THE ROLE OF HISTONE LOCUS BODY (HLB) ASSEMBLY AND THE CELL CYCLE IN HISTONE mRNA BIOSYNTHESIS

Esteban Alejandro Terzo

A dissertation submitted in the faculty of the University of North Carolina at Chapel Hill in partial fulfillment of the requirements for the degree of Doctor of Philosophy in the Department of Biology.

Chapel Hill
2015

Approved by:

Robert J. Duronio

William F. Marzluff

Jeff Sekelsky

Jean Cook

Greg Matera

ABSTRACT

Esteban Alejandro Terzo: The role of histone locus body (HLB) assembly and the cell cycle in histone mRNA biosynthesis
(Under the direction of Robert J. Duronio)

The spatial compartmentalization of nuclear processes such as transcription, ribosome biogenesis, cellular response to stress, and histone mRNA biosynthesis into discrete territories is essential for precisely orchestrating diverse gene expression programs. The genome organizes structures called nuclear bodies (NBs) that concentrate factors (proteins, RNA, and Ribonucleoproteins) required for the efficient regulation of gene expression. The levels of the molecular components that constitute NBs fluctuate due to a continuous exchange with the nucleoplasm in response to diverse physiological inputs. Therefore, gaining insight into how NBs assemble and function is essential for understanding their contribution to genome function. We used the *Drosophila* histone locus body (HLB) as a model to ask how HLB assembly and function contribute to the replication-dependent histone gene expression. HLBs form at the histone locus and concentrate factors required for histone mRNA biosynthesis. In addition, CycE/Cdk2 is known to be required for cell cycle-dependent histone mRNA biosynthesis. Our laboratory previously demonstrated that the Multi Sex Combs (Mxc) protein is required for HLB assembly and efficient histone mRNA biosynthesis and is also phosphorylated by CycE/Cdk2. However, the role that Mxc and its CycE/Cdk2-dependent phosphorylation play in HLB assembly and how these contribute to histone gene

expression was not known. To explore the molecular mechanism by which Mxc assembles an HLB and contributes to histone gene expression, we combined biochemical assays, confocal microscopy, and live imaging. We show that Mxc is a multivalent protein that plays an essential role in HLB assembly and in the concentration of histone mRNA biosynthetic factors at the histone locus. We demonstrate that Mxc self-interacts via a heterotypic binding between two N-terminal domains: a LisH domain and a novel self-interaction facilitator (SIF) domain. Further, we show that Mxc's self-interaction is essential for its accumulation at the histone locus and for HLB assembly. We also show that a region between amino acids 721 and 1481 is crucial for HLB assembly independent of the LisH and SIF domains. The last 195 amino acids of Mxc are required for the recruitment of FLASH, a fundamental histone pre-mRNA processing factor, to the HLB. In addition, preliminary data strongly suggest that Mxc harbors multiple CycE/Cdk2 phosphoepitopes required for its concentration at the histone locus and for the assembly of a stable HLB. Combining the present work with our previous findings, we propose that Mxc is capable of forming a three-dimensional scaffold, the HLB, by employing multiple domains, and presumably multiple CycE/Cdk2 phosphoepitopes, to accumulate at the histone locus and to concentrate the required mRNA biosynthetic components at optimal levels to assemble a stable, mature HLB that promotes and maintains histone gene expression throughout S phase.

ACKNOWLEDGEMENTS

Advisor: Dr. Robert Duronio

Committee: Dr. William Marzluff, Dr. Jeff Sekelsky, Dr. Jean Cook, Dr. Greg Matera

Duronio, Marzluff, Sekelsky, and Matera labs

Pamela Malek, Julian Curiel, Robert Williams, Tony Perdue

Dr. Shawn Lyons, Dr. John Poulton, Dr. Brenda Temple, and other collaborators

To my mentor Dr. Robert (Bob) Duronio for his guidance throughout my entire graduate career, his constructive criticism, and his always honest opinion. Bob has played an instrumental role in the development of my communications skills. I truly believe that, today, I am a better communicator of my research, that I can think critically about my project, and write about it much more effectively than I ever thought I would, as a result of his constant effort and fair opinion and evaluation of my work. I have been very fortunate to have Bob as my mentor during this important step of my scientific career. I will always be grateful to him for conveying to me his passion for science and for helping me to become an independent thinker.

To all the members of my thesis committee for their valuable help and guidance throughout my graduate career. I will always be grateful to all of them for having given me not only their support when I needed it, but also their professional advice and for sharing their extensive and diverse scientific expertise.

To my mother Amanda and my father Tito who always guided me through life and taught me to fight for my dreams, no matter how daunting the challenges were.

To my sisters Alejandrina and Mariana and my nephew Alejandro whose support has always been inspirational during hard times.

To my dear wife and friend Lauren for her unconditional love and for her always candid and valuable advice.

To my parents in law Bruce and Carmen for their loving care and genuine interest in helping me to succeed in each of my endeavors.

To my dear friends Katy Curry and Stephen Klusza for their priceless friendship, for their continuous support, and their immeasurable kindness and care.

TABLE OF CONTENTS

LIST OF FIGURES……..…..xi

LIST OF ABBREVIATIONS……..…….xiii

Chapter

I. INTRODUCTION…….....…...1

Histone locus body: A versatile model to study nuclear body assembly…….....4

The link between histone mRNA biosynthesis and HLB assembly……..….5

The cell cycle tightly regulates replication-dependent histone gene expression…….…..…..8

Nuclear bodies: Dynamic and complex molecular assemblages……...11

Dissertation goals……...15

References……..…17

II. DISTINCT SELF-INTERACTION DOMAINS PROMOTE MULTI SEX COMBS ACCUMULATION IN AND FORMATION OF THE *DROSOPHILA* HISTONE LOCUS BODY.....24

Overview……....25

Introduction……....26

Results……..…….29

Two NH2-terminal domains are required for Mxc concentration in HLBs……...….29

The Mxc N-terminus promotes Mxc self-interation……..…..35

Multiple Mxc domains including the N-terminal self-interaction domains are required for HLB formation in vivo and for the completion of development……..….39

The C-terminus of Mxc recruits HLB components required for histone mRNA synthesis……...……..42

Both the LisH and SIF domains are required for efficient accumulation of Mxc in the HLB…….....49

Discussions……...52

Multiple domains of Mxc are required for HLB assembly...……..…….52

Self-interaction between different Mxc molecules is required for HLB assembly…….....53

Mxc's requirement for histone mRNA biosynthesis correlates with HLB assembly……..…...57

Conclusions……..….60

Materials and Methods……..….60

Immunofluorescence……..….60

Amylose Pull-down assay……..…….61

Imaging……..…..62

Embryo *in situ* hybridization……...…….63

Molecular biology……..…….64

Western Blotting……..…...64

Northern Blotting…….....65

Statistical Analysis……..….65

Computational Analysis of Mxc's self-interaction…….....65

References…….....66

III. EXPLORING THE ROLE OF CYCE/CDK2-DEPENDENT MXC PHOSPHORYLATION IN HLB ASSEMBLY AND HISTONE mRNA BIOSYNTHESIS……..……...73

Introduction……..…….73

Identifying Mxc amino acid(s) subject to CycE/Cdk2-dependent phosphorylation…….....75

Role of Mxc's phosphorylation in HLB assembly and histone mRNA biosynthesis…….....87

Discussion and future work……..….89

Materials and Methods……..…….93

Drosophila strains……...…….93

Immunofluorescence……..…….93

Imaging……..…….94

Molecular biology……..…….94

Northern Blotting……..…...95

S1 Nuclease Protection Assay……..……95

References……..…96

IV. DISCUSSION AND FUTURE DIRECTIONS…….....100

HLB assembly is dynamic and it involves multiple protein-protein interactions……..…101

HLB assembly conforms to the “seed and grow” model……..…..106

HLB behavior and histone mRNA biosynthesis during *Drosophila* embryogenesis……..….110

Can phase transition mediate HLB assembly?.....113

Phase separation under normal and pathological conditions……...….122

A fully assembled HLB is essential for histone mRNA biosynthesis efficiency……..…..125

References…….....127

LIST OF FIGURES

Figure

1.1 Schematic representation of a nucleus displaying a diverse group of known nuclear bodies.…….……...…..2

1.2 Schematic representation of a nuclear body (NB) concentrating factors to facilitate faster reactions…….....12

2.1 Mxc structure/function analysis……...…...31

2.2 The Mxc N-terminus is required for concentration in the HLB in cultures cells……..….33

2.3 Mxc self-interaction requires teo N-terminal domains……..36

2.4 Mxc requires the LisH and SIF domains to concentrate at the histone locus in transgenic flies……..….40

2.5 The 721-1481 regionof Mxc is necessary for HLB assembly and completion of development…….....43

2.6 Mxc concentration in the HLB is required for histone mRNA biosynthesis…….....45

2.7 The C-terminus of Mxc is required to concentrate essential histone mRNA processing factors at the histone locus……..…48

2.8 Mxc's LisH and SIF domains promote HLB localization…...…...51

2.9 LisH domain/SIF domain self-interaction model……...……56

2.10 HLB assembly model……...….58

3.1 HLB accumulation and MPM-2 reactivity of Mxc's CycE/Cdk2 phosphorylation mutants……..…….77

3.2 Both GFP-Mxc^{AP12} and GFP-Mxc^{AP22} localize at the histone locus……..…...79

3.3 Both GFP-Mxc^{AP12} and GFP-Mxc^{AP22} co-localize

at the histone locus with FLASH…….....81

3.4 Both GFP-Mxc^{AP12} and GFP-Mxc^{AP22} display MPM-2 foci……..….82

3.5 CycE/Cdk2 phosphoepitopes affect Mxc's accumulation at the histone locus…….....84

3.6 Both GFP-Mxc^{AP12} and GFP-Mxc^{AP22} display defective histone mRNA biosynthesis…….....88

4.1 HLB assembly is dynamic and requires an Mxc scaffold and multiple protein-protein interactions…….....102

4.2 Phase separation-dependent HLB assembly model…….....121

LIST OF ABBREVIATIONS

CB	Cajal body
Cdk	Cyclin-dependent kinase
DAPI	4',6-diamidino-2-phenylindole
DNA	deoxyribonucleic acid
FISH	fluorescent *in situ* hybridization
FLASH	FLICE-associated huge protein
GFP	green fluorescent protein
HLB	Histone locus body
Lsm	Sm-like
MPM-2	Mitotic protein monoclonal 2
Mute	Muscle wasted
Mxc	Multi sex combs
NPAT	Nuclear Protein Ataxia-Telangiectasia locus
RNA	ribonucleic acid
SDS-PAGE	sodium dodecyl sulfate polyacrylamide gel electrophoresis
S2	Schneider 2
SLBP	stem loop binding protein
snRNP	small nuclear ribonucleoprotein
WT	wild type

CHAPTER I

INTRODUCTION

The regulation of distinct gene expression programs requires the precise spatial and temporal orchestration of an abundance of regulatory components including DNA, RNA, protein, and ribonucleoproteins. It is now widely accepted that the eukaryotic nucleus is a complex and highly dynamic organelle housing on average 2 meters of genomic DNA (Lavelle, 2014). Although the physical condensation of the genomic DNA into chromatin plays an integral role in the faithful regulation of the expression of its own genes, the genome, in addition, employs a plethora of structural components giving rise to specialized domains (Nunez et al., 2009). These nuclear membrane-less compartments concentrate multiple factors required for conducting specific biological functions inside the nucleus and are collectively known as nuclear bodies (NBs). NBs can be broadly defined as morphologically distinct regions that can be differentiated from their surroundings and are visible by light microscopy (Matera et al., 2009; Morimoto and Boerkoel, 2013). Nuclear bodies include the nucleolus, the nuclear speckle, the paraspeckle, the cajal body (CB), the promyelocytic leukemia body (PML), and the histone locus body (HLB) among others (Figure 1.1). Discoveries made throughout the past 15 years have greatly contributed to the understanding of NB organization. However, the molecular mechanisms underlying NB assembly, the dynamics of their constituents,

and therefore their functions are not well understood (Sleeman and Trinkle-Mulcahy, 2014). NBs are known to be important for gene expression regulation and therefore essential for maintaining cellular homeostasis. It is not surprising that NB dysfunction has been linked to a myriad of diseases, including distinct types of cancer and degenerative disorders (Sleeman and Trinkle-Mulcahy, 2014).

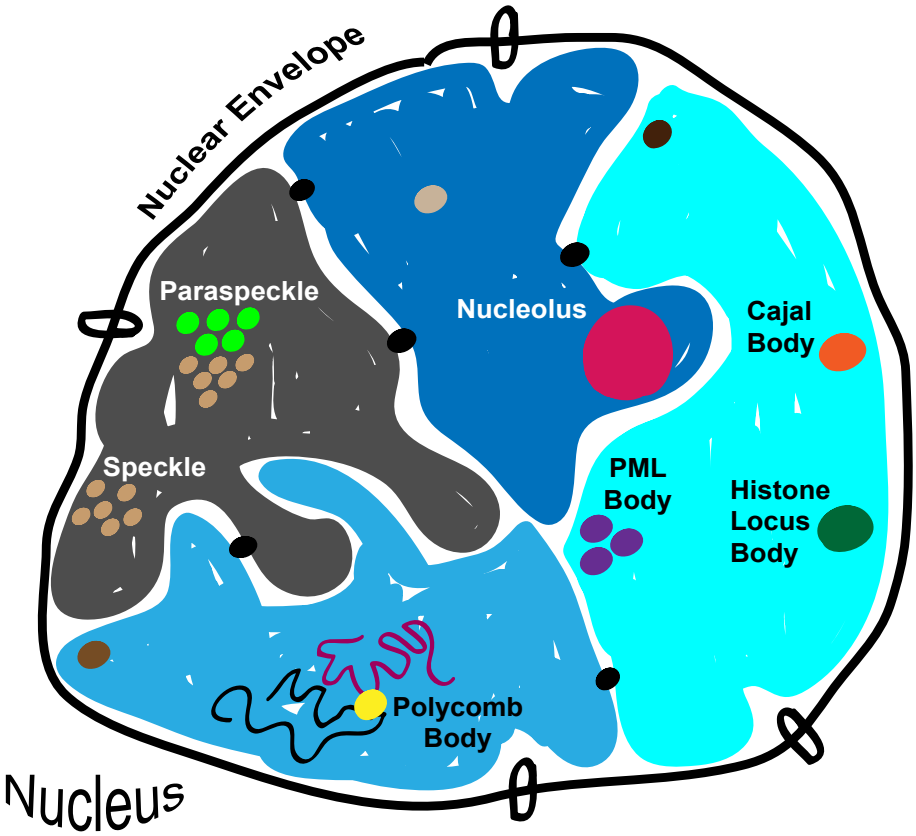


Figure 1.1. Schematic representation of a nucleus displaying a diverse group of known nuclear bodies. Nucleoli are constituted by a myriad of proteins (approximately 700 catalogued to date), most of which are involved in ribosomal RNA (rRNA) transcription and pre-mRNA processing (Morimoto and Boerkoel, 2013). Other proteins are involved in chromatin structure and messenger RNA (mRNA) metabolism as well as different RNA species, such as small nucleolar RNAS (snoRNAs). **Cajal Bodies**

(CBs) are known to associate with histone, small nuclear RNA (snRNA), and snoRNA genes in various human cancer cell lines (Matera et al., 2009). CBs are multifunctional as they serve as sites for the modification of snRNAs and snoRNAs, for the assembly and trafficking of ribonucleoproteins (RNPs), and have also been shown to be involved in telomere length regulation. **Histone locus bodies** (HLBs) are known to play an important role in the regulation of histone mRNA biosynthesis and are extensively discussed throughout the entire dissertation. **Promyelocytic leukemia** (PML) bodies are mobile nuclear compartments shown to be implicated in diverse cellular functions including apoptosis, senescence, tumor suppression, transcription, antiviral response and DNA replication and repair (Chung et al., 2011). **Polycomb Group** (PcG) bodies are the result of aggregation of some polycomb group proteins, which are known chromatin remodelers. PcG bodies tend to randomly disperse inside the nucleus and as a general rule they serve as structures in which different polycomb targets are clustered and co-repressed (Del Prete et al., 2015). **Speckles** possess many factors that are components of the splicesome and are involved in pre-mRNA splicing. For instance, members of the serine/arginine-rich splicing factor (SRSF) family are found in speckles, which participate in constitutive and alternative splicing, transcription, mRNA translation and genome stability (Morimoto and Boerkoel, 2013). **Paraspeckles** are known to participate in nuclear retention of long adenosine-to-inosine hyperedited RNAs and in the storage and rapid release of certain RNAs under stress conditions (Shelkovnikova et al., 2014).

Therefore, there is a great impetus for gaining understanding on how NBs assemble and consequently control gene expression regulation in response to diverse physiological inputs under both normal and adverse circumstances. My dissertation research employs the *Drosophila* HLB as a model of NB assembly and function to advance our understanding of how NBs contribute to gene expression regulation and ultimately to cellular homeostasis.

In the following sections of this chapter I will describe the features that make the HLB a versatile paradigm to study NB formation and function. Then, I will discuss the

intricate relationship between histone mRNA biosynthesis and HLB assembly and I will describe the findings strongly arguing in favor of such link between these two phenomena. Next, I will talk about the tight cell cycle-dependent regulation of histone gene expression and some findings that strongly support the notion of a major role for the cell cycle in histone mRNA biosynthesis. Lastly, I will discuss the dynamic nature of NBs, the current model defining a NB and its purpose and an emergent hypothesis postulating that phase separation plays a fundamental role in NB assembly. I finish this last sub-section with a few outstanding questions in the field regarding the biophysical forces dictating NB formation.

Histone locus body: A versatile model to study nuclear body assembly

Nuclear bodies enriched with factors necessary for efficient histone gene expression, such as U7 snRNP, were dubbed HLBs in a seminal study conducted in *Drosophila* melanogaster (Liu et al., 2006). HLBs are known to associate invariably with the replication-dependent histone genes, which are encoded in a single 5-kb repeat present in approximately 100 tandem copies in *Drosophila* (Liu et al., 2006; White et al., 2007). HLBs have also been shown to form in mammalian cells. Studies employing human embryonic stem (hES) cells demonstrated that NPAT, a transcription factor required to activate histone gene expression, co-localizes at the two major histone gene clusters on 6p22 and 1q21 with histone gene expression-specific factors such as FLASH and U7 snRNP (Ghule et al., 2008).

The HLB is a versatile NB assembly and function model as it exclusively concentrates all factors required for histone mRNA biosynthesis at a discrete nuclear location, the histone locus. In *Drosophila*, the tandem repeats of DNA replication-

dependent histone genes are localized in only one locus on the left arm of chromosome 2. This feature notably simplifies qualitative assays and, with the advent of powerful super-resolution microscopy techniques, will soon be true for quantitative experiments, too. With respect to functional studies, the cell cycle is a known critical input regulating HLB function, as it limits histone gene expression to S phase, when DNA is replicated. By virtue of these attributes, the HLB is a unique and powerful paradigm that will advance our understanding of NB formation and function. By utilizing the HLB as a model, we will also be able to gain a better understanding of the mechanisms linking gene expression and NB formation by further exploring the complex molecular aspects underlying the connection between histone mRNA biosynthesis and HLB assembly.

The link between histone mRNA biosynthesis and HLB assembly

Replication-dependent histone genes encode the four core histones, namely H2A, H2B, H3, and H4, which are the building blocks of the nucleosome and also the linker H1 histone, which sits on top of each nucleosome keeping in place the DNA that wraps around them (Marzluff et al., 2008). Histone genes organize in clusters in all organisms. For instance, mammals display two big clusters of histone genes. In humans, a large cluster localizes on chromosome 6 (chromosome 13 in mice), which contains more than 80% of the genes, including six histone H1 genes. A smaller histone gene cluster is found on human chromosome 1 (chromosome 3 in mice) that contains the four core histone genes above mentioned. In *Drosophila* and *Xenopus*, the four core histone genes and the histone linker H1 gene are all clustered in a tandem array (Perry et al., 1985).

In metazoans, the mRNAs that code for histone proteins are structurally unique as they are the only mRNAs that do not end with a polyA tail in their 3' end. Instead, histone mRNAs display a stem-loop whose sequence has been shown to be highly conserved in evolution (Dominski and Marzluff, 1999). The stem-loop interacts with a unique set of proteins known as the stem-loop binding proteins (SLBPs), which are known to be responsible for the bulk of the regulation of histone mRNA (Marzluff and Duronio, 2002). Another unique set of proteins required for regulating the expression of histone mRNAs includes the U7 small nuclear RNA (snRNA) and the SM-like proteins LSm10, LSm11, and the SLBP-interacting protein 1 (SLIP1). Human protein NPAT (Nuclear Protein mapped to the Ataxia-Telangiectasia locus) is constitutively associated with the histone genes throughout the cell cycle and is known to be essential for entry into S phase (Ye et al., 2003) and to possess an N-terminal LisH-like domain required to stimulate histone mRNA transcription (Ma et al., 2000; Wei et al., 2003). NPAT is phosphorylated by CycE/Cdk2 at the onset of S phase and this modified version of NPAT persists throughout S phase at the histone locus, which results in an increased expression of the canonical histone genes (Marzluff et al., 2008).

Metazoan replication-dependent histone genes lack introns. Therefore, the only processing step required to form a mature histone mRNA is an endonucleolytic cleavage that releases the nascent pre-mRNA from DNA template. Insertion of an intron into a histone gene has been shown to disrupt the normal histone mRNA processing, which elicited the formation of polyadenylated histone mRNA (Pandey et al., 1990). The endonucleolytic cleavage occurs between the stem-loop and the histone downstream element (HDE), a purine-rich sequence found approximately 15 nuleotides downstream

of the cleavage site (Dominski et al., 2005; Ryan et al., 2004; Scharl and Steitz, 1994). The first step in the formation of the histone mRNA 3' end is the binding of the SLBP protein to the stem-loop, which is followed by the binding of the HDE sequence with the U7 snRNA (a component of the U7 small nuclear ribonucleoprotein, snRNP, complex). Another factor required for the formation of the 3' end of histone mRNAs is the protein FLASH (Yang et al., 2009). FLASH together with the U7 snRNP complex are continuously present at the histone locus and their recruitment to this site does not depend on *cis* elements present on the pre-mRNA with which they associate (Liu et al., 2006; Salzler et al., 2013; White et al., 2011; White et al., 2007). These findings indicate that both histone processing factors, FLASH and U7 snRNP, must be recruited directly to the histone locus (Salzler et al., 2013). FLASH in concert with the U7 snRNP complex recruits a cleavage complex known as the Histone Cleavage complex (HCC) (Yang et al., 2013). The direct interaction between the N-termini of FLASH and the U7 snRNP component Lsm11 facilitates the association of the HCC with U7 snRNP (Sabath et al., 2013; Yang et al., 2013). It is therefore intriguing that U7 snRNP and FLASH exclusively recruit the HCC at the HLB during S phase despite their continuous localization at the histone locus.

