer Associates, Sunderland, Mass.

Golubowsky, M. D. and Ivanov, Y. N. (1972) Autosomal mutation in *Drosophila melanogaster* killing the males and connected with female sterility. *Drosophila Information Service*, **49**, 117.

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.

Grant, P.A., Duggan, L., Coté, J., Roberts, S.M., Brownell, J.E., Candau, R., Ohba, R., Owen-Hughes, T., Allis, C.D., Winston, F., Berger, S.L. and Workman, J.L. (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.

Gray, M.D., Shen, J.C., Kamath-Loeb, A.S., Blank, A., Sopher, B.L., Martin, G.M., Oshima, J. and Loeb, L.A. (1997) The Werner syndrome protein is a DNA helicase. *Nat Genet*, **17**, 100-103.

Gu, W., Szauter, P. and Lucchesi, J.C. (1998) Targeting of MOF, a putative histone acetyl transferase, to the X chromosome of *Drosophila melanogaster*. *Developmental Genetics*, **22**, 56-64.

Heard, E., Clerc, P. and Avner, P. (1997) X-chromosome inactivation in mammals. *Annu Rev Genet*, **31**, 571-610.

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.

Iizuka, M. and Stillman, B. (1999) Histone acetyltransferase HBO1 interacts with the ORC1 subunit of the human initiator protein. *Journal of Biological Chemistry*, **274**, 23027-23034.

Ito, T., Levenstein, M.E., Fyodorov, D.V., Kutach, A.K., Kobayashi, R. and Kadonaga, J.T. (1999) ACF consists of two subunits, Acf1 and ISWI, that function cooperatively in the ATP-dependent catalysis of chromatin assembly. *Genes Dev*, **13**, 1529-1539.

Jenster, G., Spencer, T.E., Burcin, M.M., Tsai, S.Y., Tsai, M.J. and O'Malley, B.W. (1997) Steroid receptor induction of gene transcription: a two-step model. *Proc Natl Acad Sci U S A*, **94**, 7879-7884.

Jin, Y., Wang, Y., Walker, D.L., Dong, H., Conley, C., Johansen, J. and Johansen, K.M. (1999) JIL-1: a novel chromosomal tandem kinase implicated in transcriptional regulation in Drosophila. *Mol Cell*, **4**, 129-135.

John, S., Howe, L., Tafrov, S.T., Grant, P.A., Sternglanz, R. and Workman, J.L. (2000) The Something About Silencing protein, Sas3, is the catalytic subunit of NuA3, a yTAF(II)30-containing HAT complex that interacts with the Spt16 subunit of the yeast CP (Cdc68/Pob3)-FACT complex. *Genes Dev*, **14**, 1196-1208.

Kamine, J., Elangovan, B., Subramanian, T., Coleman, D. and Chinnadurai, G. (1996) Identification of a cellular protein that specifically

interacts with the essential cysteine region of the HIV-1 Tat transactivator. *Virology*, **216**, 357-366.

Kazazian, H.H., Jr., Young, W.J. and Childs, B. (1965) X-linked 6-phosphogluconate dehydrogenase in drosophila: subunit associations. *Science*, **150**, 1601-1602.

Kelley, R.L., Solovyeva, I., Lyman, L.M., Richman, R., Solovyev, V. and Kuroda, M.I. (1995) Expression of msl-2 causes assembly of dosage compensation regulators on the X chromosomes and female lethality in Drosophila. *Cell*, **81**, 867-877.

Kiernan, R.E., Vanhulle, C., Schiltz, L., Adam, E., Xiao, H., Maudoux, F., Calomme, C., Burny, A., Nakatani, Y., Jeang, K.T., Benkirane, M. and Van Lint, C. (1999) HIV-1 tat transcriptional activity is regulated by acetylation. *Embo J*, **18**, 6106-6118.

Kimura, A. and Horikoshi, M. (1998) Tip60 acetylates six lysines of a specific class in core histones in vitro. *Genes to Cells*, **3**, 789-800.