In *Drosophila*, work done in our laboratory demonstrated that a sequence of approximately 300 nucleotides in length spanning the bidirectional H3-H4 promoters contains the information that is necessary and sufficient to form an HLB. Interestingly, the H2a-H2b and H1 promoters, and other regions of the histone repeat, cannot. Although the 300-nt sequence serves as a scaffold for HLB assembly, transcription activation must occur in order for the HLB to be fully formed (mature HLB) (Salzler et al., 2013). To

date, neither the precise role of histone mRNA biosynthesis in HLB formation nor the true function of the HLB are well understood. However, recent evidence supports the notion that the HLB's main function is to increase the local concentration of factors required to facilitate the biochemical reactions involved in histone mRNA biosynthesis (see **Chapter II** for a more detailed description of our recent findings) (Tatomer et al., **In revision** [I am second author on this manuscript and I am currently conducting experiments to address revisions]). In conclusion, it is becoming increasingly clear that the HLB requires a combination of *cis* elements at the histone locus in order to signal for the accumulation of specific factors necessary for concentrating other HLB components required for the initiation of histone gene expression, which further underscores the intricate relationship between histone mRNA biosynthesis and HLB assembly. Elucidating the complex link between histone gene expression and HLB assembly will be important to further interrogate if and how gene expression and NB formation are mutually orchestrated and if this reciprocal control is a global phenomenon observed in other NBs. In addition, as replication-dependent histone genes are cell cycle regulated, understanding the molecular mechanisms connecting the cell cycle to histone gene expression will further shed light on the intricate interconnection between histone mRNA biosynthesis and HLB formation.

The cell cycle tightly regulates replication-dependent histone gene expression

Replication-dependent histone genes are tightly regulated by the cell cycle and they must reach high levels of expression during S phase. It has been well documented that canonical histone mRNA transcription is increased by 35-fold at the G1/S transition

in mammalian cell lines (Borun et al., 1975; Breindl and Gallwitz, 1973; DeLisle et al., 1983; Detke et al., 1979; Harris et al., 1991; Heintz et al., 1983; Parker and Fitschen, 1980). The proper rate of the rapid accumulation of all five canonical histone mRNAs is controlled by transcription initiation rate and pre-mRNA processing as cells enter S phase (Marzluff et al., 2008). Histone genes must be highly regulated as both lower and higher than normal protein levels have pervasive effects on cellular homeostasis. For instance, the scarcity of histones has been known for years to cause cell-cycle arrest affecting viability in yeast (Han et al., 1987). On the other hand, excess of histones has also been shown to result in stalling of the replication fork (Herrero and Moreno, 2011), cytotoxicity mainly dependent on inappropriate electrostatic interactions between positively charged histones and different negatively charged molecules such as DNA and RNA (Singh et al., 2010), and persistent DNA damage response (Landais et al., 2014).

Due to the presence of the stem loop, histone mRNAs require unique transcription and processing mechanisms. In addition, the HLB is known to concentrate several factors required for histone mRNA biosynthesis and for its own assembly (Ghule et al., 2008; Liu et al., 2006; Rajendra et al., 2011; Salzler et al., 2013; White et al., 2011; White et al., 2007), a phenomenon that involves numerous protein-protein interactions (Burch et al., 2011; Yang et al., 2014). However, how these distinct steps of histone mRNA biosynthesis and HLB assembly are connected to the molecular mechanisms orchestrating cell cycle progression through the activity of cyclin-dependent kinases (Cdks) is not completely understood.

Cdks are known to contain a serine/threonine-specific catalytic core and to physically associate with their regulatory subunits, cyclins, which control the kinase

activity and substrate specificity (Lim and Kaldis, 2013). Cyclins and Cdks together form heterodimer (Cyc/Cdk) complexes that are known to be the core engines driving the eukaryotic cell cycle (Budirahardja and Gonczy, 2009). Cyc/Cdk complexes were first implicated in cell cycle control on pioneering work performed in yeast, where a single Cdk (Cdc28 in budding yeast and Cdc2 in fission yeast) was found to promote transitions between different cell cycle phases by interacting with numerous cyclins (Beach et al., 1982; Nurse and Thuriaux, 1980; Nurse et al., 1976).

The CycE/Cdk2 complex is known to promote the G1 to S transition in part by phosphorylating proteins that mediate changes in the expression of genes associated with the onset of DNA replication. One of these CycE/Cdk2 targets is the human protein NPAT, shown to physically interact with the CycE/Cdk2 complex, to co-localize with CycE at the histone locus, and to be essential for activation of histone gene expression (Ma et al., 2000; Zhao et al., 1998; Zhao et al., 2000). We (White et al., 2011) and others (Rajendra et al., 2011) have identified *Drosophila* multi sex combs (Mxc) as the functional equivalent of human NPAT.

In 1998, by staining follicle cell nuclei with the MPM-2 antibody Calvi and colleagues observed a "subnuclear sphere of unknown identity". These spheres lit up by the MPM-2 antibody coincided with the highest peak in the activity of CycE/Cdk2 during S phase (Calvi et al., 1998). Work from our laboratory demonstrated that MPM-2, widely used during the past 15 years to detect CycE/Cdk2 targets in *Drosophila*, recognizes a target co-localizing with known HLB factors and later on that Mxc is a CycE/Cdk2 phosphorylation target concentrated at the histone locus, as it directly reacts with MPM-2 (White et al., 2011; White et al., 2007). Although it is speculated that cell cycle-

dependent activation of NPAT (Ma et al., 2000; Pirngruber and Johnsen, 2010) and Mxc (White et al., 2011) is crucial for histone gene transcription, the specific role of Mxc's CycE/Cdk2-dependent phosphorylation in HLB assembly and in histone gene expression remains unknown.

To date, several research groups have shown that post-translational modifications (PTMs) play a crucial role in the assembly and disassembly of nuclear and cytoplasmic bodies (Kwon et al., 2013; Kwon et al., 2014; Wang et al., 2014; Zhu and Brangwynne, 2015). Therefore, taking into consideration the crucial role that Mxc plays in HLB assembly and histone mRNA biosynthesis and that Mxc is a target for CycE/Cdk2 phosphorylation, we set out to explore how Mxc's CycE/Cdk2 phosphoepitopes contribute to these two biological processes. In chapter III of this dissertation, I describe in detail the work that I am currently conducting and I also discuss preliminary data strongly indicating that Mxc harbors multiple CycE/Cdk2 phosphorylation sites and that they play a crucial role in Mxc's ability to normally accumulate at the histone locus and to assemble a stable HLB. Our findings support the notion that PTMs, such as phosphorylation, greatly contribute to the mechanisms underlying the dynamic NB assembly/disassembly phenomenon.

Nuclear bodies: Dynamic and complex molecular assemblages

In the field of nuclear organization, there is a current notion that states that NBs can be considered molecular depots (Figure 1.2). This notion postulates that a particular time during cell cycle substrates, reactive proteins (e.g., scaffolding proteins), and enzymes are diffused throughout the nucleoplasm. Due to a determined signal, all of these factors

concentrate in a much more reduced nuclear space giving rise to what we know as a NB. It is thought that the sole purpose of this collection of molecules is to facilitate a faster and therefore more efficient reaction (Dundr, 2012).

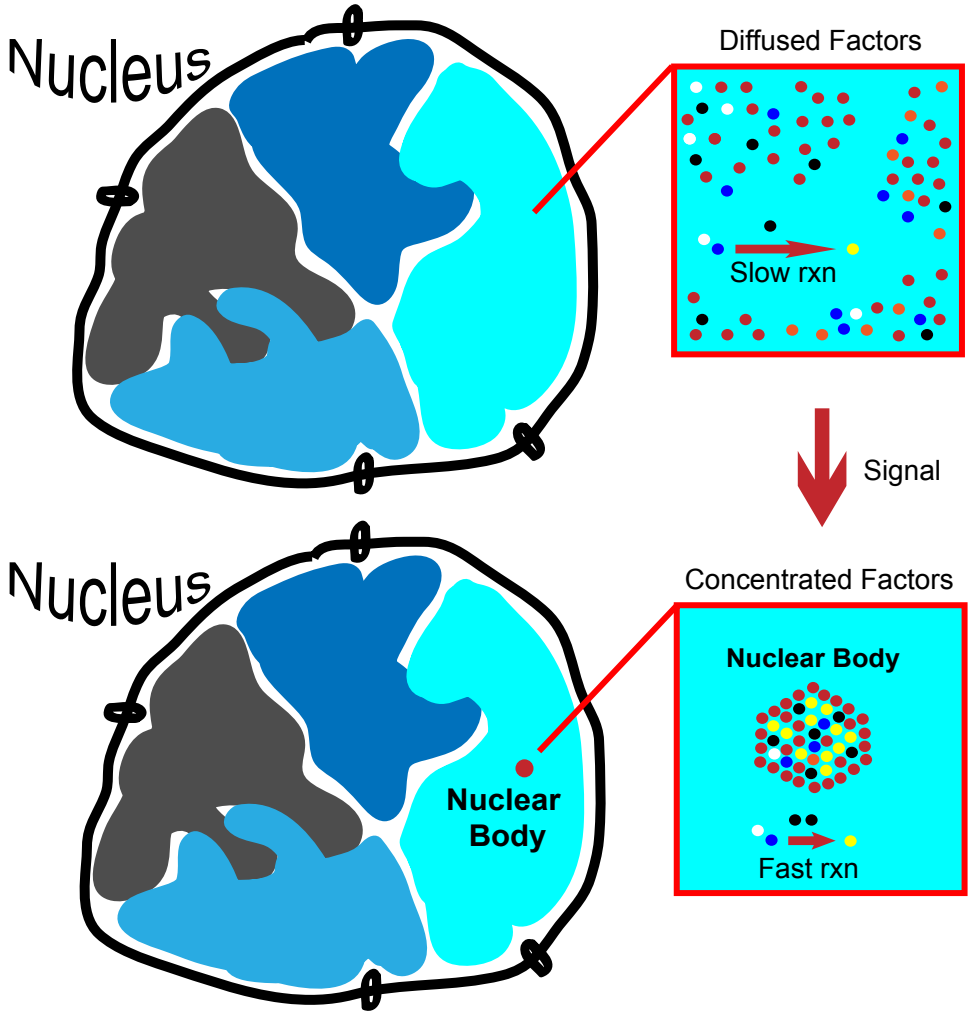


Figure 1.2. Schematic representation of a nuclear body (NB) concentrating factors to facilitate faster reactions. Diffused nuclear components are represented by full circles. Blue and white circles represent substrates of either a reactive protein (e.g., scaffolding proteins that can accelerate reactions by physical contact with other proteins) or enzymes (black circles). As a result of a determined signal, these

components concentrate in a more reduced space and give rise to what we know as a nuclear body (NB), where a faster reaction, represented by a bigger number of yellow circles takes place.

A new NB assembly hypothesis that is increasingly gaining attention in the field of nuclear organization postulates that both nuclear and cytoplasmic bodies concentrate in focal locations via intracellular phase separation (Brangwynne, 2013; Brangwynne et al., 2009; Hyman et al., 2014; Li et al., 2012; Zhu and Brangwynne, 2015). Phase transitions are common in nature and are the consequence of thermodynamic forces driving a particular system to equilibrium (Zhu and Brangwynne, 2015). *In vitro* crystallization of proteins is a well-known phenomenon in biology where proteins that are soluble have the tendency to condense into more concentrated liquid (gel-like) phases or even solid phases. It is therefore thought that this phenomenon could occur in the cell where proteins transition from a diffused, and less concentrated, nucleoplasmic state into a localized and consequently more concentrated body and vice versa.

The same NB assembly hypothesis also postulates that protein-protein and protein-RNA interactions that give rise to assemblages, such as NBs, are multivalent and dynamic and are mediated by intrinsically disordered regions (IDRs) of low-complexity and/or multivalent folded protein domains (Toretsky and Wright, 2014). The dynamic nature of these interactions is ingrained in the capacity of these multivalent proteins to homotypically or heterotypically associate and to also be able to rapidly dissociate through weak physical interactions. Accruing evidence supports this NB assembly hypothesis by demonstrating that proteins harboring multiple protein-protein or protein-RNA interaction domains (multivalent proteins) facilitate the local concentration of factors that in turn assemble into NBs (Chen et al., 1999; Good et al., 2011; Ishov et al.,

1999; Matera et al., 2009; Zaidi et al., 2007). For instance, human coilin's N-terminal domain mediates a self-interaction required for coilin to concentrate in the CB (Hebert and Matera, 2000). Another example is the work done on nuclear paraspeckle assembly. Here, the authors perform both *in vitro* and *in vivo* approaches to demonstrate that the human RBM14, an RNA-binding protein, utilizes its prion-like domain (PLD), a known low-complexity region, to connect with another PLD-containing protein, such as FUS, to assemble a paraspeckle. Moreover, the authors demonstrate that RBM14, as previously shown for fused in sarcoma (FUS) (Han et al., 2012; Kato et al., 2012), requires its PLD region to form hydrogels *in vitro*, a liquid-phase transition potential (Hennig et al., 2015). Although important advances have been made in the molecular and biophysic aspects of NB formation, great challenges lie ahead of gaining a more complete understanding of how NBs assemble. For instance, can phase separation be a common phenomenon driving the assembly of all NBs? Or, do all NBs assemble through high concentrations of multivalent proteins whose physical interactions must be weak in order for them to be reversible and therefore dynamic?

In the present dissertation, I describe our recently published findings demonstrating that Mxc, a large protein containing multiple folded domains, plays an essential role in HLB assembly and in concentrating histone mRNA biosynthetic factors at the histone locus. Moreover, Mxc's accumulation and therefore HLB assembly and enlargement during early cell cycles of *Drosophila* embryogenesis are highly dynamic, as evidenced by live imaging experiments. In addition, I mention a number of low-complexity (unfolded) domains that I have found on Mxc by performing software analysis and I discuss the possibility that these disordered domains may play a role in

Mxc's function at the HLB, that is, the formation of the three-dimensional scaffold. Therefore, these findings raise the outstanding question of whether Mxc employs both folded and unfolded domains to assemble a molecular scaffold, the HLB, by increasing its own concentration and that of all of the required HLB factors at the histone locus via phase separation.

Dissertation goals

This dissertation describes the work done exploring how HLB assembly occurs and the importance of this phenomenon in maintaining the efficiency of histone mRNA biosynthesis. More specifically, this dissertation is focused on the molecular mechanisms underlying the dynamic nature of HLB assembly. Chapter II contains data demonstrating that *Drosophila* Mxc plays a crucial role in HLB assembly by utilizing multiple domains to accumulate at the histone locus and to concentrate histone mRNA biosynthetic factors, such as FLASH. Notably, we find that Mxc harbors two N-terminal domains involved in self-interaction: the LisH domain and the novel SIF (Self-Interaction Facilitator) domain. Both LisH and SIF domains are required for Mxc's effective concentration at the histone locus, which we show that is first observed during cell cycle 9, and consequently for the enlargement of the HLB throughout the subsequent cell cycles of *Drosophila* embryogenesis.

Our laboratory previously demonstrated that Mxc is directly recognized by MPM-2 antibody, strongly suggesting that Mxc is a CycE/Cdk2 phosphorylation target (White et al., 2011). Chapter III describes the data generated and research currently conducted to explore the role of this cell cycle input in HLB assembly and histone mRNA

biosynthesis. We have generated a large number of Mxc phosphorylation mutants in which the serines or threonines preceding prolines were changed to alanines to begin the identification of CycE/Cdk2 phosphoepitopes on Mxc. Immunofluorescence and biochemical data demonstrate that full-length Mxc lacking a subset of putative CycE/Cdk2 phosphorylation sites (GFP-Mxc^{AP22}) can still co-localize at the histone locus with other known HLB factors. However, GFP-Mxc^{AP22} consistently forms a higher number of GFP, Mxc, and MPM-2 positive foci and the latter marker is dimmer than those observed for GFP-Mxc. Notably, when segregating in a null *mxc* background, GFP-Mxc^{AP22} generates sterile female flies. Based on these data, we hypothesize that Mxc harbors multiple CycE/Cdk2 sites and that these are required for Mxc's effective accumulation at the histone locus and to assemble a stable, mature HLB. Chapter IV takes an in-depth look into the relevance of the findings discussed in this dissertation and addresses fundamental questions highlighted by our research and the work conducted by our laboratory and others. I also introduce, and extensively describe, the concept of phase separation, which is proposed, by the work of several research groups, to be the biophysical force governing NB formation. Chapter IV also discusses the relevance of understanding NB formation and function in a normal developmental context and under aberrant physiological conditions and describes analyses and technologies that could be implemented to further explore the biophysical mechanisms underlying NB formation.

REFERENCES

Beach, D., B. Durkacz, and P. Nurse. 1982. Functionally homologous cell cycle control genes in budding and fission yeast. *Nature*. 300:706-709.

Borun, T.W., F. Gabrielli, K. Ajiro, A. Zweidler, and C. Baglioni. 1975. Further evidence of transcriptional and translational control of histone messenger RNA during the HeLa S3 cycle. *Cell*. 4:59-67.

Brangwynne, C.P. 2013. Phase transitions and size scaling of membrane-less organelles. *J Cell Biol*. 203:875-881.

Brangwynne, C.P., C.R. Eckmann, D.S. Courson, A. Rybarska, C. Hoege, J. Gharakhani, F. Julicher, and A.A. Hyman. 2009. Germline P granules are liquid droplets that localize by controlled dissolution/condensation. *Science*. 324:1729-1732.

Breindl, M., and D. Gallwitz. 1973. Identification of histone messenger RNA from HeLa cells. Appearance of histone mRNA in the cytoplasm and its translation in a rabbit-reticulocyte cell-free system. *Eur J Biochem*. 32:381-391.

Budirahardja, Y., and P. Gonczy. 2009. Coupling the cell cycle to development. *Development*. 136:2861-2872.

Burch, B.D., A.C. Godfrey, P.Y. Gasdaska, H.R. Salzler, R.J. Duronio, W.F. Marzluff, and Z. Dominski. 2011. Interaction between FLASH and Lsm11 is essential for histone pre-mRNA processing in vivo in Drosophila. *RNA*. 17:1132-1147.

Calvi, B.R., M.A. Lilly, and A.C. Spradling. 1998. Cell cycle control of chorion gene amplification. *Genes Dev*. 12:734-744.

Chen, T., F.M. Boisvert, D.P. Bazett-Jones, and S. Richard. 1999. A role for the GSG domain in localizing Sam68 to novel nuclear structures in cancer cell lines. *Mol Biol Cell*. 10:3015-3033.

Chung, I., H. Leonhardt, and K. Rippe. 2011. De novo assembly of a PML nuclear subcompartment occurs through multiple pathways and induces telomere elongation. *J Cell Sci*. 124:3603-3618.

Del Prete, S., P. Mikulski, D. Schubert, and V. Gaudin. 2015. One, Two, Three: Polycomb Proteins Hit All Dimensions of Gene Regulation. *Genes (Basel)*. 6:520-542.

DeLisle, A.J., R.A. Graves, W.F. Marzluff, and L.F. Johnson. 1983. Regulation of histone mRNA production and stability in serum-stimulated mouse 3T6 fibroblasts. *Mol Cell Biol*. 3:1920-1929.

Detke, S., A. Lichtler, I. Phillips, J. Stein, and G. Stein. 1979. Reassessment of histone gene expression during cell cycle in human cells by using homologous H4 histone cDNA. *Proc Natl Acad Sci U S A*. 76:4995-4999.

Dominski, Z., and W.F. Marzluff. 1999. Formation of the 3' end of histone mRNA. *Gene*. 239:1-14.

Dominski, Z., X.C. Yang, and W.F. Marzluff. 2005. The polyadenylation factor CPSF-73 is involved in histone-pre-mRNA processing. *Cell*. 123:37-48.

Dundr, M. 2012. Nuclear bodies: multifunctional companions of the genome. *Curr Opin Cell Biol*. 24:415-422.

Ghule, P.N., Z. Dominski, X.C. Yang, W.F. Marzluff, K.A. Becker, J.W. Harper, J.B. Lian, J.L. Stein, A.J. van Wijnen, and G.S. Stein. 2008. Staged assembly of histone gene expression machinery at subnuclear foci in the abbreviated cell cycle of human embryonic stem cells. *Proc Natl Acad Sci U S A*. 105:16964-16969.

Good, M.C., J.G. Zalatan, and W.A. Lim. 2011. Scaffold proteins: hubs for controlling the flow of cellular information. *Science*. 332:680-686.

Han, M., M. Chang, U.J. Kim, and M. Grunstein. 1987. Histone H2B repression causes cell-cycle-specific arrest in yeast: effects on chromosomal segregation, replication, and transcription. *Cell*. 48:589-597.

Han, T.W., M. Kato, S. Xie, L.C. Wu, H. Mirzaei, J. Pei, M. Chen, Y. Xie, J. Allen, G. Xiao, and S.L. McKnight. 2012. Cell-free formation of RNA granules: bound RNAs identify features and components of cellular assemblies. *Cell*. 149:768-779.

Harris, M.E., R. Bohni, M.H. Schneiderman, L. Ramamurthy, D. Schumperli, and W.F. Marzluff. 1991. Regulation of histone mRNA in the unperturbed cell cycle: evidence suggesting control at two posttranscriptional steps. *Mol Cell Biol*. 11:2416-2424.

Hebert, M.D., and A.G. Matera. 2000. Self-association of coilin reveals a common theme in nuclear body localization. *Mol Biol Cell*. 11:4159-4171.

Heintz, N., H.L. Sive, and R.G. Roeder. 1983. Regulation of human histone gene expression: kinetics of accumulation and changes in the rate of synthesis and in the half-lives of individual histone mRNAs during the HeLa cell cycle. *Mol Cell Biol*. 3:539-550.

Hennig, S., G. Kong, T. Mannen, A. Sadowska, S. Kobelke, A. Blythe, G.J. Knott, K.S. Iyer, D. Ho, E.A. Newcombe, K. Hosoki, N. Goshima, T. Kawaguchi, D. Hatters, L. Trinkle-Mulcahy, T. Hirose, C.S. Bond, and A.H. Fox. 2015. Prion-like domains in RNA binding proteins are essential for building subnuclear paraspeckles. *J Cell Biol*. 210:529-539.

Herrero, A.B., and S. Moreno. 2011. Lsm1 promotes genomic stability by controlling histone mRNA decay. *EMBO J*. 30:2008-2018.

Hyman, A.A., C.A. Weber, and F. Julicher. 2014. Liquid-liquid phase separation in biology. *Annu Rev Cell Dev Biol*. 30:39-58.

Ishov, A.M., A.G. Sotnikov, D. Negorev, O.V. Vladimirova, N. Neff, T. Kamitani, E.T. Yeh, J.F. Strauss, 3rd, and G.G. Maul. 1999. PML is critical for ND10 formation and recruits the PML-interacting protein daxx to this nuclear structure when modified by SUMO-1. *J Cell Biol*. 147:221-234.

Kato, M., T.W. Han, S. Xie, K. Shi, X. Du, L.C. Wu, H. Mirzaei, E.J. Goldsmith, J. Longgood, J. Pei, N.V. Grishin, D.E. Frantz, J.W. Schneider, S. Chen, L. Li, M.R. Sawaya, D. Eisenberg, R. Tycko, and S.L. McKnight. 2012. Cell-free formation of RNA granules: low complexity sequence domains form dynamic fibers within hydrogels. *Cell*. 149:753-767.

Kwon, I., M. Kato, S. Xiang, L. Wu, P. Theodoropoulos, H. Mirzaei, T. Han, S. Xie, J.L. Corden, and S.L. McKnight. 2013. Phosphorylation-regulated binding of RNA polymerase II to fibrous polymers of low-complexity domains. *Cell*. 155:1049-1060.

Kwon, I., S. Xiang, M. Kato, L. Wu, P. Theodoropoulos, T. Wang, J. Kim, J. Yun, Y. Xie, and S.L. McKnight. 2014. Poly-dipeptides encoded by the C9orf72 repeats bind nucleoli, impede RNA biogenesis, and kill cells. *Science*. 345:1139-1145.

Landais, S., C. D'Alterio, and D.L. Jones. 2014. Persistent replicative stress alters polycomb phenotypes and tissue homeostasis in Drosophila melanogaster. *Cell reports*. 7:859-870.

Lavelle, C. 2014. Pack, unpack, bend, twist, pull, push: the physical side of gene expression. *Curr Opin Genet Dev*. 25:74-84.

Li, P., S. Banjade, H.C. Cheng, S. Kim, B. Chen, L. Guo, M. Llaguno, J.V. Hollingsworth, D.S. King, S.F. Banani, P.S. Russo, Q.X. Jiang, B.T. Nixon, and M.K. Rosen. 2012. Phase transitions in the assembly of multivalent signalling proteins. *Nature*. 483:336-340.

Lim, S., and P. Kaldis. 2013. Cdks, cyclins and CKIs: roles beyond cell cycle regulation. *Development*. 140:3079-3093.

Liu, J.L., C. Murphy, M. Buszczak, S. Clatterbuck, R. Goodman, and J.G. Gall. 2006. The Drosophila melanogaster Cajal body. *J Cell Biol*. 172:875-884.

Ma, T., B.A. Van Tine, Y. Wei, M.D. Garrett, D. Nelson, P.D. Adams, J. Wang, J. Qin, L.T. Chow, and J.W. Harper. 2000. Cell cycle-regulated phosphorylation of p220(NPAT) by cyclin E/Cdk2 in Cajal bodies promotes histone gene transcription. *Genes Dev*. 14:2298-2313.

Marzluff, W.F., and R.J. Duronio. 2002. Histone mRNA expression: multiple levels of cell cycle regulation and important developmental consequences. *Curr Opin Cell Biol*. 14:692-699.

Marzluff, W.F., E.J. Wagner, and R.J. Duronio. 2008. Metabolism and regulation of canonical histone mRNAs: life without a poly(A) tail. *Nat Rev Genet*. 9:843-854.

Matera, A.G., M. Izaguire-Sierra, K. Praveen, and T.K. Rajendra. 2009. Nuclear bodies: random aggregates of sticky proteins or crucibles of macromolecular assembly? *Dev Cell*. 17:639-647.

Morimoto, M., and C.F. Boerkoel. 2013. The role of nuclear bodies in gene expression and disease. *Biology (Basel)*. 2:976-1033.

Nunez, E., X.D. Fu, and M.G. Rosenfeld. 2009. Nuclear organization in the 3D space of the nucleus - cause or consequence? *Curr Opin Genet Dev*. 19:424-436.

Nurse, P., and P. Thuriaux. 1980. Regulatory genes controlling mitosis in the fission yeast Schizosaccharomyces pombe. *Genetics*. 96:627-637.

Nurse, P., P. Thuriaux, and K. Nasmyth. 1976. Genetic control of the cell division cycle in the fission yeast Schizosaccharomyces pombe. *Mol Gen Genet*. 146:167-178.

Pandey, N.B., N. Chodchoy, T.J. Liu, and W.F. Marzluff. 1990. Introns in histone genes alter the distribution of 3' ends. *Nucleic acids research*. 18:3161-3170.

Parker, I., and W. Fitschen. 1980. Histone mRNA metabolism during the mouse fibroblast cell cycle. *Cell Differ*. 9:23-30.

Perry, M., G.H. Thomsen, and R.G. Roeder. 1985. Genomic organization and nucleotide sequence of two distinct histone gene clusters from Xenopus laevis. Identification of novel conserved upstream sequence elements. *J Mol Biol*. 185:479-499.

Pirngruber, J., and S.A. Johnsen. 2010. Induced G1 cell-cycle arrest controls replication-dependent histone mRNA 3' end processing through p21, NPAT and CDK9. *Oncogene*.

Rajendra, T.K., K. Praveen, and A.G. Matera. 2011. Genetic Analysis of Nuclear Bodies: From Nondeterministic Chaos to Deterministic Order. *Cold Spring Harb Symp Quant Biol*. in press.

Ryan, K., O. Calvo, and J.L. Manley. 2004. Evidence that polyadenylation factor CPSF-73 is the mRNA 3' processing endonuclease. *RNA*. 10:565-573.

Sabath, I., A. Skrajna, X.C. Yang, M. Dadlez, W.F. Marzluff, and Z. Dominski. 2013. 3'-End processing of histone pre-mRNAs in Drosophila: U7 snRNP is associated with FLASH and polyadenylation factors. *RNA*. 19:1726-1744.

Salzler, H.R., D.C. Tatomer, P.Y. Malek, S.L. McDaniel, A.N. Orlando, W.F. Marzluff, and R.J. Duronio. 2013. A sequence in the Drosophila H3-H4 Promoter triggers histone locus body assembly and biosynthesis of replication-coupled histone mRNAs. *Dev Cell*. 24:623-634.

Scharl, E.C., and J.A. Steitz. 1994. The site of 3' end formation of histone messenger RNA is a fixed distance from the downstream element recognized by the U7 snRNP. *EMBO J*. 13:2432-2440.

Shelkovnikova, T.A., H.K. Robinson, C. Troakes, N. Ninkina, and V.L. Buchman. 2014. Compromised paraspeckle formation as a pathogenic factor in FUSopathies. *Human molecular genetics*. 23:2298-2312.

Singh, R.K., D. Liang, U.R. Gajjalaiahvari, M.H. Kabbaj, J. Paik, and A. Gunjan. 2010. Excess histone levels mediate cytotoxicity via multiple mechanisms. *Cell Cycle*. 9:4236-4244.

Sleeman, J.E., and L. Trinkle-Mulcahy. 2014. Nuclear bodies: new insights into assembly/dynamics and disease relevance. *Curr Opin Cell Biol*. 28:76-83.

Toretsky, J.A., and P.E. Wright. 2014. Assemblages: functional units formed by cellular phase separation. *J Cell Biol*. 206:579-588.

Wang, J.T., J. Smith, B.C. Chen, H. Schmidt, D. Rasoloson, A. Paix, B.G. Lambrus, D. Calidas, E. Betzig, and G. Seydoux. 2014. Regulation of RNA granule dynamics by phosphorylation of serine-rich, intrinsically disordered proteins in C. elegans. *Elife*. 3:e04591.

Wei, Y., J. Jin, and J.W. Harper. 2003. The cyclin E/Cdk2 substrate and Cajal body component p220(NPAT) activates histone transcription through a novel LisH-like domain. *Mol Cell Biol*. 23:3669-3680.

White, A.E., B.D. Burch, X.C. Yang, P.Y. Gasdaska, Z. Dominski, W.F. Marzluff, and R.J. Duronio. 2011. Drosophila histone locus bodies form by hierarchical recruitment of components. *J Cell Biol*. 193:677-694.

White, A.E., M.E. Leslie, B.R. Calvi, W.F. Marzluff, and R.J. Duronio. 2007. Developmental and cell cycle regulation of the Drosophila histone locus body. *Mol Biol Cell*. 18:2491-2502.

Yang, X.C., B.D. Burch, Y. Yan, W.F. Marzluff, and Z. Dominski. 2009. FLASH, a proapoptotic protein involved in activation of caspase-8, is essential for 3' end processing of histone pre-mRNAs. *Mol Cell*. 36:267-278.

Yang, X.C., I. Sabath, J. Debski, M. Kaus-Drobek, M. Dadlez, W.F. Marzluff, and Z. Dominski. 2013. A complex containing the CPSF73 endonuclease and other polyadenylation factors associates with U7 snRNP and is recruited to histone pre-mRNA for 3'-end processing. *Mol Cell Biol*. 33:28-37.

Yang, X.C., I. Sabath, L. Kunduru, A.J. van Wijnen, W.F. Marzluff, and Z. Dominski. 2014. A conserved interaction that is essential for the biogenesis of histone locus bodies. *J Biol Chem*. 289:33767-33782.

Ye, X., Y. Wei, G. Nalepa, and J.W. Harper. 2003. The cyclin E/Cdk2 substrate p220(NPAT) is required for S-phase entry, histone gene expression, and Cajal body maintenance in human somatic cells. *Mol Cell Biol*. 23:8586-8600.

Zaidi, S.K., D.W. Young, A. Javed, J. Pratap, M. Montecino, A. van Wijnen, J.B. Lian, J.L. Stein, and G.S. Stein. 2007. Nuclear microenvironments in biological control and cancer. *Nature reviews. Cancer*. 7:454-463.

Zhao, J., B. Dynlacht, T. Imai, T. Hori, and E. Harlow. 1998. Expression of NPAT, a novel substrate of cyclin E-CDK2, promotes S-phase entry. *Genes Dev*. 12:456-461.

Zhao, J., B.K. Kennedy, B.D. Lawrence, D.A. Barbie, A.G. Matera, J.A. Fletcher, and E. Harlow. 2000. NPAT links cyclin E-Cdk2 to the regulation of replication-dependent histone gene transcription. *Genes Dev*. 14:2283-2297.

Zhu, L., and C.P. Brangwynne. 2015. Nuclear bodies: the emerging biophysics of nucleoplasmic phases. *Curr Opin Cell Biol*. 34:23-30.

CHAPTER II

DISTINCT SELF-INTERACTION DOMAINS PROMOTE MULTI SEX COMBS ACCUMULATION IN AND FORMATION OF THE *DROSOPHILA* HISTONE LOCUS BODY

This work was previously published in the journal Molecular Biology of the Cell (MBoC). I performed the experimental component of this paper in its entirety. Shawn Lyons and I jointly conducted the pull-down experiments shown in Figure 2.3. John Poulton and I jointly designed and conducted the live-imaging experiments using *Drosophila* syncytial embryos shown in Figure 2.8. My advisor Dr. Robert Duronio designed the project. He, Dr. William Marzluff, and I wrote the manuscript and analyzed the data. Supplemental data and movies can be found on the journal's website.

Terzo, E.A., Lyons, S.M., Poulton, J.S., Temple, B. R. S., Marzluff, W.F., and Duronio, R.J. (2015)

Distinct self-interaction domains promote Multi Sex Combs accumulation in and formation of the *Drosophila* histone locus body.

Mol. Biol. Cell. 26:1559-74.

Overview

Nuclear bodies (NBs) are structures that concentrate proteins, RNAs, and ribonucleoproteins that perform functions essential to gene expression. How NBs assemble is not well understood. We studied the *Drosophila* histone locus body (HLB), a NB that concentrates factors required for histone mRNA biosynthesis at the replication-dependent histone gene locus. We coupled biochemical analysis with confocal imaging of both fixed and live tissues to demonstrate that the *Drosophila* Multi-Sex Combs (Mxc) protein contains multiple domains necessary for HLB assembly. An important feature of this assembly process is the self-interaction of Mxc via two conserved N-terminal domains: a LisH domain and a novel SIF (*S*elf *I*nteraction *F*acilitator) domain immediately downstream of the LisH domain. Molecular modeling suggests that the LisH and SIF domains directly interact, and mutation of either the LisH or SIF domains severely impairs Mxc function *in vivo* resulting in reduced histone mRNA accumulation. A region of Mxc between amino acids 721 and 1481 is also necessary for HLB assembly independent of the LisH and SIF domains. Lastly, the C-terminal 195 amino acids of Mxc are required for recruiting FLASH, an essential histone mRNA processing factor, to the HLB. We conclude that multiple domains of the Mxc protein promote HLB assembly in order to concentrate factors required for histone mRNA biosynthesis.

Introduction

Numerous levels of molecular organization within the nucleus facilitate the highly regulated expression of the genome. One level of organization is the concentration of proteins, RNAs, and ribonucleoproteins into structures known as nuclear bodies (NBs) that are visible by light microscopy (Dundr and Misteli, 2001; Dundr and Misteli, 2010; Gall, 2000; Matera, 1999; Matera et al., 2009; Misteli, 2001; Misteli, 2005; Parada et al., 2004). NBs include well-known structures such as Cajal bodies and the nucleolus, and less well understood structures including PML bodies, speckles, paraspeckles and histone locus bodies (HLBs). An attractive hypothesis for NB function posits that NBs concentrate factors to accelerate reactions that would otherwise take longer if these factors were dispersed throughout the nucleus (Dundr, 2012). This hypothesis has gained support from studies of vertebrate Cajal bodies, which promote efficient spliceosomal snRNP assembly (Klingauf et al., 2006; Machyna et al., 2014; Novotny et al., 2011; Strzelecka et al., 2010). However, *Drosophila* snRNA modification by scaRNAs, which are localized to Cajal bodies, does not require Cajal body assembly (Deryusheva and Gall, 2009). Thus, the general applicability and further tests of this hypothesis require additional study.

An understanding of NB function requires detailed knowledge of NB composition and assembly. We have been exploring this issue by studying how HLB assembly contributes to the expression of replication-dependent histone genes, which encode the only known cellular mRNAs that are not polyadenylated (Marzluff et al., 2008). HLBs were defined by Gall and coworkers as a NB associated with the *Drosophila* histone gene locus that contained U7 snRNP (Liu et al., 2006), a factor essential for generating the

unique histone mRNA 3' end (Mowry and Steitz, 1987; Strub and Birnstiel, 1986). Similar factors necessary for histone transcription and pre-mRNA processing are found in both vertebrate and *Drosophila* HLBs, including human NPAT (Nuclear Protein mapped to the mutated Ataxia Telangectasia locus), which was identified as a Cyclin E/Cdk2 substrate essential for histone mRNA expression (Ma et al., 2000; Miele et al., 2005; Wei et al., 2003; (Zhao et al., 2000). The *multi sex combs* (*mxc*) locus encodes the *Drosophila* ortholog of NPAT. Mxc, like NPAT, is phosphorylated by Cyclin E/Cdk2, co-localizes with U7 snRNP at the histone locus, and is required for both HLB assembly and histone gene expression (White et al., 2011). Other known HLB components include FLASH and Mute. FLASH was identified in mammals as co-localizing with NPAT (Bongiorno-Borbone et al., 2008) and subsequently shown to interact with U7 snRNP and to be essential for histone pre-mRNA processing (Yang et al., 2009). *Mute* was identified as a *Drosophila* HLB component in a screen for factors required for muscle development, but its biochemical function is not known (Bulchand et al., 2010).

Our previous experiments on *Drosophila* HLBs suggest that Mxc is critical for HLB assembly. Mxc and FLASH localize to the histone locus immediately before the beginning of histone gene expression in syncytial embryos, and prior to this time HLBs are not detected. Loss of Mxc results in a failure to localize other HLB components, including FLASH and U7 snRNP (White et al., 2011). The *Drosophila* HLB is present in all cells, independent of whether they are cycling (Liu et al., 2006; White et al., 2007). The 5 canonical *Drosophila* histone genes (H1, H2A, H2B, H3, and H4) are clustered together in a 5kb sequence that is repeated approximately 100 times at a single locus on chromosome 2. The 300 base pair bidirectional promoter of the H3-H4 gene pair within

this cluster is necessary and sufficient for HLB assembly and is necessary for expression of the adjacent H2A-H2B gene pair (Salzler et al., 2013). This 300 bp sequence is also sufficient to recruit Mxc and FLASH, consistent with Mxc playing an integral role in HLB assembly and histone gene expression. How Mxc participates in coordinating these processes remains unclear.

The *mxc* locus was originally described by an allelic series of mutations in which null alleles resulted in a failure of cell proliferation and lethality. Knocking out NPAT in cultured mammalian cells is similarly lethal (Ye et al., 2003). In contrast, viable, hypomorphic *mxc* alleles cause homeotic transformations in adult males (giving rise to the gene name) (Santamaria and Randsholt, 1995). Whether there is any causal relationship between histone gene expression and the homeotic transformations observed in *mxc* hypomorphs is unknown. Two *mxc* hypomorphic alleles encode nonsense mutations at residues K1482 and Q1643 of the 1837-amino acid long Mxc protein (White et al., 2011). The resulting truncated mutant proteins support histone gene expression (Landais et al., 2014) while an amorphic *mxc* allele that does not produce Mxc protein does not (White et al., 2011). The Q1643→Stop mutation (mxc^{G46}) partially disrupts Mxc function resulting in replication stress and a persistent DNA damage response that contributes to the loss of germ line stem cells through mis-regulation of histone gene expression (Landais et al., 2014).

Studies in human cell culture indicate that distinct domains of NPAT are required to activate histone gene expression and allow entry into S phase (Wei et al., 2003). These data suggest that Mxc/NPAT may contain multiple domains that organize HLB assembly and coordinate histone mRNA biosynthesis. Proteins harboring multiple protein-protein

interaction domains are likely a critical feature of NBs (Matera et al., 2009). The focal organization provided by the multiple interaction domains could facilitate a more efficient and rapid physiological response to distinct stimuli (Bian et al., 2012; Cortese et al., 2008; Foray et al., 2003; Good et al., 2011; Matera et al., 2009; Nussinov et al., 2013; Zaidi et al., 2007). Here, we identify functional domains of Mxc required for localization of Mxc to the HLB in the presence of full-length Mxc using cultured *Drosophila* S2 cells. We use *mxc* mutant animals expressing different Mxc mutant transgenic proteins to explore the function of Mxc *in vivo*, and time-lapse imaging of early embryos expressing GFP-Mxc to assess the dynamics of Mxc localization to the HLB. Our data indicate that Mxc requires multiple domains for complete function *in vivo*, and that two self-interaction domains of Mxc are essential for HLB assembly, which in turn promotes histone mRNA biosynthesis.