Koonin, E.V., Zhou, S. and Lucchesi, J.C. (1995) The chromo superfamily: new members, duplication of the chromo domain and possible role in delivering transcription regulators to chromatin. *Nucleic Acids Res*, **23**, 4229-4233.

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.

Lachaise, D., Cariou, M-L, David, J.R., Lemeunier, F., Tsacas, L., Ashburner, M. (1988) Historical Biogeography of the *Drosophila melaogaster* species subgroup. *Evolutionary Biology* **22**, 159-225

Lanz, R.B., McKenna, N.J., Onate, S.A., Albrecht, U., Wong, J., Tsai, S.Y., Tsai, M.J. and O'Malley, B.W. (1999) A steroid receptor coactivator, SRA, functions as an RNA and is present in an SRC-1 complex. *Cell*, **97**, 17-27.

Lavender, J.S., Birley, A.J., Palmer, M.J., Kuroda, M.I. and Turner, B.M. (1994) Histone H4 acetylated at lysine 16 and proteins of the Drosophila dosage compensation pathway co-localize on the male X chromosome through mitosis. *Chromosome Res*, **2**, 398-404.

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

Lee, C.G., da Costa Soares, V., Newberger, C., Manova, K., Lacy, E. and Hurwitz, J. (1998a) RNA helicase A is essential for normal gastrulation. *Proceedings of the National Academy of Sciences of the United States of America*, **95**, 13709-13713.

Lee, C.G., Eki, T., Okumura, K., da Costa Soares, V. and Hurwitz, J. (1998b) Molecular analysis of the cDNA and genomic DNA encoding mouse RNA helicase A. *Genomics*, **47**, 365-371.

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

Li, J., Tang, H., Mullen, T.M., Westberg, C., Reddy, T.R., Rose, D.W. and Wong-Staal, F. (1999) A role for RNA helicase A in post-transcriptional regulation of HIV type 1. *Proc Natl Acad Sci U S A*, **96**, 709-714.

Lucchesi, J.C. (1983) Dosage compensation in Drosophila. *Isozymes Curr Top Biol Med Res*, **9**, 179-188.

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

Meller, V.H., 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.

Mizzen, C.A. and Allis, C.D. (1998) Linking histone acetylation to transcriptional regulation. *Cell Mol Life Sci*, **54**, 6-20.

Mukherjee, A.S. and Beermann, W. (1965) Synthesis of ribonucleic acid by the X-chromosomes of *Drosophila melanogaster* and the problem of dosage compensation. *Nature*, **207**, 785-786.

Muller, H.J. (1932) Further studies on the nature and causes of gene mutations. *Proc. 6th Int Congr. Genet. Res.*, **1**, 213-255.

Nakajima, T., Uchida, C., Anderson, S.F., Lee, C.G., Hurwitz, J., Parvin, J.D. and Montminy, M. (1997) RNA helicase A mediates association of CBP with RNA polymerase II. *Cell*, **90**, 1107-1112.

Neal, K.C., Pannuti, A., Smith, E.R. and Lucchesi, J.C. (2000) A new human member of the MYST family of histone acetyl transferases with high sequence similarity to Drosophila MOF . *Biochim Biophys Acta*, **1490**, 170-174.

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.

Ohno, S. (1967) *Sex chromosomes and sex-linked genes*. Springer-Verlag, New York.

Orphanides, G., LeRoy, G., Chang, C.H., Luse, D.S. and Reinberg, D. (1998) FACT, a factor that facilitates transcript elongation through nucleosomes. *Cell*, **92**, 105-116.

Ott, M., Schnolzer, M., Garnica, J., Fischle, W., Emiliani, S., Rackwitz, H.R. and Verdin, E. (1999) Acetylation of the HIV-1 Tat protein by p300 is important for its transcriptional activity. *Curr Biol*, **9**, 1489-1492.

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.

Paro, R. (1990) Imprinting a determined state into the chromatin of Drosophila. *Trends Genet*, **6**, 416-421.