Results

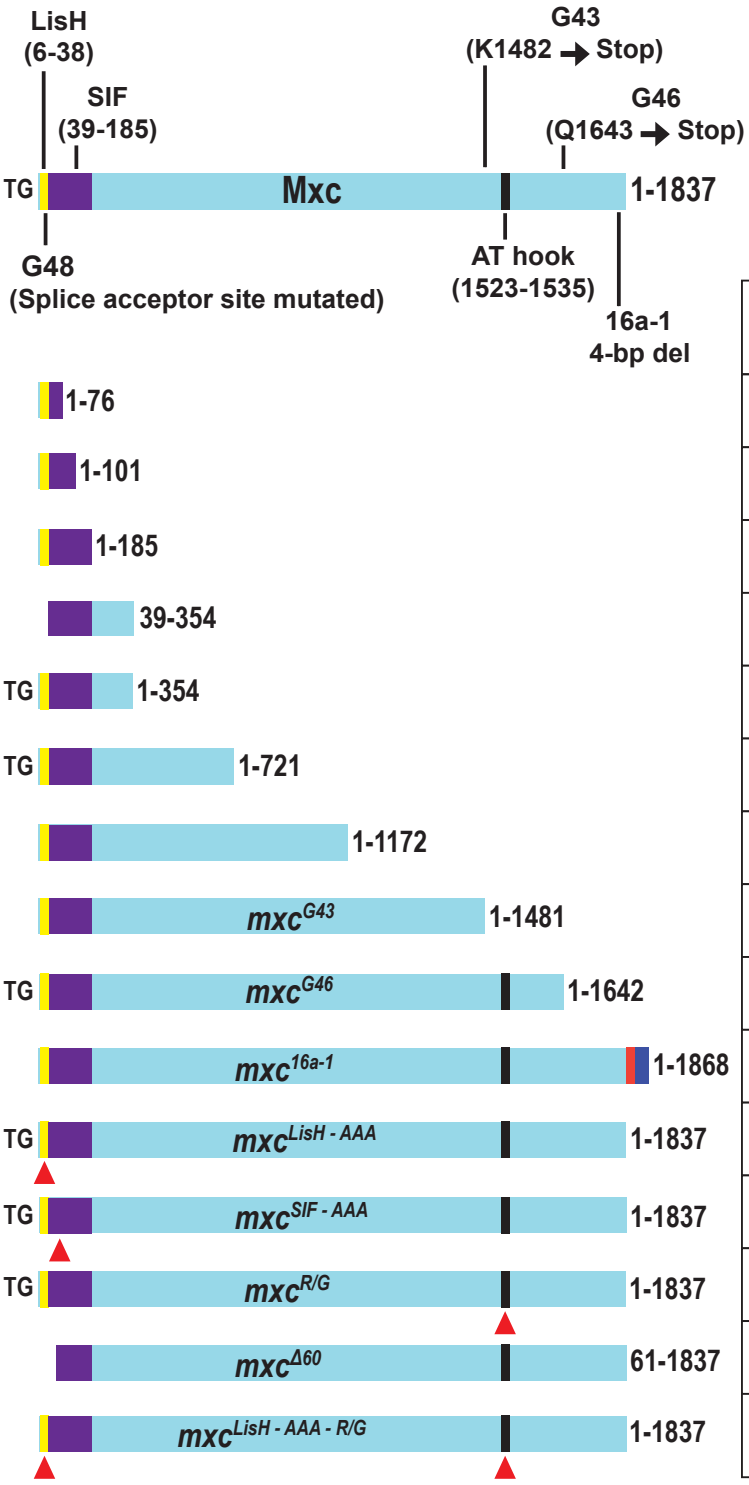
Two NH_2-terminal domains are required for Mxc concentration in HLBs

The design of the Mxc mutants we analyzed for this study was directed by both homology to previously characterized protein domains and the properties of a collection of *mxc* mutant alleles (Santamaria and Randsholt, 1995) that we have recently sequenced (White et al., 2011). The 1837-amino acid Mxc protein contains only two small domains recognizable by primary sequence, a LisH domain at the N-terminus (amino acids 6 to 38) and a 13-amino acid long AT-hook motif toward the C-terminus (amino acids 1523-1535). LisH domains are 33-amino acid motifs readily identifiable by primary sequence homology, with invariant hydrophobic residues at positions 9 and 13 and an aromatic

residue at position 12 (Figure S1A, arrowheads). Structural and biochemical analyses indicate that LisH domains adopt a characteristic fold that mediates protein-protein interactions, including homo-dimerization (Cerna and Wilson, 2005; Gerlitz et al., 2005; Kim et al., 2004; Mikolajka et al., 2006). AT hook domains bind the minor groove of DNA at AT-rich stretches and are characterized by an invariant peptide core motif of R-G-R-P that is well conserved and that is flanked on both sides by positively charged amino acids (Aravind, 1998; Harrer et al., 2004; Reeves, 1990). The mxc^{G48} null allele, with a AG to AA splice acceptor mutation at the intron 1/exon 2 border, does not produce detectable full length Mxc protein. The three other alleles we sequenced (mxc^{16a-1}, mxc^{G43}, and mxc^{G46}) are predicted to generate altered Mxc proteins (Figure 2.1). The hypomorphic mxc^{G43} and mxc^{G46} alleles each have a mutation resulting in a premature stop codon and are predicted to express 1481- and 1642-amino acid long proteins, respectively. While both of these alleles are viable, mxc^{G43} has a stronger phenotype (consistent with having a larger deletion) with fewer progeny developing to adulthood than mxc^{G46} (Santamaria and Randsholt, 1995). The mxc^{16a-1} mutant contains a 4-base pair deletion/frame shift near the end of the open reading frame resulting in replacement of the last 14 amino acids of Mxc with 45 residues not normally present in wild type Mxc. Interestingly, the mxc^{16a-1} mutant is not viable, although it should produce an Mxc protein with a relatively small alteration at the C-terminus, compared to the more extensive deletions in Mxc^{G43} and Mxc^{G46}.

To determine the regions of Mxc necessary for concentration in the HLB, we designed constructs encoding GFP-Mxc^{16a-1}, GFP-Mxc^{G43}, GFP-Mxc^{G46}, and three

additional larger deletion mutants (GFP-Mxc$^{1-354}$, GFP-Mxc$^{1-721}$, GFP-Mxc$^{1-1172}$) as N-



	Mxc Accumulation in the HLB	
Expressed protein	wt Mxc present	wt Mxc absent
1-76	-	ND
1-101	+	ND
1-185	+	ND
39-354	-	ND
1-354	+	-
1-721	+	-
1-1172	+	ND
G43	+	+
G46	+	+
16-a1	+	-
LisH-AAA	↓	-
SIF-AAA	↓	-
R/G	+	+
Δ60	-	ND
LisH-AAA R/G	-	ND

Figure 2.1. Mxc structure/function analysis. Top, full-length Mxc displaying LisH (yellow), SIF (green), and AT hook (black) domains and previously described mutations of *mxc*. Numbers to the left of each Mxc

fragment indicate length in amino acids. Red arrowheads indicate V14, Y17, L18; L58, I61, I62; and R1528 point mutations in LisH, SIF, and AT hook domains, respectively. **TG** indicates fragments used to generate transgenic flies. Table compiles results regarding the ability of each Mxc fragment to form a detectable HLB either in the presence or absence of wild-type (wt) Mxc. (+) Accumulation; (-) No accumulation; (ê) Decreased accumulation; and (ND) Not determined.

terminal GFP fusion proteins and expressed them in S2 cells (Figures 2.1 and 2.2). We stained transfected S2 cells with antibodies against GFP (to detect exogenous Mxc) and antibodies against FLASH or Mxc to mark the endogenous S2 cell HLB. Note that our Mxc antibody was raised against the last 169 amino acids of Mxc and therefore does not detect GFP-Mxc fusion proteins lacking the C-terminus of Mxc. The five deletion mutants, including the smallest, GFP-Mxc$^{1\text{-}354}$, were capable of concentrating in the endogenous HLB (Figure 2.2A-G). The GFP-Mxc$^{16a\text{-}1}$ protein behaved differently than the deletion mutants: while we could detect some co-localization with FLASH in S2 cells, GFP-Mxc$^{16a\text{-}1}$ was also mis-localized in large foci throughout the nucleus. In addition, FLASH was also mis-localized in these cells (Figure 2.2H). This result indicates that the altered C-terminus encoded by *mxc$^{16a\text{-}1}$* disrupts both Mxc and FLASH concentration in the HLB, an issue that we explore further below.

We also tested whether the Mxc AT hook domain was necessary for concentrating Mxc in the HLB in S2 cells. Mutation of the second conserved Arginine of the Pro-Arg-Gly-Arg-Pro AT hook consensus motif to Glycine in High-mobility-group protein A1a (HMGA1a) results in redistribution of HMGA1a within interphase nuclei (Harrer et al., 2004). We therefore changed Arg^{1528} of the Arg-Gly-Arg-Pro Mxc AT hook motif to Gly (GFP-Mxc$^{R/G}$). GFP-Mxc$^{R/G}$ concentrated in the HLB in S2 cells similarly to control GFP

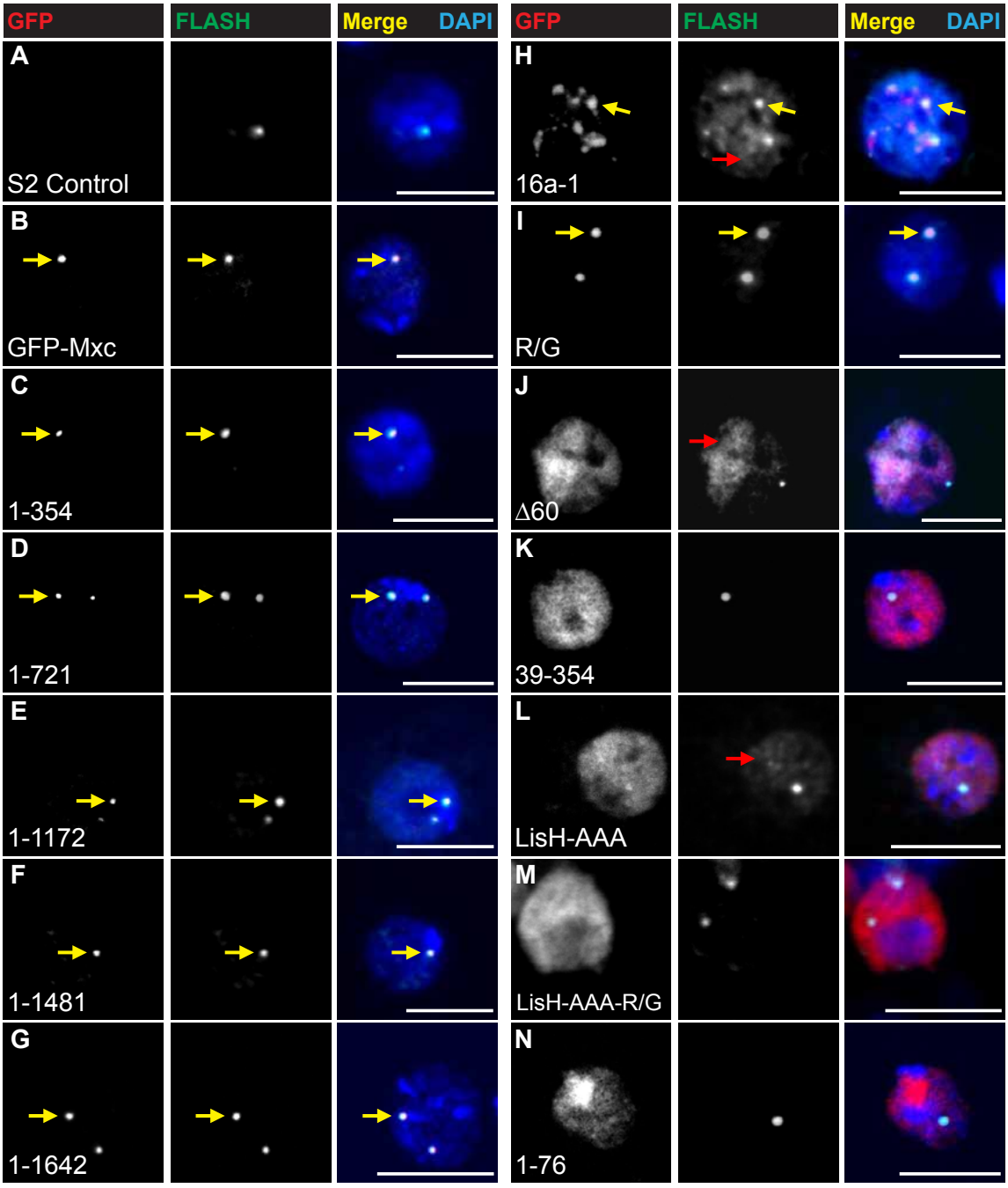


Figure 2.2. The Mxc N-terminus is required for concentration in the HLB in cultured cells.

Untransfected S2 cells (A) or S2 cells transfected with constructs expressing the indicated GFP-Mxc proteins (B-N) were stained with anti-GFP and anti-FLASH antibodies. Yellow arrows indicate foci of co-localizing GFP-Mxc and FLASH. Note that transfection of mis-localized Mxc proteins with an intact C terminus result in mis-localized FLASH (red arrows in H, J, and L). Bars: 10μM.

Mxc (Figure 2.2B and I), suggesting that the AT hook domain is not necessary to concentrate exogenous Mxc in the HLB, consistent with the results of the C-terminal deletion experiments which remove the AT hook domain.

To more precisely define the sequences required for concentration in the HLB in S2 cells, we used N-terminal deletions to explore whether the LisH domain plays a role in concentrating Mxc in the HLB. We tested if an otherwise full length Mxc lacking the first 60 amino acids encompassing the LisH motif (GFP-Mxc$^{\Delta 60}$) would be able to concentrate in the HLB. GFP-Mxc$^{\Delta 60}$ did not concentrate in the HLB in S2 cells, but rather was found throughout the nucleus (Figure 2.2J). A GFP-Mxc fragment lacking the LisH domain (GFP-Mxc$^{39\text{-}354}$) also failed to concentrate in the HLB (Figure 2.2K). These data suggest that the LisH domain is required for concentration of Mxc in the HLB. Note that endogenous FLASH also became partially mis-localized after expression of GFP-Mxc$^{\Delta 60}$, but not after expression of GFP-Mxc$^{39\text{-}354}$, which lacks the C-terminus (Figure 2.2J and K). This result suggests that the presence of the C-terminus in a mis-localized Mxc can result in the mis-localization of FLASH, perhaps because the C-terminus of Mxc binds to FLASH.

Eleven of the first twelve amino acids in the LisH domain of Mxc are identical in all vertebrate NPATs, and overall human NPAT LisH domain is 51% identical to that of Mxc. Three amino acids, Val14, Tyr17, and Leu18, are conserved between the NPAT and Mxc LisH domains and well conserved among LisH domains in other proteins (Figure S1A). Mutating these three amino acids to Ala, which is not expected to sterically hinder LisH domain formation (Kim et al., 2004), abrogated Mxc's concentration in the HLB in S2 cells (Figure 2.2L). GFP-Mxc$^{\text{LisH-AAA}}$ accumulated throughout the nucleus, and caused

some FLASH mis-localization, similar to GFP-Mxc$^{\Delta 60}$ (Figure 2.2J and L). In addition, a GFP-Mxc$^{\text{LisH-AAA-R/G}}$ double mutant behaved similarly to the GFP-Mxc$^{\text{LisH-AAA}}$ single mutant (Figure 2.2M). Taken together, the LisH deletion and point mutant data indicate that the LisH domain plays an essential role in concentrating Mxc in the HLB. However, the LisH domain is not sufficient for HLB localization, as GFP-Mxc$^{1\text{-}76}$ did not concentrate in the HLB in S2 cells (Figure 2.2N). Together these data demonstrate that the LisH domain and residues between 76-354 of Mxc provide critical determinants for concentrating exogenous Mxc in the HLB in the presence of endogenous Mxc.

The Mxc N-terminus promotes Mxc self-interaction

Structural studies have shown that some LisH domains directly interact with each other, mediating dimerization (Kim et al., 2004). We postulated that the N-terminal 354 amino acids of Mxc function to promote the self-interaction of Mxc molecules, and that the LisH domain was part of this interaction. To explore this possibility, we conducted *in vitro* pull-down assays using a recombinant protein fragment expressed in *E. coli* as Maltose Binding Protein (MPB) fused to the first 101 amino acids of Mxc (MBP-Mxc^{101}) (Fig. 2.3A-15 and B). We expressed ^{35}S-Met labeled Mxc fragments by in vitro translation in rabbit reticulocyte lysates and tested their ability to interact with recombinant MBP-Mxc^{101}. We efficiently pulled down ^{35}S-Met-labeled Mxc$^{1\text{-}354}$ using MBP-Mxc^{101} but not with MBP alone, indicating that the N-terminus of Mxc interacts with itself (Figure 2.3A-1). Two shorter fragments of Mxc, Mxc$^{1\text{-}185}$ and Mxc$^{1\text{-}101}$, were also pulled down by MBP-Mxc^{101} (Figure 2.3A-2 and 3) and were capable of concentrating in the S2 cell HLB as effectively as Mxc$^{1\text{-}354}$ (Figure S3A, B). Interestingly,

an Mxc fragment lacking the LisH domain (Mxc$^{39-354}$) was also pulled down by MBP-Mxc^{101}, indicating that the Mxc self-interaction does not require LisH domain homodimerization.

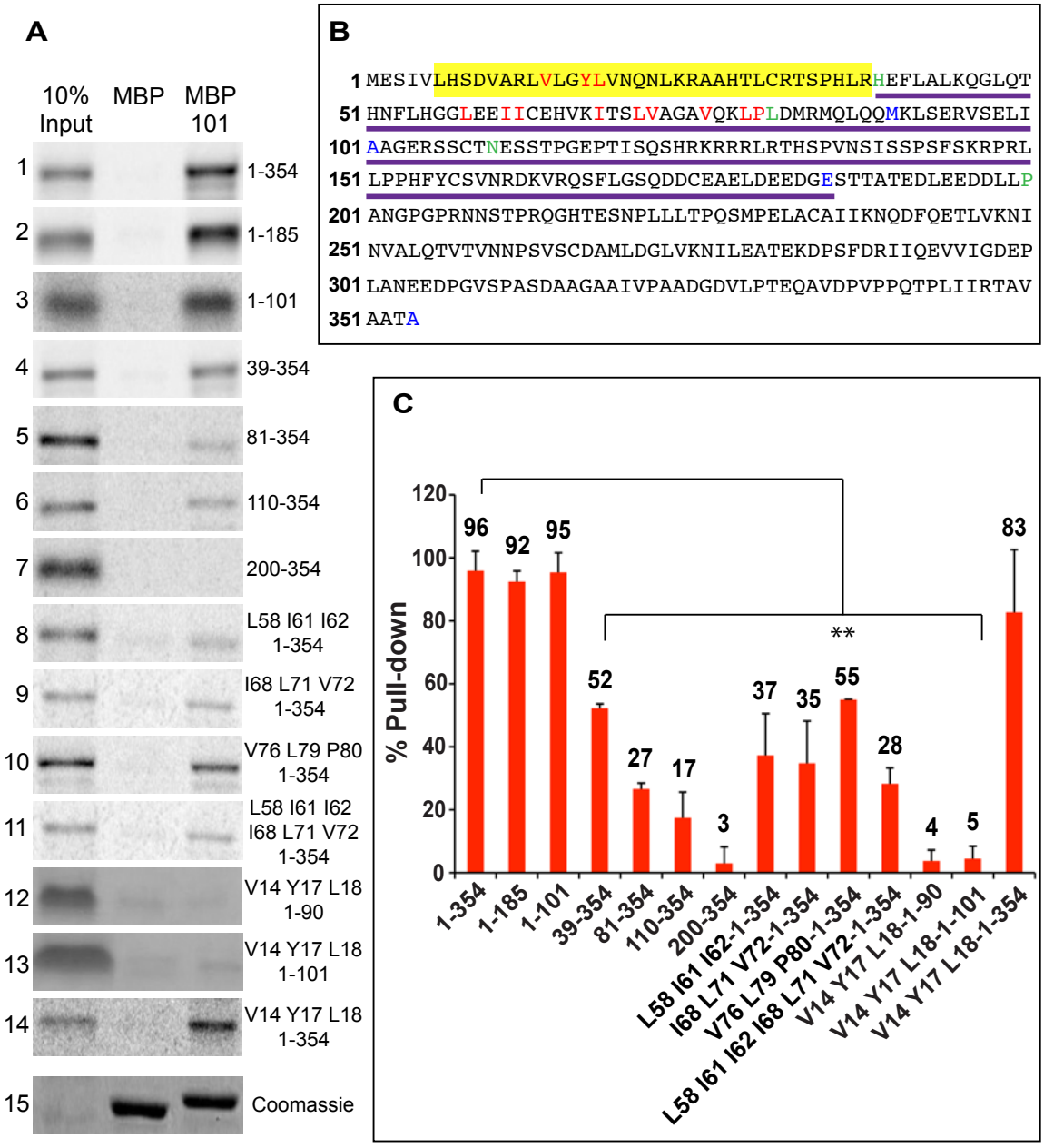


Figure 2.3. Mxc self-interaction requires two N-terminal domains. (A1-14) ^{35}S-Met-labeled, *in vitro* translated Mxc fragments (indicated at right) precipitated with MBP-Mxc^{101} (MBP 101) and run side-by-side with a 10% input to compare percentage of pull-down. (A15) Coomassie-stained gel showing similar

loading of recombinant MBP proteins. (B) Sequence of first 354 amino acids of Mxc with the LisH domain (L6-R38) indicated in yellow and the SIF domain (H39-E185) underlined. Amino acids in red indicate Ala substitution mutations in LisH (V14, Y17, and L18) and SIF (L58, I61, I62, I68, L71, V72, V76, L79, and P80) domains. Residues in green indicate N-terminal amino acids in the 39-354 (L39), 81-354 (L81), 110-354 (N110), and 200-354 (P200) fragments. Residues in blue indicate C-terminal amino acid in the 1-90 (M90), 1-101 (A101), 1-185 (E185), and 1-354 (A354) fragments. (C) Bar graph showing percentage of pull-down for each Mxc fragment. Error bars represent SEM. The double asterisk indicates all statistically significant ($p < 0.001$) differences in binding compared to Mxc^{1-354}.

Mxc^{39-354} was pulled down about half as efficiently as Mxc^{1-354}, Mxc^{1-185}, and Mxc^{1-101} (Figure 2.3A1-4, and C). These results suggest that sequences in addition to the LisH domain can promote Mxc self-interaction. Indeed, further deletion of the N-terminus (Mxc^{81-354} and $Mxc^{110-354}$) further reduced, but did not abolish, binding to MBP-Mxc^{101} (Figure 2.3A-5 and 6, and C). A fragment from amino acid 200 to 354 ($Mxc^{200-354}$) did not bind MBP-Mxc^{101} (Figure 2.3A-7 and C). Taken together, these data suggest that residues downstream of the LisH domain between amino acids 39 and 101 are necessary for high affinity Mxc self-interaction.

To identify candidate residues in this region, we performed an *in silico* structural analysis. LisH domains consist of a helix-turn-helix motif that typically homodimerizes. A homodimer of the Mxc LisH domain was modeled based on the crystallographic homodimer of the LisH domain of TBL1X (PDB ID 2XTC). Analysis of the modeled Mxc homodimer revealed the possibility of a steric clash between His-7 of one LisH domain and Tyr-17 of the second LisH domain, suggesting that Mxc LisH domains do not homodimerize consistent with our pull down data. We hypothesized that additional structural motifs within the N-terminal 101 amino acids of Mxc would interact with the

LisH domain and would also contain residues in a similar helical structure that correspond to the highly conserved VxxYL hydrophobic residues within LisH domains that typically drive LisH homodimerization. We identified three sets of hydrophobic residues between amino acids 58 and 80 of Mxc that were in helical regions, and which might drive self-interaction. These three motifs were LxxII (L58-I61-I62), IxxLV (I68-L71-V72) and VxxLP (V76-L79-P80) (Figure 2.3B). We therefore constructed three different sets of triple Ala substitution mutations in Mxc^{1-354} and measured binding to MBP-Mxc^{101} using the pull down assay. The binding of Mxc^{1-354} fragments containing L58A-I61A-I62A, I68A-L71A-V72A, or V76A-L79A-P80A mutations to MBP-Mxc^{101} was reduced 50-65% relative to wild type Mxc^{1-354} (Figure 2.3A-8-10, B and C). An Mxc^{1-354} fragment carrying both L58A-I61A-I62A and I68A-L71A-V72A mutations did not further reduce binding to MBP-Mxc^{101} (Figure 2.3A-11). These data indicate that specific residues between amino acids 58 and 80 are required for efficient Mxc self-interaction, perhaps through a heterologous interaction with the LisH domain (Figure S1B).

To further interrogate the role of the LisH domain in Mxc self-interaction, we generated Mxc^{1-90} and Mxc^{1-101} fragments containing the LisH triple Ala mutation (V14A-Y17A-L18A; Figure S1A). Either fragment harboring a mutagenized LisH domain showed >90% reduction in binding to MBP-Mxc^{101} (Figure 2.3A-12, 13 and 2.3C). In contrast, the LisH domain mutation in Mxc^{1-354} did not significantly affect binding to MBP-Mxc^{101} (Figure 2.3A-14, and C). These data indicate that a mutant LisH domain had little affect on Mxc self-interaction when additional downstream residues are present. Because $Mxc^{200-354}$ does not bind MBP-Mxc^{101} while Mxc^{1-185} binds very well, we

conclude that the C-terminal boundary of these additional amino acids is before residue 185. When all of our biochemical data are considered together (Figure 2.1), the results indicate that high affinity Mxc self-interaction requires two distinct regions, the LisH domain (residues 6-38) and sequences between amino acids 39 and 185, which we designated the Mxc *S*elf-*I*nteraction *F*acilitator (SIF) domain.

Multiple Mxc domains including the N-terminal self-interaction domains are required for HLB formation in vivo and for the completion of development

To identify domains of Mxc required for *in vivo* function, we determined which of our Mxc transgenes (Figure 2.1) encoded proteins that concentrate in the HLB and whether they were capable of rescuing the lethality of the mxc^{G48} null allele. These transgenes utilize the *ubiquitin-63E* promoter to ubiquitously express proteins with GFP fused to the N-terminus of wild type or mutant Mxc. We first determined whether these proteins were present in the HLB in the presence of endogenous Mxc (Figure 2.4). GFP-Mxc concentrated in the HLB in embryos and ovarian follicle cells in the presence of the endogenous Mxc, as did the transgenic GFP-$Mxc^{1\text{-}354}$ and GFP-$Mxc^{1\text{-}721}$ proteins (Figure 2.4A-D, and G-J). In contrast, transgenic full-length $Mxc^{LisH\text{-}AAA}$ or $Mxc^{SIF\text{-}AAA}$ proteins did not localize to the HLB in the presence of endogenous Mxc (Figure 2.4E, F, K, and L).

Embryos that are homozygous for mxc^{G48} hatch, develop to second-instar larvae, and then die. Expressing full length GFP-Mxc in the homozygous mxc^{G48} background completely rescued mxc^{G48} lethality (i.e. supported development to adulthood) and resulted in assembly of HLBs that were indistinguishable from wild type (Figure S2A). In

fact, we can maintain a stock containing GFP-Mxc as the only functional copy of Mxc. GFP-Mxc$^{R/G}$, which contains a point mutation in the A/T hook domain, is also capable of

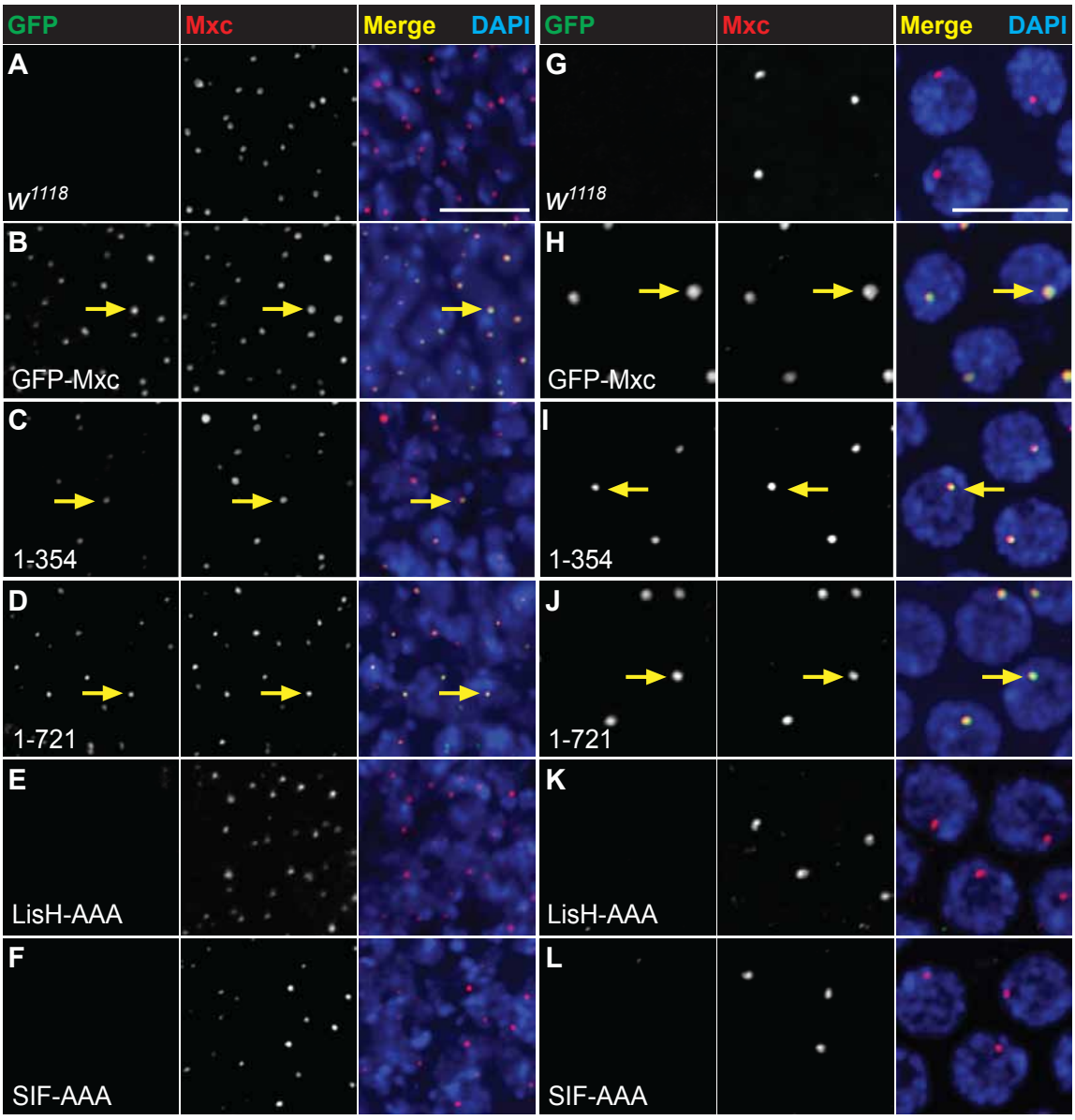


Figure 2.4. Mxc requires the LisH and SIF domains to concentrate at the histone locus in transgenic flies. 8-10-hour old embryos (A-F) and ovarian follicle cells (G-L) expressing the indicated transgenes were stained with anti-GFP and anti-Mxc antibodies. Note that GFP-Mxc$^{1-354}$ and GFP-Mxc$^{1-721}$ localize with endogenous Mxc (yellow arrows), but form smaller foci than GFP-Mxc. Bars: 10μM (A-F) and 5μM (G-L).

rescuing *mxc*^{G48}. In contrast full-length Mxc harboring either a mutant LisH domain (Mxc$^{LisH\text{-}AAA}$) or a mutant SIF domain (Mxc$^{SIF\text{-}AAA}$) could not rescue *mxc*^{G48} lethality or support HLB assembly (Figure S2B, C). The Mxc$^{SIF\text{-}AAA}$ allele contains the L58A-I61A-I62A mutation that reduces self-interaction in the pull down assay (Figure 2.3A-8, C), and we selected this mutation to test *in vivo* because these residues are conserved in human NPAT (Figure S1C). Western blotting revealed that the Mxc$^{LisH\text{-}AAA}$ and Mxc$^{SIF\text{-}AAA}$ mutant proteins accumulate to levels similar to wild type GFP-Mxc (Figure S3C). These data indicate that the self-interaction domains we identified *in vitro* are required for Mxc function *in vivo*.

The GFP-Mxc$^{1\text{-}354}$ and GFP-Mxc$^{1\text{-}721}$ deletion mutants, which contain wild type LisH and SIF domains and localized to the HLB in S2 cells and wild type embryos and follicle cells, failed to rescue lethality of *mxc*^{G48} or to support HLB assembly in the absence of endogenous Mxc (Figure S2B-E). This observation demonstrates that the concentration of GFP-Mxc$^{1\text{-}354}$ and GFP-Mxc$^{1\text{-}721}$ in HLBs in S2 cells and wild type embryos requires interaction with endogenous, full length Mxc. Moreover, these results demonstrate that sequences between amino acid 721 and the C-terminus of Mxc are necessary for HLB assembly and *Drosophila* development.

To explore in more detail the functional domains within the 721-1837 region, we analyzed the hypomorphic mutants *mxc*^{G46} and *mxc*^{G43}, which expresses Mxc proteins truncated at amino acid 1642 and 1481, respectively (Figure 2.1). Because our anti-Mxc antibody was raised against the last 169 amino acids of Mxc, it will not detect these proteins. We therefore generated a transgenic GFP-Mxc^{G46} protein to determine if this

Mxc truncation is able to concentrate in the HLB. In a wild-type background, GFP-Mxc^{G46} protein co-localizes at the histone locus with endogenous Mxc and FLASH (Figure 2.5A). In an *mxc^{G48}* null background, GFP-Mxc^{G46} forms foci resembling HLBs and rescues *mxc^{G48}* lethality (Figure 2.5B-E). In addition, *mxc^{G48}* males expressing GFP-Mxc^{G46} are sterile. Nuclear foci were detected in brains from *mxc^{G46}* third instar larvae or from *mxc^{G48}* larvae expressing GFP-Mxc^{G46} after staining with MPM-2, a monoclonal antibody that recognizes Cyclin E/Cdk2-dependent phosphorylation sites in Mxc (Figure 2.5F-H). Similarly, we detected MPM-2 foci in brains from *mxc^{G43}* mutant larvae (Figure 2.5I,J), which express an Mxc protein truncated at amino acid 1481 (Figure 2.1). A small fraction of *mxc^{G43}* mutants survive to adulthood (Remillieux-Leschelle et al., 2002; Saget et al., 1998; Santamaria and Randsholt, 1995). Together these results demonstrate that a mutant Mxc with a C-terminal truncation to amino acid 1481 is capable of assembling into an HLB nuclear body and supporting the completion of development, although inefficiently.

The C-terminus of Mxc recruits HLB components required for histone mRNA synthesis

To determine which domains of Mxc are necessary for histone mRNA transcription and pre-mRNA processing, and the relationship between these processes and HLB formation, we measured total accumulation of histone H3 mRNA in our panel of mutants by fluorescence in situ hybridization of 8-10 hour old embryos using a probe from the coding region of H3 (H3-cod). By 8 hours of embryogenesis the maternal stores of Mxc are substantially depleted as assayed by immunofluorescence (Figure 2.5C), as

we previously reported (White et al., 2011), allowing us to assess the capability of the different mutant GFP-Mxc proteins to support histone mRNA synthesis. In control

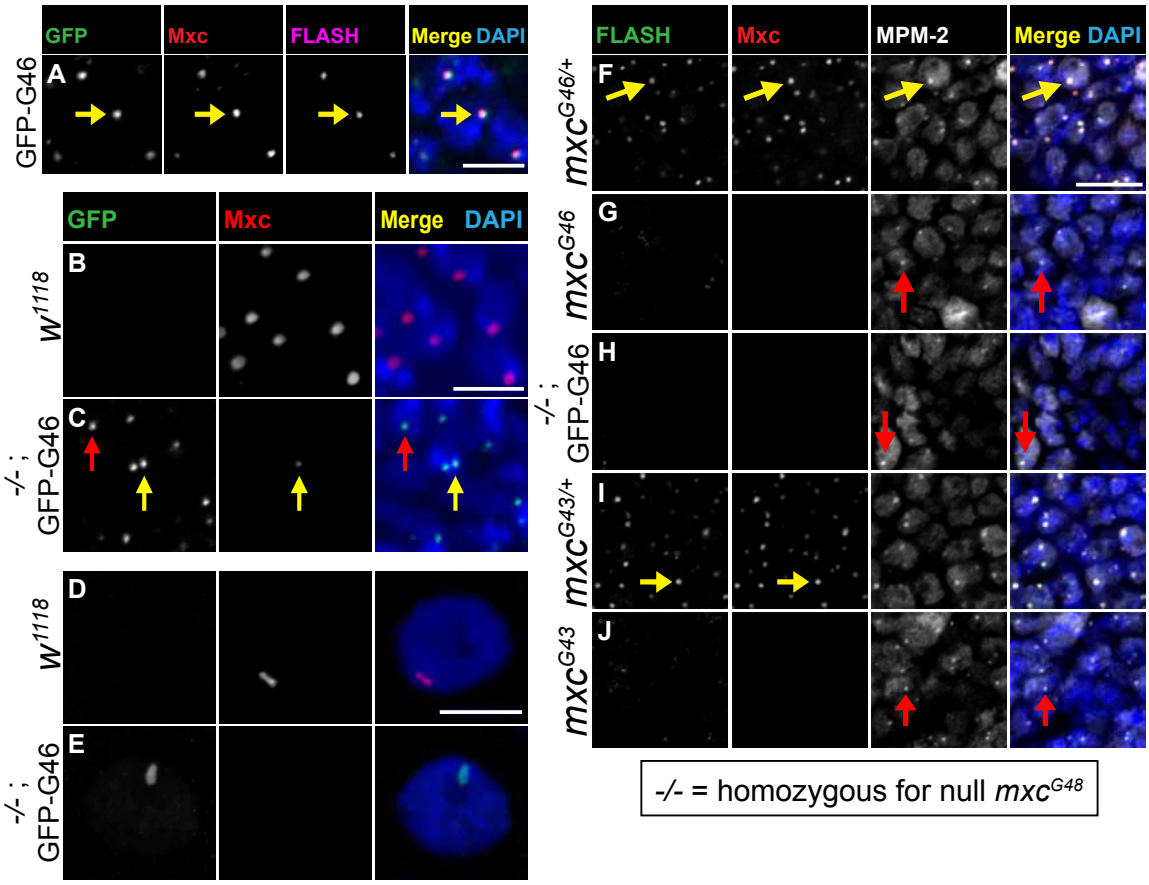


Figure 2.5. The 721-1481 region of Mxc is necessary for HLB assembly and completion of development. (A) 8-10-hour old transgenic embryo expressing GFP-Mxc^{G46} stained with anti- GFP, FLASH, and Mxc antibodies. Yellow arrow indicates GFP-Mxc^{G46} localization to the endogenous HLB. (B and C) *w*^{1118} control and a homozygous *mxc*^{G48} mutant embryo expressing GFP-Mxc^{G46} stained with anti-GFP and Mxc antibodies. The yellow arrow in (C) indicates a cell with an HLB containing a small amount of full length, maternal Mxc still present (our anti-Mxc antibody does not recognize the trunctated Mxc^{G46} protein). The red arrow in (C) indicates a cell where maternal Mxc has been depleted from the HLB. (D and E) Nuclei from salivary glands of *w*^{1118} and *mxc*^{G48}*; gfp-G46* third-instar larvae stained with anti- GFP and Mxc antibodies. (F-J) Third-instar larval brains of the indicated genotypes stained with anti-FLASH, anti-Mxc and MPM-2 antibodies. Yellow arrows indicate foci containing all three HLB markers. Red arrows

indicate MPM-2 positive foci containing Mxc^{G46}, GFP-Mxc^{G46}, and Mxc^{G43} (G, H, and J, respectively). Note that MPM-2 also detects other proteins in the nucleus. Bar: 5μM (A-C) and 10μM (D-J).

embryos that have endogenous wild type Mxc concentrated in the HLB, histone H3 mRNA accumulates in the cytoplasm of actively cycling cells (Figure 2.6A). In mxc^{G48} null mutant embryos, histone H3 mRNA levels are reduced and HLBs are not detectable with anti-Mxc antibodies that would only detect maternal Mxc (Figure 2.6B). Expression of GFP-Mxc restores HLB assembly and histone H3 mRNA expression (Figure 2.6C). GFP-$Mxc^{LisH\text{-}AAA}$ and GFP-$Mxc^{SIF\text{-}AAA}$ transgenic proteins in the mxc^{G48} null background fail to concentrate in the HLB and fail to support normal accumulation of histone H3 mRNA (Figures 2.6D and E). We observed a similar phenotype with GFP-$Mxc^{1\text{-}354}$ and GFP-$Mxc^{1\text{-}721}$ (Figures 26F and G), indicating that either mutation of the Mxc self-interaction domains or C-terminal truncation to amino acid 721 eliminates the ability of Mxc to support HLB formation and normal histone gene expression. In contrast, we detected slightly higher accumulation of histone H3 mRNA in the viable, hypomorphic truncation alleles mxc^{G43} (residues 1-1481) and mxc^{G46} (residues 1-1642) (Figures 2.6A, C, H, and I).

The in situ hybridization results were corroborated by northern blot analysis of RNA extracted from 15-18 hour old embryos or from early second instar larvae, near the lethal phase of the mxc^{G48} null allele (Figures 2.6K and L, respectively). Quantification of these data reveal a reproducible ~25% reduction of H3 mRNA accumulation in mxc^{G48} relative to control embryos, and that the GFP-$Mxc^{LisH\text{-}AAA}$, GFP-$Mxc^{SIF\text{-}AAA}$, GFP-$Mxc^{1\text{-}354}$, GFP-$Mxc^{1\text{-}721}$ alleles are indistinguishable from null (Figure S4). These data suggest that very small amounts of maternal Mxc in the HLB that is undetectable by

immunofluorescence can support some histone mRNA expression, or that there is a basal level of expression that can occur in the absence of Mxc. H3 mRNA accumulation in *mxc*^{G43} and *mxc*^{G46} embryos is reduced relative to control, but less so than the null alleles.

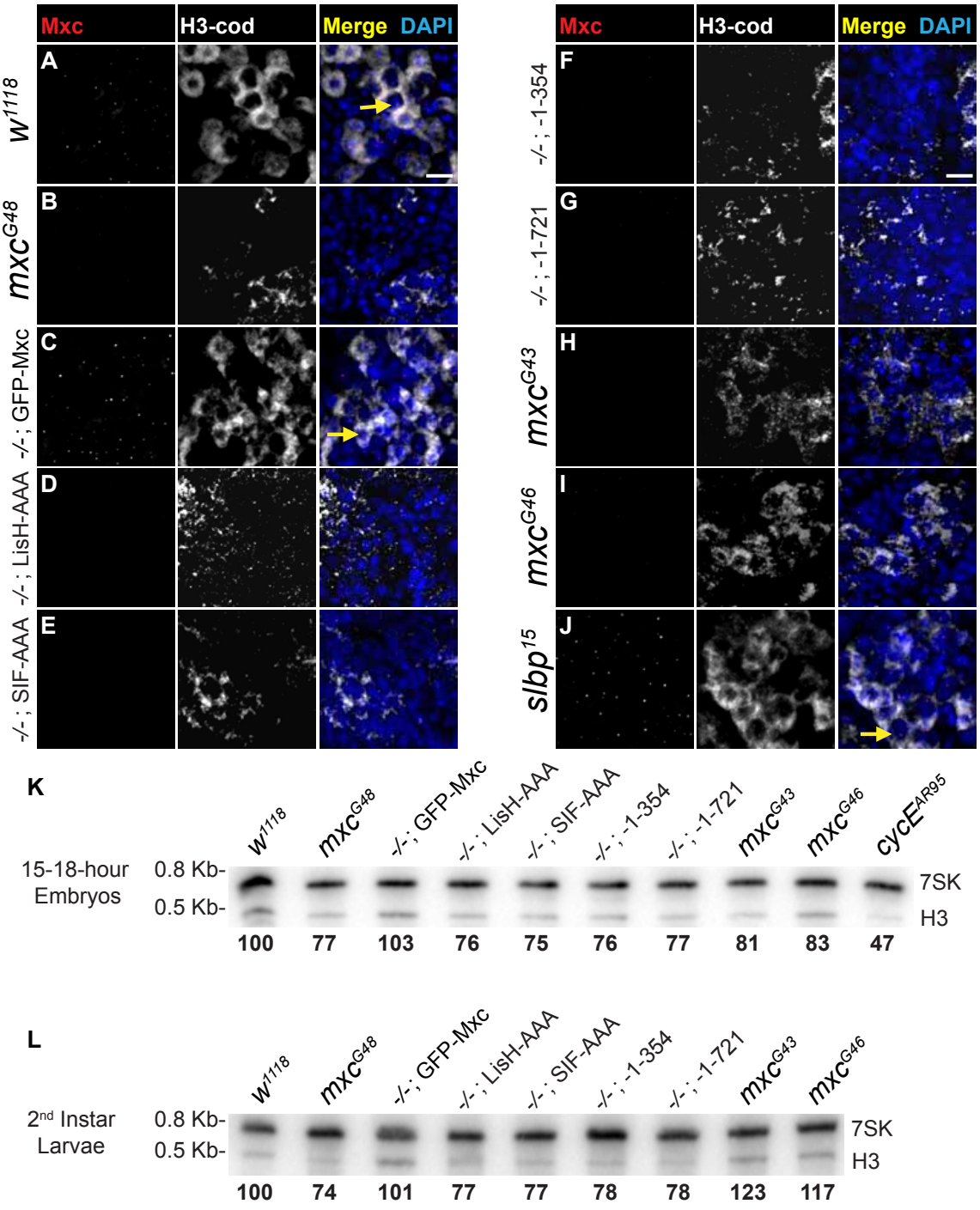


Figure 2.6. Mxc concentration in the HLB is required for Histone mRNA biosynthesis. (A-J) 8-10-hour old embryos of the indicated genotypes were subjected to fluorescence *in situ* hybridization (FISH) with an RNA probe generated from the H3 coding region (H3-cod) and co-stained with anti-Mxc antibodies. Images of epithelial cells were specifically obtained from the cephalic region. Yellow arrows indicate Mxc foci in the nuclei of actively cycling cells that accumulate histone H3 mRNA in the cytoplasm. Note that mxc^{G43} and mxc^{G46} mutant embryos (H and I) accumulate detectable amounts of H3 mRNA. Bars: 10μM. (K and L) Northern blot analysis of histone H3 transcript levels from two developmental stages of different *mxc* mutants. 1μg of total RNA from 15-18-hour old embryos (K) and 5μg of total RNA from 2nd instar larvae (L) per well were run on a 6% acrylamide 8M Urea denaturing gel. 7SK RNA was used as a loading control on both gels. Numbers below each lane represent the averaged percentage of histone H3 transcript levels obtained from three independent experiments. Homozygous $cycE^{AR95}$ mutant embryos were used as a control as Cyclin E is known to be required for DNA replication and cell cycle progression in dividing and endocyling cells after cycle 16 (Knoblich et al., 1994) and also for histone mRNA expression (Lanzotti et al., 2004a).