Paro, R. and Hogness, D.S. (1991) The Polycomb protein shares a homologous domain with a heterochromatin- associated protein of Drosophila. *Proc Natl Acad Sci U S A*, **88**, 263-267.

Prakash, S.K., Van den Veyver, I.B., Franco, B., Volta, M., Ballabio, A. and Zoghbi, H.Y. (1999) Characterization of a novel chromo domain gene in xp22.3 with homology to Drosophila msl-3. *Genomics*, **59**, 77-84.

Richter, L., Bone, J.R. and Kuroda, M.I. (1996) RNA-dependent association of the Drosophila maleless protein with the male X chromosome. *Genes to Cells*, **1**, 325-336.

Sambrook, J., Fritsch, E.F. and Maniatis, T. (1989) Molecular cloning: A Laboratory manual, 2nd edition. Cold Spring Harbor Press, Cold Spring Harbor, New York.

Schnitzler, G., Sif, S. and Kingston, R.E. (1998) Human SWI/SNF interconverts a nucleosome between its base state and a stable remodeled state. *Cell*, **94**, 17-27.

Siddique, H., Zou, J.P., Rao, V.N. and Reddy, E.S. (1998) The BRCA2 is a histone acetyltransferase. *Oncogene*, **16**, 2283-2285.

Sliva, D., Zhu, Y.X., Tsai, S., Kamine, J. and Yang, Y.C. (1999) Tip60 interacts with human interleukin-9 receptor alpha-chain. *Biochemical & Biophysical Research Communications*, **263**, 149-155.

Smith, E.R., Eisen, A., Gu, W., Sattah, M., Pannuti, A., Zhou, J., Cook, R.G., Lucchesi, J.C. and Allis, C.D. (1998) ESA1 is a histone acetyltransferase that is essential for growth in yeast. *Proc Natl Acad Sci U S A*, **95**, 3561-3565.

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.

Spencer, T.E., Jenster, G., Burcin, M.M., Allis, C.D., Zhou, J., Mizzen, C.A., McKenna, N.J., Onate, S.A., Tsai, S.Y., Tsai, M.J. and O'Malley, B.W. (1997) Steroid receptor coactivator-1 is a histone acetyltransferase. *Nature*, **389**, 194-198.

Stokes, D.G. and Perry, R.P. (1995) DNA-binding and chromatin localization properties of CHD1. *Molecular & Cellular Biology*, **15**, 2745-2753.

Takechi, S. and Nakayama, T. (1999) Sas3 is a histone acetyltransferase and requires a zinc finger motif. *Biochem Biophys Res Commun*, **266**, 405-410.

Tanaka, A., Fukunaga, A. and Oishi, K. (1976) Studies on the sex-specific lethals of *Drosophila melanogaster*. II. Further studies on a male-specific lethal gene, maleless. *Genetics*, **84**, 257-266.

Tsukiyama, T. and Wu, C. (1995) Purification and properties of an ATP-dependent nucleosome remodeling factor. *Cell*, **83**, 1011-1020.

Uchida S., U.T., Oishi K. (1981) Studies on the sec-specific lethals of *Drosophila melanogaster*. III. A thirst chromosome male-specific lethal mutant. *Jpn. J. Genet.*, **56**, 523-527.

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 [published erratum appears in Nature 1997 Oct 30;389(6654):1003]. *Nature*, **388**, 598-602.

Wei, W., Twell, D. and Lindsey, K. (1997) A novel nucleic acid helicase gene identified by promoter trapping in Arabidopsis. *Plant Journal*, **11**, 1307-1314.

Wilson, E.B. (1905) The chromosomes in relation to the determination of sex in insects. *Science*, **22**, 500-502.

Woodage, T., Basrai, M.A., Baxevanis, A.D., Hieter, P., Collins, F.S. (1997) Characterization of the CHD family of proteins. *Proc. Natl. Acad. Sci.*, **94**, 11472-11477.