Curiously, in the mxc^{G43} and mxc^{G46} larval samples H3 mRNA accumulation is greater than in control, consistent with a previous study (Landais et al., 2014). The mechanistic basis for the mxc^{G43} and mxc^{G46} H3 larval mRNA expression phenotype is not known.

We next determined if histone pre-mRNA processing was disrupted in our panel of Mxc mutants. Loss of histone pre-mRNA processing factors, such as the Stem Loop Binding Protein (SLBP) or U7 snRNP, results in transcription past the normal processing site, utilization of cryptic, downstream polyadenylation signals, and the accumulation of cytoplasmic poly A+ histone mRNA (Godfrey et al., 2006; Sullivan et al., 2001). These aberrant histone H3 transcripts are readily detected by in situ hybridization using a probe (H3-ds) derived from sequences downstream of the normal H3 mRNA 3' end (Lanzotti et al., 2002; Lanzotti et al., 2004a). The H3-ds probe does not hybridize to wild type, mxc^{G48}

null mutant; *mxc*^{G48}; GFP-Mxc rescued embryos; or the cytoplasm of *mxc*^{G43} or *mxc*^{G46} embryos because all of the histone mRNA is processed normally, as judged by Northern blotting (Figure 2.6K and L), and the H3-ds probe does not detect processed H3 mRNA (Figure 2.7A-C). In contrast, in *Slbp*^{15} null mutant embryos the H3-ds probe detects nascent, read through H3 transcripts in the nucleus that co-localize with Mxc-positive HLBs (Figure 2.7D, red arrow). These read-through transcripts are processed to poly A+ H3 mRNA and exported to the cytoplasm where they are detected with either the H3-cod (Figure 2.6J) or H3-ds probes (Figure 2.7D, yellow arrowhead) (Lanzotti et al., 2002; Lanzotti et al., 2004b).

Using this assay we determined if any of the Mxc mutants accumulate unprocessed H3 mRNA at the site of transcription. As expected, the H3-ds probe did not hybridize to Mxc mutant embryos with GFP-Mxc$^{\text{LisH-AAA}}$, GFP-Mxc$^{\text{SIF-AAA}}$, GFP-Mxc$^{1\text{-}354}$, and GFP-Mxc$^{1\text{-}721}$ transgenic proteins that failed to assemble an HLB and consequently fail to express H3 mRNA in the *mxc*^{G48} null background (data not shown). In contrast, we detected robust nuclear foci with the H3-ds probe in both *mxc*^{G43} and *mxc*^{G46} mutant embryos (Figure 2.7E, F). *mxc*^{G46} embryos reproducibly contained more and brighter H3-ds foci than homozygous *mxc*^{G43} embryos (Figure 2.7E, F), suggesting a higher rate of histone gene transcription in the *mxc*^{G46} mutant (Figure 2.6H, I, and K). In each mutant the H3-ds foci were fewer and dimmer than in *Slbp*^{15} mutant embryos (Figures 2.7D-F), perhaps because wild type Mxc and the normal HLB in the *Slbp*^{15} mutants drives more transcription than the Mxc^{G43} and Mxc^{G46} mutant proteins.

The FLASH protein, which is essential for histone pre-mRNA processing, is not concentrated in the HLB in *mxc*^{G43} and *mxc*^{G46} mutants (Figure 2.5G, J) (Rajendra et al., 2011), providing a possible explanation for the presence of misprocessed H3 mRNA.

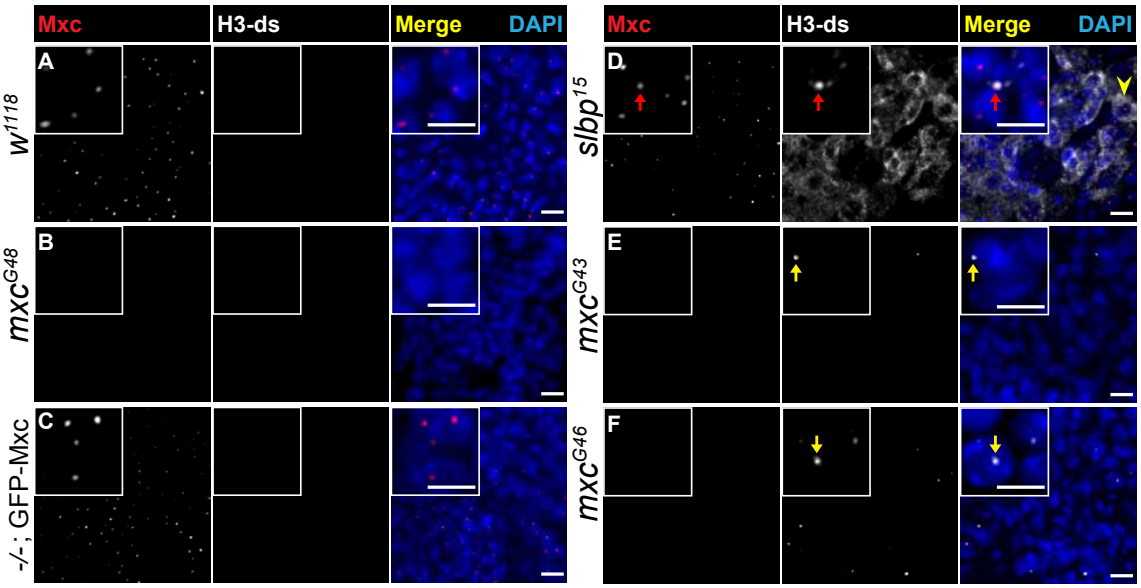


Figure 2.7. The C-terminus of Mxc is required to concentrate essential histone mRNA processing factors at the histone locus. (A-F) 8-10-hour old embryos of the indicated genotypes were subjected to fluorescence *in situ* hybridization (FISH) with an RNA probe generated from a region downstream of the normal H3 pre-mRNA processing site (H3-ds) and co-stained with anti-Mxc antibodies. Images were obtained of epithelial cells from the cephalic region. Insets show a higher magnification of nuclei. Red arrows in (D) indicate foci of nascent, unprocessed H3 RNA that co-localize with Mxc at the histone locus in *slbp*^{15} homozygous mutant embryos (Lanzotti et al., 2004a). Note that in *slbp*^{15} mutant embryos the H3 RNA is mis-processed to poly A+ mRNA that is exported to and accumulates in the cytoplasm (yellow arrowhead in D). Yellow arrows in (E and F) indicate foci of nascent, unprocessed H3 RNA in nuclei of *mxc*^{G43} (E) and *mxc*^{G46} (F) mutant embryos. Scale bars: 10μM and 5μM in insets (A-F).

Interestingly, using the H3-ds probe we did not detect misprocessed H3 mRNA in the cytoplasm of *mxc*^{G43} and *mxc*^{G46} mutant cells, although we did detect histone mRNA with

the H3-cod probe (Figure 2.6H, I). This result suggests that the nascent read-through transcripts are ultimately processed at the normal site and exported. We explore this observation in more depth in a separate study (D. Tatomer, E.T, W.F.M, and R.J.D., submitted).

Both the LisH and SIF domains are required for efficient accumulation of Mxc in the HLB

While the severe phenotypes observed with large C-terminal deletions of Mxc are not surprising, two different 3 amino acid changes ($Mxc^{LisH\text{-}AAA}$ and $Mxc^{SIF\text{-}AAA}$) effectively inactivated the 1837 residue Mxc protein. To more carefully investigate the effects that mutating the self-interaction domains has on Mxc localization and behavior *in vivo*, we conducted time lapse imaging experiments on live embryos expressing GFP-$Mxc^{LisH\text{-}AAA}$ and GFP-$Mxc^{SIF\text{-}AAA}$. We focused on the first 2 hours of embryogenesis when HLB formation first occurs (White et al., 2011). At this time, *Drosophila* embryos are a syncytium in which nuclei undergo thirteen rapid, synchronous cycles composed only of S phase and Mitosis (Swanhart et al., 2005). With our previous imaging of fixed embryos, we first detected Mxc nuclear foci during cycle 10, one cycle before histone gene expression begins (White et al., 2011). By imaging live embryos expressing GFP-Mxc and H2Av-RFP to visualize chromosomes, we detected small GFP-Mxc nuclear foci as early as interphase of cycle 9 (Figure 2.8A-A', and Figure 8 movie1), suggesting that the live imaging approach is more sensitive. These foci become larger in each subsequent cycle, as more defined and much brighter GFP-Mxc foci become visible during interphase of cycle 10, and again in cycle 11, when the mature HLB has formed (Figure

2.8B, C, J, and Figure 8 movie1). Additionally, we detected small GFP-Mxc foci associated with mitotic chromosomes (Figure S5A-A''), as we previously observed in fixed embryos (White et al., 2011). Our live imaging also revealed a low level of GFP-Mxc signal coincident with the H2Av-RFP signal from condensed mitotic chromosomes (Fig S5B-B''). One possibility for this observation is that Mxc associates with all chromosomes during mitosis, and then becomes concentrated in the HLB at the histone locus during interphase. However, we cannot eliminate the possibility that this chromosome interaction results from over-expression of GFP-Mxc relative to endogenous Mxc and does not normally happen. The increase in intensity of GFP-Mxc foci during cycles 9-11 (Figure 2.8J) suggests that the HLB expands in size after initial nucleation or "seeding" as early as cycle 9. To test whether this HLB expansion requires Mxc self-interaction, we performed live-imaging experiments with GFP-Mxc$^{LisH\text{-}AAA}$ and GFP-Mxc$^{SIF\text{-}AAA}$. To our surprise, we observed discrete foci of GFP-Mxc$^{LisH\text{-}AAA}$ and GFP-Mxc$^{SIF\text{-}AAA}$, again suggesting that our live imaging is more sensitive than our imaging of fixed embryos. The GFP-Mxc$^{LisH\text{-}AAA}$ and GFP-Mxc$^{SIF\text{-}AAA}$ foci were considerably dimmer than those formed by GFP-Mxc, and were first detectable in cycle 11 rather than cycle 9 (Figure 2.8D-I', J, Figure 2.8 movie2, and Figure 2.8 movie3). These results indicate that Mxc$^{LisH\text{-}AAA}$ and GFP-Mxc$^{SIF\text{-}AAA}$ are defective for HLB accumulation during the syncytial cycles. Both mutant proteins also associated with mitotic chromosomes, but again these signals were weaker than that obtained with GFP-Mxc (Figure S5C-F''). Out of necessity these experiments were performed in the presence of maternal supplies of wild type Mxc; therefore the small foci and mitotic chromosome association may result from a weak interaction between endogenous Mxc and either GFP-Mxc$^{LisH\text{-}AAA}$ or GFP-Mxc$^{SIF\text{-}AAA}$.

These data indicate that GFP-Mxc$^{LisH\text{-}AAA}$ and GFP-Mxc$^{SIF\text{-}AAA}$ are defective in HLB localization and suggest that Mxc self-interaction is a critical component of HLB assembly during development.

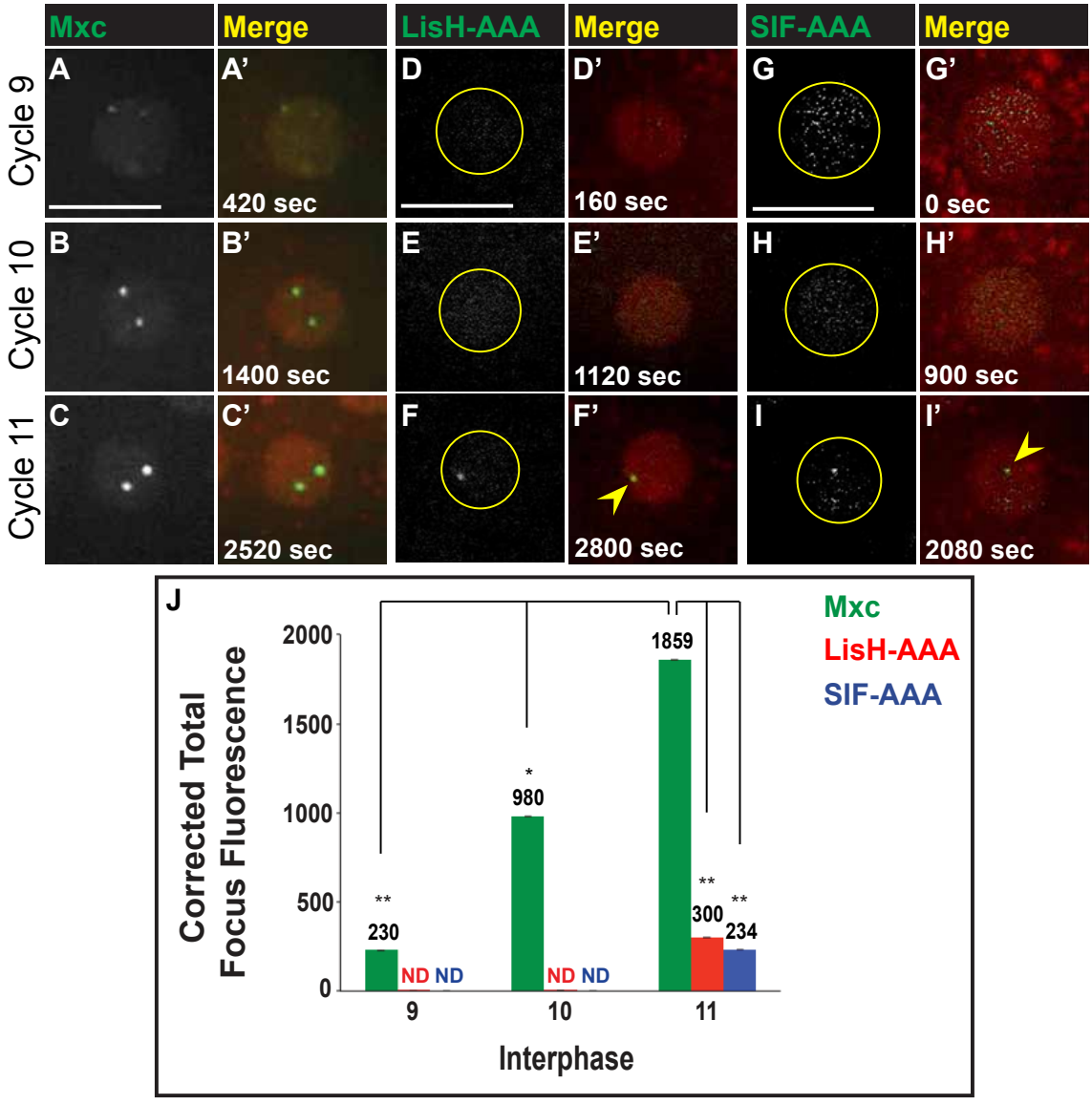


Figure 2.8. Mxc's LisH and SIF domains promote HLB localization. (A-I') Still images take from time-lapse movies of syncytial *H2Av-rfp/gfp-mxc*, *H2Av-rfp/gfp-mxc*$^{LisH\text{-}AAA}$, and *H2Av-rfp/gfp-mx*$^{SIF\text{-}AAA}$ transgenic embryos. Yellow circles denote the nuclear periphery. Yellow arrowheads point to dim foci of Mxc$^{LisH\text{-}AAA}$ (F') and Mxc$^{SIF\text{-}AAA}$ mutant proteins (I'). Note that red signal outside of nuclei are lipid droplets

containing maternally supplied H2Av protein (Li et al., 2014). Bars: 10μM. (J) Bar graph showing corrected total focus fluorescence values from interphase of cycles 9 to 11. Error bars represent the standard error of the mean (SEM). Numbers above bars represent averaged CTFF values. ND = Not determined CTFF values due to undetectable foci. Significant differences are indicated either by a single ($p < 0.05$) or double ($p < 0.001$) asterisk.

Discussion

HLBs assemble at replication-dependent histone loci and provide a distinct compartment in the nucleus that promotes efficient transcription and processing of histone mRNA, likely by concentrating histone biosynthetic factors as well as excluding factors specifically required for polyadenylation (Dundr, 2012). In this study we show that multiple protein domains are necessary for Mxc to support HLB assembly and histone mRNA biosynthesis, and ultimately normal *Drosophila* development.

Multiple domains of Mxc are required for HLB assembly

Whether NBs form by an ordered assembly process, by random association of components, or by a combination of each of these processes is not clear for most NBs (Matera et al., 2009). In the case of the HLB, we have demonstrated that hierarchical assembly contributes to NB formation, with Mxc and FLASH part of a complex that initially forms at a specific sequence at the histone locus (Salzler et al., 2013; White et al., 2011). Here we defined two regions in the N-terminus of Mxc, the LisH domain and a novel domain we have named the SIF domain, both of which are necessary for GFP-Mxc to concentrate in the HLB in the presence of endogenous Mxc and to support HLB assembly in the absence of endogenous Mxc.

Although GFP-Mxc$^{1-354}$ and GFP-Mxc$^{1-721}$, which contain both LisH and SIF domains, are incorporated into the HLB in the presence of endogenous Mxc, they do not support formation of a complete HLB in the absence of endogenous Mxc and cannot rescue the lethality caused by an *mxc* null mutation. Thus sequences in addition to the LisH and SIF domains are required for HLB formation. Truncated Mxc proteins encoded by the viable, hypomorphic *mxc^{G43}* and *mxc^{G46}* alleles (1481 and 1642 amino acids, respectively) form nuclear bodies (HLBs) as judged by staining tissues with the MPM-2 antibody, which recognizes phosphorylated Mxc, and formation of nuclear foci by GFP-Mxc^{G46} protein in the absence of endogenous Mxc. Thus, there is a region of Mxc between amino acids 721 and 1481 that together with the N-terminus is required for HLB formation. The larger Mxc proteins likely contain elements necessary for recruitment of Mxc to the H3-H4 intergenic region of the histone locus that is essential for HLB formation (Salzler et al., 2013). However, because maternal supplies of wild type Mxc initially establish the HLB in the early embryo prior to the zygotic expression of *mxc^{G43}* and *mxc^{G46}*, we cannot be certain that Mxc^{G43} and Mxc^{G46} proteins are capable of forming an HLB de novo. Finally, Mxc likely contains binding sites for other HLB components, such as FLASH, U7 snRNP or Mute, and is regulated by phosphorylation by Cyclin E/Cdk2.

Self-interaction between different Mxc molecules is required for HLB assembly

LisH domains are found in a variety of multi-protein complexes, and promote protein-protein interactions important for the assembly of these complexes (Cerna and Wilson, 2005; Gerlitz et al., 2005; Kim et al., 2004; Mikolajka et al., 2006). Some LisH

domain proteins dimerize through their LisH domains, and a structure of a LisH domain homodimer has been solved (Kim et al., 2004). We find that the Mxc N-terminus promotes interaction of two Mxc molecules, but that this interaction does not occur by LisH domain homodimerization. In Mxc there is a possible steric clash between His-7 of one LisH domain and Tyr-17 of a second LisH domain that may explain why the Mxc LisH domains do not homodimerize in a manner typical of other LisH domains. Instead, Mxc self-interaction requires a region downstream of the LisH domain between amino acids 39 and 185 (the SIF domain), and three amino acids (Leu52, Ile61, and Ile62) in this region conserved between flies and vertebrates are required for HLB assembly *in vivo* and for rescuing the lethality of an *mxc* null mutation. Furthermore, live imaging revealed dramatically reduced concentration of GFP-Mxc$^{LisH\text{-}AAA}$ and GFP-Mxc$^{SIF\text{-}AAA}$ in HLBs in the presence of endogenous Mxc, consistent with reduced binding affinity between the mutant and wild type Mxc molecules.

Thus, the LisH domain of one molecule of Mxc binds the SIF domain (i.e. amino acids 39-185) of another molecule of Mxc. Our molecular modeling suggests that this interaction may be mediated by direct binding between the LisH domain and the LxxII motif of the SIF domain (Figure 2.9). In addition to the LxxII motif, the SIF domain contains other amino acids that contribute to efficient Mxc self-interaction. These multiple interaction sites indicate that each Mxc molecule can potentially interact with at least two, and possibly more Mxc molecules, raising the possibility that the N-terminal region of Mxc can promote formation of a three-dimensional lattice that is likely an essential component of HLB structure (Figure 2.10). Similarly, an N-terminal domain of Coilin that mediates self interaction is necessary for Coilin accumulation in the CB

(Hebert and Matera, 2000), suggesting that oliogomerization is a common feature of NB formation.

Many LisH domain-containing proteins also contain a CTLH domain (C-terminus to LisH) defined in both ProSite and SMART (Adams, 2002; Emes and Ponting, 2001; Umeda et al., 2003), which is often but not always immediately C-terminal to the LisH domain. Other than the prediction that this domain contains a-helical regions there is no structural information on the CTLH domain. The CTLH domains of several proteins have recently been shown to participate in protein-protein interactions important for the assembly of multi-protein complexes (Kobayashi et al., 2007; Menssen et al., 2012; Salemi et al., 2014; Sun et al., 2013). The Mxc SIF domain that we have identified functions similarly to the CTLH domain but is clearly distinct from the CTLH domain. The SMART and ProSite CTLH domain logos each contain a conserved glycine (G) at position 16, a conserved phenylalanine (F) at position 46, a conserved leucine (L) at position 48, a conserved glutamic acid (E) at position 55 (numbering of SMART logo), none of which are present in the SIF domain of Mxc. Thus the region in Mxc C-terminal to the LisH domain is distinct from the CTLH domain.

Harper and colleagues previously demonstrated that human NPAT is essential for cell proliferation and histone gene expression, and that the NPAT LisH domain was necessary for stimulating His4 and H2B promoter activation in cell culture based transfection/reporter assays (Wei et al., 2003; Ye et al., 2003).