Yamamoto, T. and Horikoshi, M. (1997) Novel substrate specificity of the histone acetyltransferase activity of HIV-1-Tat interactive protein Tip60. *Journal of Biological Chemistry*, **272**, 30595-30598.

Young, W.J. (1966) X-linked electrophoretic variation in 6-phosphogluconate dehydrogenase in *Drosophila melanogaster*. *J Hered*, **57**, 58-60.

Yu, J.X., Chao, L., Ward, D.C. and Chao, J. (1996) Structure and chromosomal localization of the human prostasin (PRSS8) gene. *Genomics*, **32**, 334-340.

Zhang, S., Maacke, H. and Grosse, F. (1995) Molecular cloning of the gene encoding nuclear DNA helicase II. A bovine homologue of human RNA helicase A and Drosophila Mle protein. *J Biol Chem*, **270**, 16422-16427.

Zhou, S., Yang, Y., Scott, M.J., Pannuti, A., Fehr, K.C., Eisen, A., Koonin, E.V., Fouts, D.L., Wrightsman, R., Manning, J.E. and et al. (1995) Male-specific lethal 2, a dosage compensation gene of Drosophila, undergoes

sex-specific regulation and encodes a protein with a RING finger and a metallothionein-like cysteine cluster. *Embo J*, **14**, 2884-2895.

EMORY UNIVERSITY
GRADUATE SCHOOL OF ARTS AND SCIENCES

Graduate Division of Biological and Biomedical Sciences

Program in Genetics and Molecular Biology

Characterization of Human and Drosophila Homologues of Two Drosophila Dosage Compensation Genes

Karama Carrol Neal
B.A., Swarthmore College, 1993

Wednesday, June 14, 2000 at 10 am
Room 308, Dental School Building

An Oral Presentation of a Dissertation
submitted to the faculty of the Graduate School of
Emory University
in partial fulfillment of the requirements for the degree of
Doctor of Philosophy

Committee in Charge:

John C. Lucchesi, Ph.D., Advisor
Jeremy M. Boss, Ph.D.
Victoria Finnerty, Ph.D.
Sue Jinks-Robertson, Ph.D.
Barry Yedvobnick, Ph.D.

ENVOI

I conclude this volume with the following two pieces of information which I place here simply because they need to be recorded:

First, although Drosophila melanogaster are now quite cosmopolitan, they likely originated in West Africa (Lachaise, 1988).

Second, during the first lecture of the 1996 Cold Spring Harbor Course on Advanced *Drosophila* Genetics, Professor Michael Ashburner suggested that *Drosophila melanogaster* migrated to the Western Hemisphere during the trade in African people (Ashburner, pers. comm.).

Abstract:

Dosage compensation is the mechanism by which the amount of X-linked gene product is equalized between the males and females. Drosophila accomplish this by a two fold hypertranscription of the X chromosome in males. Five proteins are known to regulate this process and functional absence of any one of the five causes male specific lethality. In the last few years, the genes known to be involved in dosage compensation, collectively known as male specific lethals or MSLs, have been cloned and characterized. Biochemical analyses of the Drosophila dosage compensation machinery have linked this process to more widely conserved processes of chromatin modification and remodeling. The MSLs are associated in a multiprotein complex that binds hundreds of sites on the X chromosome of male flies. The male X chromatin also has a specific isoform of histone H4, acetylated at lysine 16. The MSL protein MOF, a histone acetyltransferase (HAT), has been shown to be responsible for this modification. It is thought that this modification, in combination with the activities of the other MSLs including an RNA helicase, results in the hypertranscription of X-linked genes in males.

Several of these MSL genes are evolutionarily conserved. Homologues of the RNA helicase MLE are found in mammals. Additionally, homologues of MOF, a MYST family histone acetyltransferase, and MSL3, a chromo and chromoshadow domain-containing protein, have been found in yeasts, mammals and flies. Because of the conserved nature of these proteins, I hypothesize that MYST family HATs associate specifically with MSL3-like proteins, as is the case in the Drosophila dosage compensation complex. In this work, I explore this hypothesis in human and Drosophila by examining activity, localization and interaction partners of MOF and MSL3 homologues in humans and Drosophila. Specifically, I present evidence that the Drosophila homologues are present in multiprotein complexes and characterize the HAT activity of hMOF, a human MYST HAT and homologue of Drosophila MOF. Possible cellular roles for the human and Drosophila proteins are discussed.