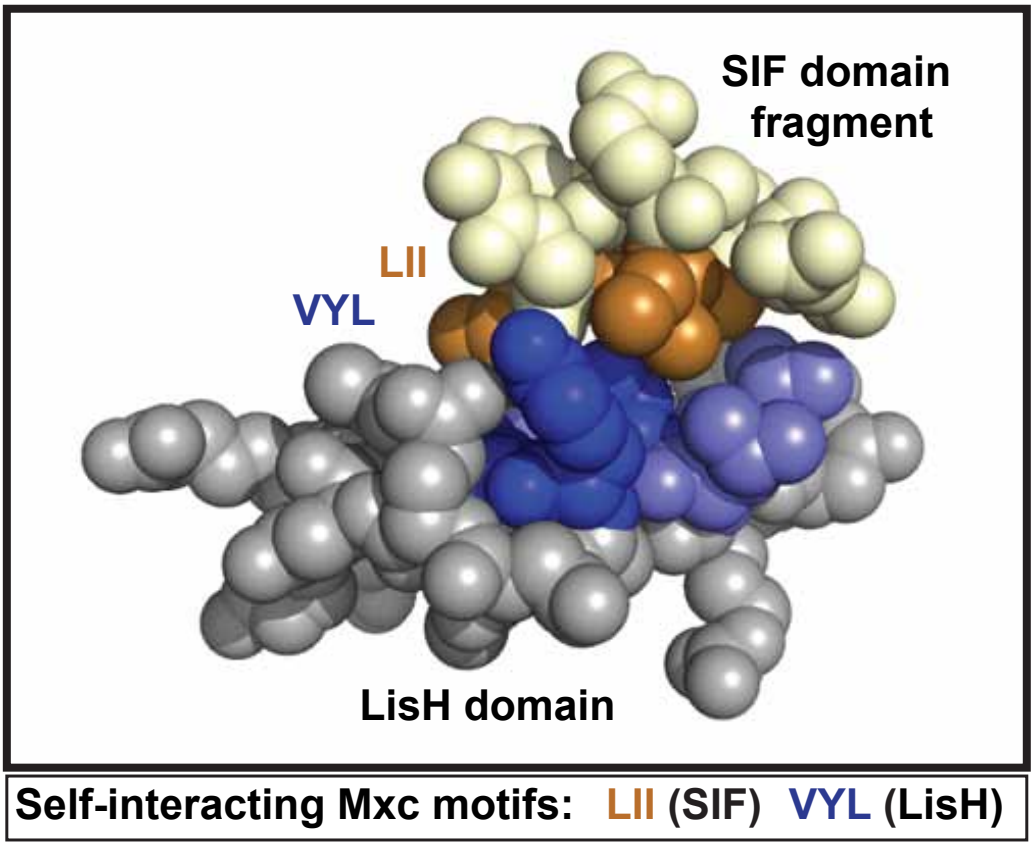


Figure 2.9. LisH domain/SIF domain self-interaction model. Space filling model of the proposed interaction of the Mxc LisH domain with the Mxc SIF domain. The LisH domain of Mxc is light gray with the VYL motif colored dark blue and neighboring hydrophobic residues in LisH colored slate blue. A small fragment of the SIF domain (GGLEEIICE) rendered in PyMOL is colored light yellow with the critical LII motif colored copper.

They also reported that a LisH domain mutant NPAT protein could localize to Coilin-positive NBs (a subset of which are likely to be HLBs) (Wei et al., 2003). However these experiments were performed by transfecting RAT1 cells containing endogenous NPAT, and the role of the LisH domain in NB formation, cell proliferation and histone gene expression was not examined in the absence of endogenous NPAT. In addition, mutations

of the NPAT SIF domain were not generated and analyzed in these previous studies. Based on our results and the similarity between the N-termini of mammalian NPAT and Mxc (Figure S1), we suspect that human NPAT LisH domain mutants can interact with endogenous NPAT via the SIF and/or other domains. We propose that the N-terminus of human NPAT promotes interaction between multiple NPAT molecules.

Mxc's requirement for histone mRNA biosynthesis correlates with HLB assembly

Our prior imaging of fixed embryos and our live imaging reported here indicate that maternal Mxc and FLASH co-localize in nuclear foci prior to the initiation of zygotic histone gene transcription in the syncytial embryo (Salzler et al., 2013; White et al., 2011; White et al., 2007). Once histone transcription initiates these foci enlarge into mature HLBs as detected by increased intensity of both Mxc and FLASH staining as well as recruitment of other HLB components U7 snRNP and Mute. We previously reported that *mxc* null mutant 1st instar larvae fail to accumulate normal amounts of histone H3 mRNA, supporting a role for Mxc in histone gene expression (White et al., 2011). Here we show that the maternal supply of Mxc (as determined by detection of HLBs by immunofluorescence) is depleted in most cells by 8 hrs of embryogenesis, and that this depletion is accompanied by a decrease in histone H3 transcript levels. In spite of reduced levels of histone mRNA, *mxc* null mutant embryos hatch. Thus, as the maternal supply of Mxc is depleted in *mxc* mutant embryos, histone gene expression drops resulting in larval death in early larval stages.

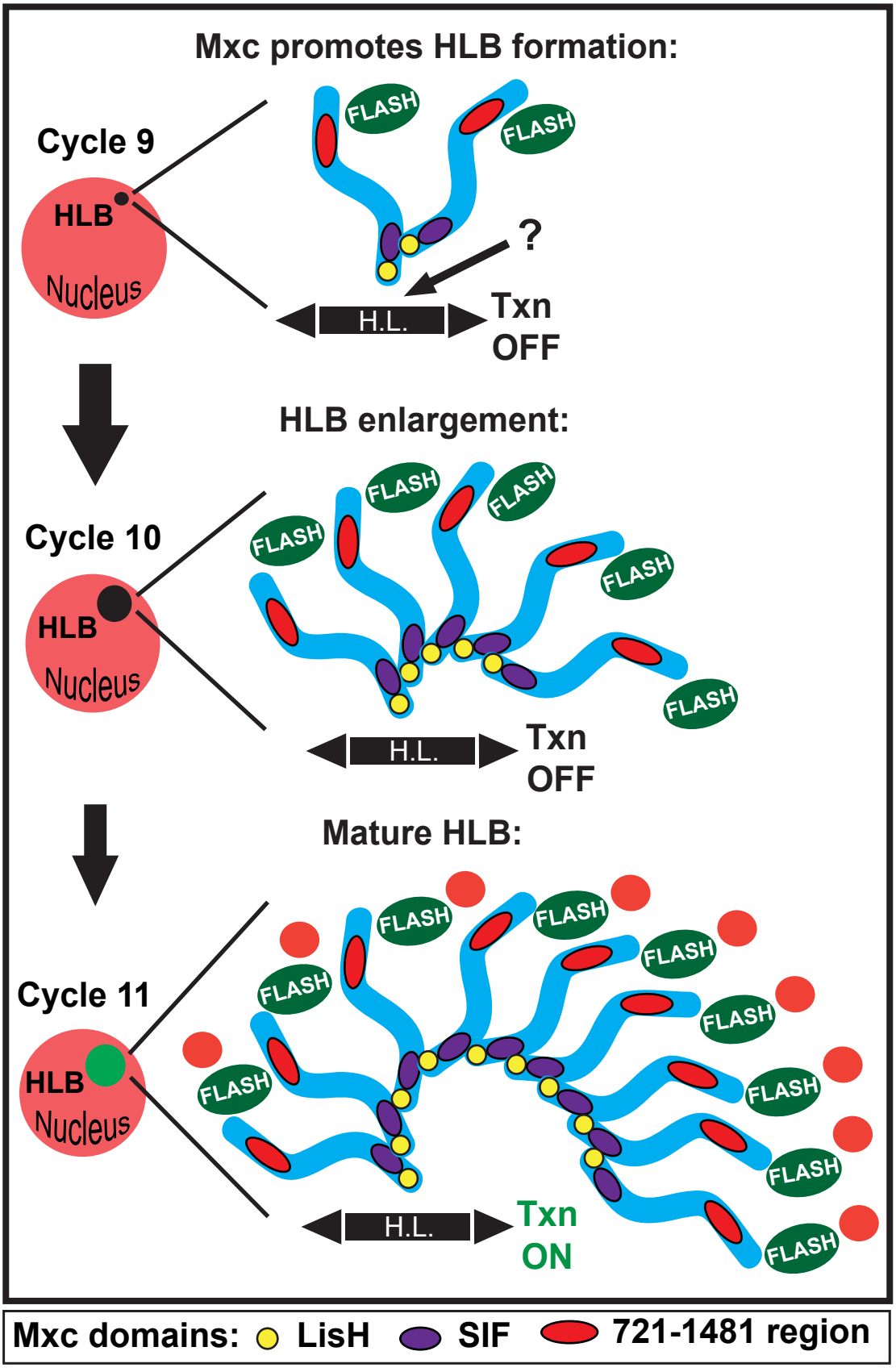

Mxc promotes HLB formation:
Cycle 9
HLB
Nucleus
FLASH
FLASH
?
H.L.
Txn
OFF
HLB enlargement:
Cycle 10
HLB
Nucleus
FLASH
FLASH
FLASH
FLASH
FLASH
H.L.
Txn
OFF
Mature HLB:
Cycle 11
HLB
Nucleus
FLASH
FLASH
FLASH
FLASH
FLASH
FLASH
FLASH
FLASH
H.L.
Txn
ON
Mxc domains:
LisH
SIF
721-1481 region

Figure 2.10. HLB assembly model. Mxc along with FLASH first associate with the histone locus (H.L.), but if and how Mxc binds DNA directly or through a DNA-binding protein remains unknown (**?**). Once chromosome-associated, Mxc initiates the formation of a 3-dimensional HLB lattice by accumulating more Mxc molecules from the nucleoplasm via a LisH domain (yellow circle)/SIF domain (purple oval)-dependent self-interaction as early as syncytial nuclear cycle 9. During cycle 10, when histone transcription has not yet begun (black nuclear circle), an ~4-fold enlargement of the HLB lattice is observed, which requires Mxc's N-terminus (LisH and SIF domains) and a region between amino acids 721 and 1481 (red oval). During cycle 11, when zygotic histone gene expression begins (green nuclear circle), the HLB lattice further increases its size and other histone mRNA biosynthetic factors are recruited to the HLB (orange circles).

In contrast to the null allele, hypomorphic *mxc* mutant embryos (mxc^{G43} and mxc^{G46}) develop to adults and hence are capable of supporting histone mRNA biosynthesis, consistent with previous observations (Landais et al., 2014). In ovaries the 1642 amino acid Mxc^{G46} protein fails to recruit FLASH to HLBs (Rajendra et al., 2011), and results in accumulation of small amounts of misprocessed histone H3 mRNA (D. Tatomer, E.T, W.F.M, and R.J.D., submitted). Here we report that unprocessed histone H3 RNA accumulates at the histone locus in mxc^{G43} and mxc^{G46} mutant embryos. This nascent, unprocessed H3 RNA was detected by *in situ* hybridization with a probe derived from sequence downstream of the normal H3 mRNA 3' end. We do not detect these unprocessed RNAs in wild type embryos. Thus, loss of the last 195 amino acids from Mxc may reduce the efficiency of normal histone mRNA 3' end formation.

Conclusions

Several lines of evidence suggest that proteins with multiple protein-protein interaction domains mediate the localized concentration of components that give rise to NBs (Foray et al., 2003; Good et al., 2011; Matera et al., 2009; Zaidi et al., 2007). NB components can exchange with the nucleoplasm (Deryusheva and Gall, 2004; Dundr et al., 2004), suggesting there are multiple relatively weak protein-protein interactions between components of nuclear bodies, a property that is shared with other cellular bodies (e.g. P-bodies and stress granules in the cytoplasm) (Dundr and Misteli, 2010; Voronina et al., 2011). Together with our previous work (Salzler et al., 2013; White et al., 2011), we propose a model in which Mxc together with FLASH helps drive formation of a large (i.e. visible by light microscopy) 3-dimensional lattice, the HLB, containing components necessary for efficient transcription and processing of histone mRNA (Figure 2.10). Gaining additional insight into the biogenesis of NBs will further our understanding of the assembly and function of regulatory machineries required to effectively control gene expression, and is crucial to understand how these complex structures respond to diverse physiological stimuli during normal and pathological circumstances.

Materials and Methods

Immunofluorescence

For embryos, larval brains, larval salivary glands and ovaries, the following primary antibodies were used: monoclonal mouse MPM-2, (1:2000, Millipore); chicken anti-GFP (1:1000, Millipore); affinity-purified polyclonal rabbit anti-FLASH (1:2000),

affinity purified rabbit and guinea pig anti-Mxc (1:2000) (White et al., 2011; Yang et al., 2009), and monoclonal mouse anti-Lamin (DSHB). For S2 cells, immunostaining was performed as described (White et al., 2011). The secondary antibodies used (1:2000) in all experiments were: goat anti-rabbit IgG labeled with Alexa fluor 488 (Abcam) or Cy5 (Jackson); goat anti-mouse IgG Cy3 (Jackson); donkey anti-chicken Cy2 (Jackson); and goat anti-guinea pig IgG Cy3 or Cy5 (Jackson). DNA was detected by incubating tissue in 1ug/ml DAPI (DAKO Corporation, Carpinteria, CA) for 1 min. Embryos were dechorionated, fixed in a 1:1 mixture of 7% formaldehyde:heptane for 20 minutes and incubated with primary overnight at 4ºC and secondary for 1 hour at 25ºC. Brains and salivary glands were dissected from third instar larvae in Grace's medium, fixed in 4% paraformaldehyde and 3.7% formaldehyde, respectively, for 20 minutes. Brains were permeabilized in 0.2% Tween-20 for 20 minutes prior to immunostaining.

Amylose Pull-down assay

6x-Histidine-tagged MBP and MBP-Mxc^{101} proteins (pDest-566 Gateway Destination vector. Addgene plasmid 11517 courtesy of Dr. Dominic Esposito, Frederick National Laboratory for Cancer Research) were expressed in *E. coli* and subsequently affinity purified through Nickel-NTA resin columns (Qiagen). Fragments of Mxc were labeled with ^{35}S-met *by in vitro* translation using Promega's TNT coupled rabbit reticulocyte kit. 5 μg of recombinant MBP proteins were incubated at 4ºC with pre-equilibrated amylose resin (GE lifesciences) in 100 μl of TEN100 buffer (20 mM Tris [pH 7.5], 0.1 mM EDTA, 100 mM NaCl). Unbound protein was removed by washing 2X with 250 μL TEN100. 10 μL of *in vitro* translated protein was added to beads along with

10 μL of 10X TEN100 buffer, 14 μL of GDB buffer (10% Glycerol, 10 mM DTT, 0.05 mg/mL BSA) and 76 μL of dH_2O. Proteins were allowed to bind for 2 hours at 4ºC while rotating. Amylose beads were washed 4 times with 1 mL of TEN100 buffer. 25 μL of 2X SDS loading dye (4% SDS, 10% β-mercaptoethanol, 0.125 M Tris [pH 6.8], 20% glycerol, 0.2% Bromophenol Blue) was added to beads and boiled for 10 minutes. The supernatant was loaded onto an SDS-PAGE gel. Gels were stained with coomassie blue to confirm pull-down of recombinant MBP protein. Gels were dried and visualized by autoradiography.

Imaging

Confocal images for embryo in situ hybridization were obtained at a zoom of 1.0-5.0 with a 20x PlanNeofluar (NA 0.5) and 40x PlanApochromat (NA 1.3) objectives using the ZEN data acquisition software on a laser-scanning confocal microscope (510; Carl Zeiss). Confocal images for embryo, adult and larval tissue immunostaining and high magnification embryo in situ hyrbidization were obtained at a zoom of 1.0-5.0 with a 63x PlanAchromat (NA 1.4) objective using the ZEN data acquisition software on a laser-scanning confocal microscope (710; Carl Zeiss). Confocal images for *Drosophila* S2 cells were taken at a zoom of 2.0-5.0 with a 40x (NA 1.25) Plan Apochromat objective on a laser-scanning confocal microscope (SP5; Leica, Exton, PA).

For live imaging, transgenic flies harboring GFP-Mxc were generated and crossed to flies carrying a transgenic histone H2Av variant fused to the red fluorescent protein tag (H2Av-RFP). Female virgins carrying one copy of GFP-Mxc (White et al., 2011) and one of H2Av-RFP (Poulton et al., 2014) were selected and crossed to their male siblings

to assure one copy of each transgene maternally supplied to the embryos to be analyzed. Syncytial Drosophila embryos were mounted on a lumox porous-surfaced dish (Sarstedt) and covered with halocarbon oil 700 (Sigma). Images from the surface of the embryo body were acquired at approximately 21°C on a Nikon TE2000-E microscope with Visitech Infinity-Hawk multi-point array scanner, using 100× Nikon objectives, a Ludl emission filter wheel with Semrock filters, and Hamamatsu ORCA R2 camera. Excitation was by 491 nm (GFP) and 561 nm (RFP) lasers. Movies and stills were processed in ImageJ. Fluorescence intensity was calculated for all foci on a single z plane with the highest integrated intensity values in the region of interest. A circle was drawn around each focus and in areas inside 5 nuclei without fluorescence on the same z plane to be used for background readings. To calculate the corrected total focus fluorescence (CTFF) using ImageJ software, we analyzed data from three embryos representing three independent experiments and adapted the following formula: CTFF = Integrated Density – (Area of selected focus x Mean fluorescence of background readings) (Burgess et al., 2010; Potapova et al., 2011).

Embryo *in situ* hybridization

w^{1118}, $Slbp^{15}$, and *mxc* mutant embryos were collected and aged at room temperature until 8-10 hours old. Embryos were fixed in a 1:1 mixture of 7% formaldehyde/heptane for 20 minutes and rehydrated in 1x PBS 0.1% Tween-20. Histone H3 transcripts were detected by fluorescent in situ hybridization using digoxigenin-labeled H3-coding or H3-ds probes (Lanzotti et al., 2002; White et al., 2007).

Molecular biology

Mxc fragments used for immunostaining and live-imaging experiments were expressed in *Drosophila* cultured S2 cells or as transgenes in Gateway-compatible vectors (Carnegie Institution) as previously described (White et al., 2011). Mxc fragments used for pull-down assays were all expressed in the pxFRM vector (Lyons et al., 2014). The primers used to amplify all Mxc fragments are listed in supplemental Table 1.

Western Blotting

Ovary protein lysates were obtained from w^{1118}; *GFP-Mxc*, w^{1118}; *GFP-Mxc*$^{LisH\text{-}AAA}$, and w^{1118}; *GFP-Mxc*$^{SIF\text{-}AAA}$ female flies dissected in 1x Tris PBS. Ovaries were snap-froze in dry ice and ethanol for 10 minutes and stored at -20ºC overnight. Ovaries were resuspended in buffer containing 4% SDS and dissociated with 20 strokes of a dounce homogenizer on ice. Equal amounts of protein were run on a 7.5% acrylamide gel (BioRad) and then transferred to a polyvinylidene fluoride (PVDF) membrane (Millipore) pre-soaked in methanol for 15 minutes at room temperature. Membranes were incubated overnight at 4ºC in primary rabbit anti-GFP (Abcam) antibody to detect GFP-tagged proteins (~250 kDa in size) and in primary mouse anti-Lamin (DSHB) antibody to detect *Drosophila* Lamin (74 kDa in size). Lamin and an anti-GFP antibody cross-reacting band were used as loading controls. Subsequently, membranes were incubated in secondary antibody HRP-conjugated donkey-anti rabbit IgG (GE) and HRP-conjugated goat-anti mouse IgG (GE) for 2 hours at room temperature to detect GFP and Lamin, respectively. The signal was enhanced using Enhance Signal West Dura (Thermo Scientific) and

visualized using an 8-megapixel EX Sigma camera in a BioSpectrum imaging system (UVP) after a 25-minute exposure.

Northern Blotting

Northern blotting was performed using a 6% 7M urea acrylamide gel to resolve histone mRNAs and 7SK RNA (Nguyen et al., 2012) as previously described (Mullen and Marzluff, 2008).

Statistical Analysis

Standard error of the mean (SEM) calculated by dividing standard deviation (SD) by the square root of number of samples (n). Statistical significance between different samples was calculated using the Student's t-test.

Computational Analysis of Mxc's self-interaction

The structure of the LisH domain of TBL1X (PDB ID 2XTC) was identified by HHpred (http://toolkit.tuebingen.mpg.de/hhpred) (Soding, 2005) as a structural template for homology modeling of the LisH domain of Mxc using the MODELLER software program (Eswar et al., 2006).

REFERENCES

Adams, J.C. 2002. Characterization of a Drosophila melanogaster orthologue of muskelin. *Gene*. 297:69-78.

Aravind, L., Landsman, D. 1998. AT-hook motifs identified in a wide variety of

DNA-binding proteins. *Oxford University Press*. 26:4413-4421.

Bian, C., R. Wu, K. Cho, and X. Yu. 2012. Loss of BRCA1-A complex function in RAP80 null tumor cells. *PLoS One*. 7:e40406.

Bulchand, S., S.D. Menon, S.E. George, and W. Chia. 2010. Muscle wasted: a novel component of the Drosophila histone locus body required for muscle integrity. *J Cell Sci*. 123:2697-2707.

Burgess, A., S. Vigneron, E. Brioudes, J.C. Labbe, T. Lorca, and A. Castro. 2010. Loss of human Greatwall results in G2 arrest and multiple mitotic defects due to deregulation of the cyclin B-Cdc2/PP2A balance. *Proc Natl Acad Sci U S A*. 107:12564-12569.

Cerna, D., and D.K. Wilson. 2005. The structure of Sif2p, a WD repeat protein functioning in the SET3 corepressor complex. *J Mol Biol*. 351:923-935.

Cortese, M.S., V.N. Uversky, and A.K. Dunker. 2008. Intrinsic disorder in scaffold proteins: getting more from less. *Progress in biophysics and molecular biology*. 98:85-106.

Deryusheva, S., and J.G. Gall. 2004. Dynamics of coilin in Cajal bodies of the Xenopus germinal vesicle. *Proc Natl Acad Sci U S A*. 101:4810-4814.

Deryusheva, S., and J.G. Gall. 2009. Small Cajal body-specific RNAs of Drosophila function in the absence of Cajal bodies. *Mol Biol Cell*. 20:5250-5259.

Dundr, M. 2012. Nuclear bodies: multifunctional companions of the genome. *Curr Opin Cell Biol*. 24:415-422.

Dundr, M., M.D. Hebert, T.S. Karpova, D. Stanek, H. Xu, K.B. Shpargel, U.T. Meier, K.M. Neugebauer, A.G. Matera, and T. Misteli. 2004. In vivo kinetics of Cajal body components. *J Cell Biol*. 164:831-842.