Publications:

KC Neal, A Pannuti, ER Smith and JC Lucchesi (2000) A new human member of the MYST family of histone acetyl transferases with high sequence similarity to Drosophila MOF *Biochemica et Biophysica Acta* **1490**(1): 170-174

Eisen, A, M Sattah, T Gazitt, K Neal, P Szauter, JC Lucchesi. (1998) A novel DEAD-box RNA helicase exhibits high sequence conservation from yeast to humans and localizes to the cytoplasm in Drosophila. *Biochemica et Biophysica Acta* **1397**(2): 131-136.

Benes, H., KC Neal, RL Willis, D Gadde, AB Castleberry, SE Korochkina. (1996) Overlappling *Lsp-2* gene sequences target expression to both the larval and adult *Drosophila* fat body. *Insect Molecular Biology* **5**(1) 39-49.

Gordesky-Gold, B., JM Warrick, DP Kutzler, KC Neal, CM Coughlin, L Tompkins. (1996) Interstrain variability of larval photokinesis in *Drosophila melanogaster. Behavior Genetics* **26**: 49-54.

Benes, H., DW Spivey, J Miles, K Neal, RG Edmondson. (1990) Fat body specific expression of the Drosophila *Lsp-2* gene. *SAAS Bulletin of Biochemistry and Biotechnology* **3**: 129-133.

Presentation:

Eisen, A, KC Neal, T Gazitt, JC Lucchesi. (1997) A Drosophila homologue of a putative murine DEAD-box RNA helicase. Presented at the National Drosophila Conference, Chicago, IL.

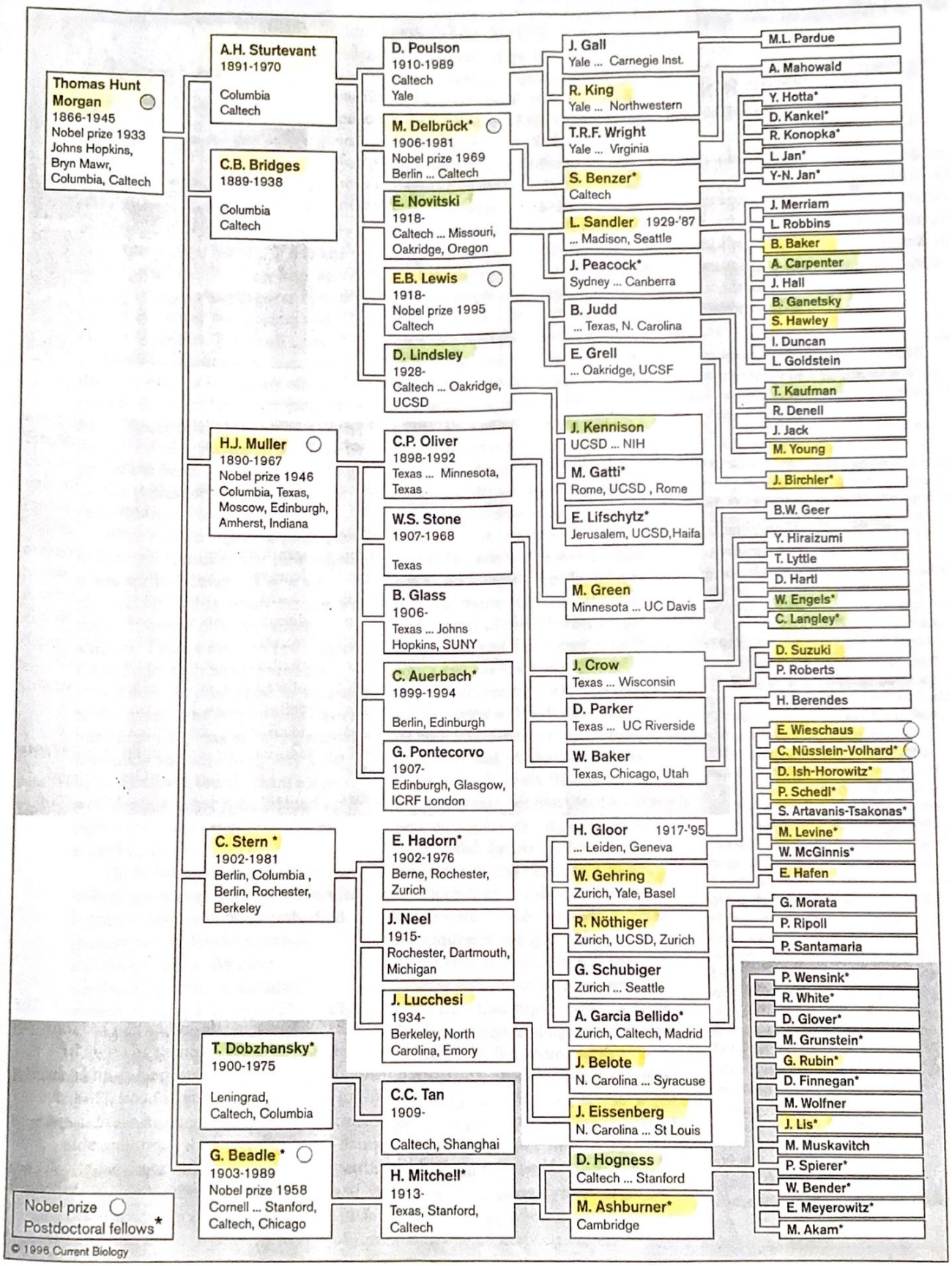
Honors and Awards:

NIH Minority Predoctoral Fellowship, Fall 1996-Fall 1999

Ford Foundation Doctoral Fellowship for Minorities, Honorable Mention, Spring 1996.

NIH Training Grant, Emory University, Summer 1994-Summer 1996.

Who's Who among Students at American Colleges and Universities, Spring 1996.


Thomas Hunt Morgan
1866-1945
Nobel prize 1933
Johns Hopkins, Bryn Mawr, Columbia, Caltech
A.H. Sturtevant
1891-1970
Columbia
Caltech
C.B. Bridges
1889-1938
Columbia
Caltech
H.J. Muller
1890-1967
Nobel prize 1946
Columbia, Texas, Moscow, Edinburgh, Amherst, Indiana
C. Stern *
1902-1981
Berlin, Columbia , Berlin, Rochester, Berkeley
T. Dobzhansky*
1900-1975
Leningrad, Caltech, Columbia
G. Beadle *
1903-1989
Nobel prize 1958
Cornell ... Stanford, Caltech, Chicago
D. Poulson
1910-1989
Caltech
Yale
M. Delbrück*
1906-1981
Nobel prize 1969
Berlin ... Caltech
E. Novitski
1918-
Caltech ... Missouri, Oakridge, Oregon
E.B. Lewis
1918-
Nobel prize 1995
Caltech
D. Lindsley
1928-
Caltech ... Oakridge, UCSD
C.P. Oliver
1898-1992
Texas ... Minnesota, Texas
W.S. Stone
1907-1968
Texas
B. Glass
1906-
Texas ... Johns Hopkins, SUNY
C. Auerbach*
1899-1994
Berlin, Edinburgh
G. Pontecorvo
1907-
Edinburgh, Glasgow, ICRF London
E. Hadorn*
1902-1976
Berne, Rochester, Zurich
J. Neel
1915-
Rochester, Dartmouth, Michigan
J. Lucchesi
1934-
Berkeley, North Carolina, Emory
C.C. Tan
1909-
Caltech, Shanghai
H. Mitchell*
1913-
Texas, Stanford, Caltech
J. Gall
Yale ... Carnegie Inst.
R. King
Yale ... Northwestern
T.R.F. Wright
Yale ... Virginia
S. Benzer*
Caltech
L. Sandler 1929-'87
... Madison, Seattle
J. Peacock*
Sydney ... Canberra
B. Judd
... Texas, N. Carolina
E. Grell
... Oakridge, UCSF
J. Kennison
UCSD ... NIH
M. Gatti*
Rome, UCSD , Rome
E. Lifschytz*
Jerusalem, UCSD, Haifa
M. Green
Minnesota ... UC Davis
J. Crow
Texas ... Wisconsin
D. Parker
Texas ... UC Riverside
W. Baker
Texas, Chicago, Utah
H. Gloor 1917-'95
... Leiden, Geneva
W. Gehring
Zurich, Yale, Basel
R. Nöthiger
Zurich, UCSD, Zurich
G. Schubiger
Zurich ... Seattle
A. Garcia Bellido*
Zurich, Caltech, Madrid
J. Belote
N. Carolina ... Syracuse
J. Eissenberg
N. Carolina ... St Louis
D. Hogness
Caltech ... Stanford
M. Ashburner*
Cambridge
M.L. Pardue
A. Mahowald
Y. Hotta*
D. Kankel*
R. Konopka*
L. Jan*
Y-N. Jan*
J. Merriam
L. Robbins
B. Baker
A. Carpenter
J. Hall
B. Ganetsky
S. Hawley
I. Duncan
L. Goldstein
T. Kaufman
R. Denell
J. Jack
M. Young
J. Birchler*
B.W. Geer
Y. Hiraizumi
T. Lyttle
D. Hartl
W. Engels*
C. Langley*
D. Suzuki
P. Roberts
H. Berendes
E. Wieschaus
C. Nüsslein-Volhard*
D. Ish-Horowitz*
P. Schedl*
S. Artavanis-Tsakonas*
M. Levine*
W. McGinnis*
E. Hafen
G. Morata
P. Ripoll
P. Santamaria
P. Wensink*
R. White*
D. Glover*
M. Grunstein*
G. Rubin*
D. Finnegan*
M. Wolfner
J. Lis*
M. Muskavitch
P. Spierer*
W. Bender*
E. Meyerowitz*
M. Akam*
Nobel prize
Postdoctoral fellows*
© 1996 Current Biology

Open questions

A forest of principles

Horace Barlow

I am astonished at Lewis Wolpert's claim that all the big principles in biology are understood — especially in developmental biology. If one knew all the principles then new facts would simply slot into their expected places, as they do in any well understood aspect of physics. But in all areas of biology bewildering new facts are discovered at an increasing rate. I frankly do not believe that Wolpert has a conceptual scheme where each new fact simply evokes an acquiescent nod of the head. Either he is goading us into protest or his 'principles' are very different from mine.

What can we expect of principles in biology? Consider 'the principle of quasi-optimal design.' There are physical limits that dictate the relative sizes of parts of different-sized animals. Thus, a mouse's eye is many times larger than an elephant eye scaled down in proportion to the size of the mouse: the latter would have a minute pupil giving appalling spatial resolution and admitting very little light. Similarly, the elephant's legs would break every time it stood up if they were simply the mouse's legs scaled up in size.

These 'design principles' are our minds' generalization from particular instances that have been embodied in genomes under the action of natural selection. We have successfully found a few such generalizations, but we emphatically do not understand all the physical limiting factors that have moulded the genomes of all species, and I have not even heard mention of limiting factors in development, although they must surely exist.

Now consider a problem in neuroscience. The neocortex of the brain is large in primates and huge in humans; comparative neuroanatomists have told us that it stores knowledge of the world, and if they are right, we can understand the selective advantage it gives us. Neurophysiologists, on the other hand, tell us how the neocortex represents sensory stimuli but say nothing about how this representation is analysed, stored, accessed or used. Only a fool could hold that no new principles might emerge from reconciling these two astonishingly different accounts.

I think my examples illustrate the general nature of biological principles. There are not just a few universal ones like the great conservation laws of physics. Instead, they form a forest, and quasi-optimal design is but one tree in it — or perhaps just a branch of the Tree of Adaptation. Wolpert might claim that adaptations are all examples of a single general principle which we already know, but this is unhelpful because we need to know what factor is limiting in each particular case: optics will not help you to understand the elephant's leg bones or neocortical size. If we did understand all the relevant limiting factors, we would be closer to knowing why a species has evolved to a certain average size, why the mass of its neocortex is a particular fraction of its body mass, and so on for many questions that we do not yet even know enough to ask. It is an absolutely safe bet that there are many new principles waiting to be found.

Mendel discovered both a whole range of new facts and the principle by which they could be understood, but few are so lucky nowadays. Thus, my request to the good fairy godmother of science would be "Please give me exact references to some facts that require new concepts for their understanding". I would add "Please pick facts pointing to concepts appropriate for my energies and abilities; I could not handle a tree of Darwinian or Mendelian size, but I would greatly enjoy some of the conceptual fruit still to be found in plenty out among the leafy branches of the Biological Forest of Principles".

Pedigree

The Morgan lineage

Guil Winchester

Intellectual pedigrees convey the longevity and continuity of scientific lineages. The *Drosophila* community founded by Thomas Hunt Morgan and the Columbia Fly Room is still flourishing after eighty-five years.

Figure 1

The pedigree is skewed to show how the *Drosophila* renaissance in the 1980s descends from the Columbia Fly Room. Descendants who founded schools in other organisms are also shown, but their 'heirs' are not (unless they move back into flies). Individuals are listed only once. Thus, interactions within labs can be deduced but not the spread of ideas and techniques via the movement of postdocs. The pedigree is divided vertically into filial generations and horizontally into sublineages. At the top left are the 'triumvirate' who 'invented' *Drosophilia* as a genetic organism, Morgan and his two graduate students Sturtevant and Bridges. All three moved to Caltech when Morgan founded the Division of Biology (1928), and the pink band contains Morgan's direct line, which formally descends through Sturtevant. 'F_2' graduate students are those selected by Sturtevant himself (*A History of Genetics*. New York: Harper and Row; 1965), plus Lindsley. Delbrück is also on Sturtevant's Caltech pedigree and illustrates the movement into 'lower' organisms in the 1930s–1940s, followed by the move back into flies in the 1960s–1970s. (Delbrück was a cofounder of the 'phage group; Benzer founded *Drosophila* neurogenetics.) The beige and yellow bands also descend from Columbia. Muller, a semi-detached member of the Fly Room, moved often and founded several schools; only his Texas and Edinburgh heirs are shown here. Stern was the most successful of the Fly Room postdocs: he and Hadorn (an amphibian embryologist who moved into flies via a postdoc with Stern) pioneered *Drosophila* developmental biology. In the lowest band are two Caltech postdocs: Dobzhansky founded a school of *Drosophila* population genetics at Columbia; Beadle 'invented' *Neurospora* as a tool for biochemical genetics and succeeded Morgan as head of Caltech's Division of Biology. Two of Beadle's heirs moved back into flies: Mitchell in the 1950s and Hogness in the 1960s. The Hogness laboratory pioneered *Drosophila* molecular biology and launched the *Drosophila* renaissance.

In presenting this dissertation as a partial fulfillment of the requirements for an advanced degree from Emory University, I agree that the Library of the University shall make it available for inspection and circulation in accordance with its regulations, governing materials of this type. I agree that permission to copy from, or to publish, this dissertation may be granted by the professor under whose direction it was written, or, in his absence, by the Dean of the Graduate School when such copying or publication is solely for scholarly purposes and does not involve potential financial gain. It is understood that any copying from or publication of this dissertation which involves potential financial gain will not be allowed without written permission.

Karama C. Neal