Dundr, M., and T. Misteli. 2001. Functional architecture in the cell nucleus. *Biochem J*. 356:297-310.

Dundr, M., and T. Misteli. 2010. Biogenesis of nuclear bodies. *Cold Spring Harb Perspect Biol*. 2:a000711.

Emes, R.D., and C.P. Ponting. 2001. A new sequence motif linking lissencephaly, Treacher Collins and oral-facial-digital type 1 syndromes, microtubule dynamics and cell migration. *Human molecular genetics*. 10:2813-2820.

Eswar, N., B. Webb, M.A. Marti-Renom, M.S. Madhusudhan, D. Eramian, M.Y. Shen, U. Pieper, and A. Sali. 2006. Comparative protein structure modeling using Modeller. *Current protocols in bioinformatics / editoral board, Andreas D. Baxevanis ... [et al.]*. Chapter 5:Unit 5 6.

Foray, N., D. Marot, A. Gabriel, V. Randrianarison, A.M. Carr, M. Perricaudet, A. Ashworth, and P. Jeggo. 2003. A subset of ATM- and ATR-dependent phosphorylation events requires the BRCA1 protein. *EMBO J*. 22:2860-2871.

Gall, J.G. 2000. Cajal bodies: the first 100 years. *Annu Rev Cell Dev Biol*. 16:273-300.

Gerlitz, G., E. Darhin, G. Giorgio, B. Franco, and O. Reiner. 2005. Novel functional features of the Lis-H domain: role in protein dimerization, half-life and cellular localization. *Cell Cycle*. 4:1632-1640.

Godfrey, A.C., J.M. Kupsco, B.D. Burch, R.M. Zimmerman, Z. Dominski, W.F. Marzluff, and R.J. Duronio. 2006. U7 snRNA mutations in Drosophila block histone pre-mRNA processing and disrupt oogenesis. *Rna*. 12:396-409.

Good, M.C., J.G. Zalatan, and W.A. Lim. 2011. Scaffold proteins: hubs for controlling the flow of cellular information. *Science*. 332:680-686.

Harrer, M., H. Luhrs, M. Bustin, U. Scheer, and R. Hock. 2004. Dynamic interaction of HMGA1a proteins with chromatin. *J Cell Sci*. 117:3459-3471.

Hebert, M.D., and A.G. Matera. 2000. Self-association of coilin reveals a common theme in nuclear body localization. *Mol Biol Cell*. 11:4159-4171.

Kim, M.H., D.R. Cooper, A. Oleksy, Y. Devedjiev, U. Derewenda, O. Reiner, J. Otlewski, and Z.S. Derewenda. 2004. The structure of the N-terminal domain of the product of the lissencephaly gene Lis1 and its functional implications. *Structure*. 12:987-998.

Klingauf, M., D. Stanek, and K.M. Neugebauer. 2006. Enhancement of U4/U6 small nuclear ribonucleoprotein particle association in Cajal bodies predicted by mathematical modeling. *Mol Biol Cell*. 17:4972-4981.

Knoblich, J.A., K. Sauer, L. Jones, H. Richardson, R. Saint, and C.F. Lehner. 1994. Cyclin E controls S phase progression and its down-regulation during Drosophila embryogenesis is required for the arrest of cell proliferation. *Cell*. 77:107-120.

Kobayashi, N., J. Yang, A. Ueda, T. Suzuki, K. Tomaru, M. Takeno, K. Okuda, and Y. Ishigatsubo. 2007. RanBPM, Muskelin, p48EMLP, p44CTLH, and the armadillo-repeat proteins ARMC8alpha and ARMC8beta are components of the CTLH complex. *Gene*. 396:236-247.

Landais, S., C. D'Alterio, and D.L. Jones. 2014. Persistent replicative stress alters polycomb phenotypes and tissue homeostasis in Drosophila melanogaster. *Cell reports*. 7:859-870.

Lanzotti, D.J., H. Kaygun, X. Yang, R.J. Duronio, and W.F. Marzluff. 2002. Developmental control of histone mRNA and dSLBP synthesis during Drosophila embryogenesis and the role of dSLBP in histone mRNA 3' end processing in vivo. *Mol Cell Biol*. 22:2267-2282.

Lanzotti, D.J., J.M. Kupsco, W.F. Marzluff, and R.J. Duronio. 2004a. string(cdc25) and cyclin E are required for patterned histone expression at different stages of Drosophila embryonic development. *Dev Biol*. 274:82-93.

Lanzotti, D.J., J.M. Kupsco, X.C. Yang, Z. Dominski, W.F. Marzluff, and R.J. Duronio. 2004b. Drosophila stem-loop binding protein intracellular localization is mediated by phosphorylation and is required for cell cycle-regulated histone mRNA expression. *Mol Biol Cell*. 15:1112-1123.

Li, Z., M.R. Johnson, Z. Ke, L. Chen, and M.A. Welte. 2014. Drosophila lipid droplets buffer the H2Av supply to protect early embryonic development. *Curr Biol*. 24:1485-1491.

Liu, J.L., C. Murphy, M. Buszczak, S. Clatterbuck, R. Goodman, and J.G. Gall. 2006. The Drosophila melanogaster Cajal body. *J Cell Biol*. 172:875-884.

Lyons, S.M., A.S. Ricciardi, A.Y. Guo, C. Kambach, and W.F. Marzluff. 2014. The C-terminal extension of Lsm4 interacts directly with the 3' end of the histone mRNP and is required for efficient histone mRNA degradation. *RNA*. 20:88-102.

Machyna, M., S. Kehr, K. Straube, D. Kappei, F. Buchholz, F. Butter, J. Ule, J. Hertel, P.F. Stadler, and K.M. Neugebauer. 2014. The Coilin Interactome Identifies Hundreds of Small Noncoding RNAs that Traffic through Cajal Bodies. *Mol Cell*. 56:389-399.

Marzluff, W.F., E.J. Wagner, and R.J. Duronio. 2008. Metabolism and regulation of canonical histone mRNAs: life without a poly(A) tail. *Nat Rev Genet*. 9:843-854.

Matera, A.G. 1999. Nuclear bodies: multifaceted subdomains of the interchromatin space. *Trends Cell Biol*. 9:302-309.

Matera, A.G., M. Izaguire-Sierra, K. Praveen, and T.K. Rajendra. 2009. Nuclear bodies: random aggregates of sticky proteins or crucibles of macromolecular assembly? *Dev Cell*. 17:639-647.

Menssen, R., J. Schweiggert, J. Schreiner, D. Kusevic, J. Reuther, B. Braun, and D.H. Wolf. 2012. Exploring the topology of the Gid complex, the E3 ubiquitin ligase involved in catabolite-induced degradation of gluconeogenic enzymes. *J Biol Chem*. 287:25602-25614.

Mikolajka, A., X. Yan, G.M. Popowicz, P. Smialowski, E.A. Nigg, and T.A. Holak. 2006. Structure of the N-terminal domain of the FOP (FGFR1OP) protein and implications for its dimerization and centrosomal localization. *J Mol Biol*. 359:863-875.

Misteli, T. 2001. The concept of self-organization in cellular architecture. *J Cell Biol*. 155:181-185.

Misteli, T. 2005. Concepts in nuclear architecture. *Bioessays*. 27:477-487.

Mowry, K.L., and J.A. Steitz. 1987. Identification of the human U7 snRNP as one of several factors involved in the 3' end maturation of histone premessenger RNA's. *Science*. 238:1682-1687.

Mullen, T.E., and W.F. Marzluff. 2008. Degradation of histone mRNA requires oligouridylation followed by decapping and simultaneous degradation of the mRNA both 5' to 3' and 3' to 5'. *Genes Dev*. 22:50-65.

Nguyen, D., B.J. Krueger, S.C. Sedore, J.E. Brogie, J.T. Rogers, T.K. Rajendra, A. Saunders, A.G. Matera, J.T. Lis, P. Uguen, and D.H. Price. 2012. The Drosophila 7SK snRNP and the essential role of dHEXIM in development. *Nucleic acids research*. 40:5283-5297.

Novotny, I., M. Blazikova, D. Stanek, P. Herman, and J. Malinsky. 2011. In vivo kinetics of U4/U6.U5 tri-snRNP formation in Cajal bodies. *Mol Biol Cell*. 22:513-523.

Nussinov, R., B. Ma, and C.J. Tsai. 2013. A broad view of scaffolding suggests that scaffolding proteins can actively control regulation and signaling of multienzyme complexes through allostery. *Biochimica et biophysica acta*. 1834:820-829.

Parada, L.A., S. Sotiriou, and T. Misteli. 2004. Spatial genome organization. *Exp Cell Res*. 296:64-70.

Potapova, T.A., S. Sivakumar, J.N. Flynn, R. Li, and G.J. Gorbsky. 2011. Mitotic progression becomes irreversible in prometaphase and collapses when Wee1 and Cdc25 are inhibited. *Molecular Biology of the Cell*. 22:1191-1206.

Poulton, J.S., J.C. Cuningham, and M. Peifer. 2014. Acentrosomal Drosophila epithelial cells exhibit abnormal cell division, leading to cell death and compensatory proliferation. *Dev Cell*. 30:731-745.

Rajendra, T.K., K. Praveen, and A.G. Matera. 2011. Genetic Analysis of Nuclear Bodies: From Nondeterministic Chaos to Deterministic Order. *Cold Spring Harb Symp Quant Biol*. in press.

Reeves, R., Nissen, M. 1990. The A-T-DNA-binding Domain of Mammalian High Mobility Group I Chromosomal Proteins. *The Journanl of Biological Chemistry*. 265:8573-8582.

Remillieux-Leschelle, N., P. Santamaria, and N.B. Randsholt. 2002. Regulation of larval hematopoiesis in Drosophila melanogaster: a role for the multi sex combs gene. *Genetics*. 162:1259-1274.

Saget, O., F. Forquignon, P. Santamaria, and N.B. Randsholt. 1998. Needs and targets for the multi sex combs gene product in Drosophila melanogaster. *Genetics*. 149:1823-1838.

Salemi, L.M., A.W. Almawi, K.J. Lefebvre, and C. Schild-Poulter. 2014. Aggresome formation is regulated by RanBPM through an interaction with HDAC6. *Biology open*. 3:418-430.

Salzler, H.R., D.C. Tatomer, P.Y. Malek, S.L. McDaniel, A.N. Orlando, W.F. Marzluff, and R.J. Duronio. 2013. A sequence in the Drosophila H3-H4 Promoter triggers histone locus body assembly and biosynthesis of replication-coupled histone mRNAs. *Dev Cell*. 24:623-634.

Santamaria, P., and N.B. Randsholt. 1995. Characterization of a region of the X chromosome of Drosophila including multi sex combs (mxc), a Polycomb group gene which also functions as a tumour suppressor. *Mol Gen Genet*. 246:282-290.

Soding, J. 2005. Protein homology detection by HMM-HMM comparison. *Bioinformatics*. 21:951-960.

Strub, K., and M.L. Birnstiel. 1986. Genetic complementation in the Xenopus oocyte: co-expression of sea urchin histone and U7 RNAs restores 3' processing of H3 pre-mRNA in the oocyte. *EMBO J*. 5:1675-1682.

Strzelecka, M., S. Trowitzsch, G. Weber, R. Luhrmann, A.C. Oates, and K.M. Neugebauer. 2010. Coilin-dependent snRNP assembly is essential for zebrafish embryogenesis. *Nat Struct Mol Biol*. 17:403-409.

Sullivan, E., C. Santiago, E.D. Parker, Z. Dominski, X. Yang, D.J. Lanzotti, T.C. Ingledue, W.F. Marzluff, and R.J. Duronio. 2001. Drosophila stem loop binding protein coordinates accumulation of mature histone mRNA with cell cycle progression. *Genes Dev*. 15:173-187.

Sun, Z., A.V. Smrcka, and S. Chen. 2013. WDR26 functions as a scaffolding protein to promote Gbetagamma-mediated phospholipase C beta2 (PLCbeta2) activation in leukocytes. *J Biol Chem*. 288:16715-16725.

Swanhart, L., J. Kupsco, and R.J. Duronio. 2005. Developmental control of growth and cell cycle progression in Drosophila. *Methods Mol Biol*. 296:69-94.

Umeda, M., H. Nishitani, and T. Nishimoto. 2003. A novel nuclear protein, Twa1, and Muskelin comprise a complex with RanBPM. *Gene*. 303:47-54.

Voronina, E., G. Seydoux, P. Sassone-Corsi, and I. Nagamori. 2011. RNA granules in germ cells. *Cold Spring Harb Perspect Biol*. 3.

Wei, Y., J. Jin, and J.W. Harper. 2003. The cyclin E/Cdk2 substrate and Cajal body component p220(NPAT) activates histone transcription through a novel LisH-like domain. *Mol Cell Biol*. 23:3669-3680.

White, A.E., B.D. Burch, X.C. Yang, P.Y. Gasdaska, Z. Dominski, W.F. Marzluff, and R.J. Duronio. 2011. Drosophila histone locus bodies form by hierarchical recruitment of components. *J Cell Biol*. 193:677-694.

White, A.E., M.E. Leslie, B.R. Calvi, W.F. Marzluff, and R.J. Duronio. 2007. Developmental and cell cycle regulation of the Drosophila histone locus body. *Mol Biol Cell*. 18:2491-2502.

Yang, X.C., B.D. Burch, Y. Yan, W.F. Marzluff, and Z. Dominski. 2009. FLASH, a proapoptotic protein involved in activation of caspase-8, is essential for 3' end processing of histone pre-mRNAs. *Mol Cell*. 36:267-278.

Ye, X., Y. Wei, G. Nalepa, and J.W. Harper. 2003. The cyclin E/Cdk2 substrate p220(NPAT) is required for S-phase entry, histone gene expression, and Cajal body maintenance in human somatic cells. *Mol Cell Biol*. 23:8586-8600.

Zaidi, S.K., D.W. Young, A. Javed, J. Pratap, M. Montecino, A. van Wijnen, J.B. Lian, J.L. Stein, and G.S. Stein. 2007. Nuclear microenvironments in biological control and cancer. *Nature reviews. Cancer*. 7:454-463.

Zhao, J., B.K. Kennedy, B.D. Lawrence, D.A. Barbie, A.G. Matera, J.A. Fletcher, and E. Harlow. 2000. NPAT links cyclin E-Cdk2 to the regulation of replication-dependent histone gene transcription. *Genes Dev*. 14:2283-2297.

CHAPTER III

EXPLORING THE ROLE OF CYCE/CDK2-DEPENDENT MXC PHOSPHORYLATION IN HLB ASSEMBLY AND HISTONE mRNA BIOSYNTHESIS

This chapter is a compilation of data generated throughout the past 3 years of my graduate career and also of research that I am currently conducting in the laboratory of my advisor, Dr. Robert Duronio. He and I designed the experiments. Dr. Robert Duronio, Dr. William Marzluff, and I analyzed the data. I have performed all of the experiments and designed all of the figures and schematic representations included in the present chapter.

Introduction

We have previously shown that Mxc is recognized by MPM-2 antibody in immunofluorescence assays and also that Mxc directly reacts with MPM-2 in Western blot analysis. Combined, these data strongly suggest that Mxc is a CycE/Cdk2 phosphorylation target (White et al., 2011). We have shown that Mxc is crucial for HLB assembly and histone mRNA biosynthesis and more recently that it plays an integral role in HLB assembly by utilizing multiple domains to accumulate at the histone locus and form a 3-dimenssional lattice required for locally concentrating histone mRNA

biosynthetic factors (Terzo et al., 2015; White et al., 2011). Whether the cell cycle input (CycE/Cdk2-dependent phosphorylation) on Mxc plays a role in HLB assembly and histone mRNA biosynthesis is not known. Mxc consists of 1837 amino acids in length and harbors multiple functional domains required for its accumulation at the histone locus, the enlargement of the HLB during early cell cycles of *Drosophila* embryogenesis, and also for recruitment of important histone pre-mRNA processing factors to the HLB, such as FLASH (Terzo et al., 2015). In addition, Mxc displays several putative CycE/Cdk2 phosphorylation sites throughout its entire protein sequence (further discussed below). Post-translational modifications (PTMs) are known to affect the stability, turnover, interaction potential, and localization of proteins within the cell (Seet et al., 2006). These aspects of PTMs are particularly relevant for proteins involved in regulation and signaling, as are many proteins harboring IDRs (Uversky, 2013b; Vucetic et al., 2007). The conformational flexibility of disordered regions as display sites for PTMs provides advantages over structured regions. For instance, flexibility facilitates the deposition of PTMs by enabling transient, but yet specific interaction with catalytic sites of modifying enzymes. This is because, upon binding, a flexible, disordered region loses more conformational freedom (i.e., entropy), which reduces the overall free energy of binding, leading to weaker and more transient binding as compared to a folded protein region that interacts with equal strength (i.e., the same biding enthalpy or equal specificity) (van der Lee et al., 2014). Proteins harboring IDRs are highly enriched for short motifs where PTMs are often located, which underlines the importance of intrinsic disorder as PTM display sites (Perkins et al., 2010). Interestingly, several lines of research have shown that multivalent proteins with the ability to self-aggregate (forming

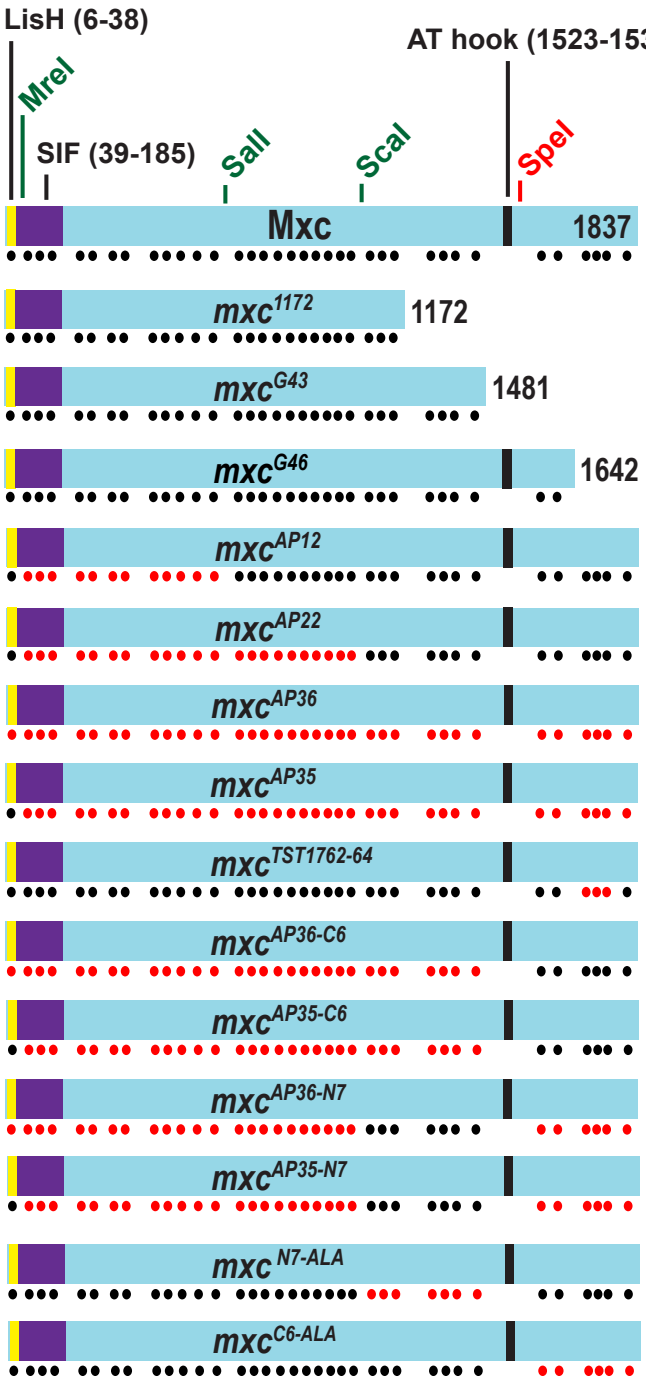
hydrogels) are highly susceptible to post-translational modifications, such as phosphorylation (Courchaine and Neugebauer, 2015; Kwon et al., 2013; Kwon et al., 2014). For instance, a recent report demonstrates that FUS (fused in sarcoma) protein carries several phosphorylation sites in its low complexity domain and that increasing phosphorylation of these sites impedes hydrogel retention (*in vitro*) of FUS (Han et al., 2012). In light of these findings, it is thought that the phenomenon of post-translational modification, e.g., phosphorylation, greatly impacts the interactions underpinning NB assembly (Hennig et al., 2015). Combined, these facts prompted us to develop powerful genetic tools to gain understanding on how Mxc's CycE/Cdk2-dependent phosphorylation contributes to HLB assembly and histone mRNA biosynthesis.

Identifying Mxc amino acid(s) subject to CycE/Cdk2-dependent phosphorylation

We set out to identify the amino acid(s) of Mxc subject to phosphorylation by the CycE/Cdk2 complex. To this end, we employed the biomarker MPM-2 antibody, which has been widely used for almost two decades to detect CycE/Cdk2 activity (Calvi et al., 1998; Davis et al., 1983). Overexpression of CycE was shown to coincide with an increase in MPM-2 foci staining, which strongly suggested that MPM-2 signal was an accurate indicator of CycE/Cdk2-dependent phosphorylation (Calvi et al., 1998). Our laboratory demonstrated that MPM-2 recognizes a target that co-localizes with known HLB factors (White et al., 2007) and later on demonstrated that MPM-2 directly reacts with Mxc on Western Blots and labels Mxc exclusively during S phase (White et al., 2011). However, where on the Mxc protein the CycE/Cdk2 phosphoepitopes lie remains unknown.

The consensus motif that Cdks typically phosphorylate is as follows: Ser/Thr-Pro-X-Lys/Arg) where "X" represents any known amino acid (Ubersax and Ferrell, 2007). We searched throughout the entire full-length Mxc protein for prolines (Pro) that were preceded by either serines (Ser) or threonines (Thr) and found 36 putative Ser/Pro and Thr/Pro CycE/Cdk2 phosphorylation sites. To begin to explore *in vivo* where Mxc's CycE/Cdk2 phosphoepitopes lie, we engineered two DNA fragments of Mxc containing 12 and 22 (out of 36) putative phosphorylation sites, in which we mutagenized the serines and threonines to alanines (Ala) (GenScript). We based our decision of choosing these 12 and 22 sites on our previous data, which strongly suggested that Mxc was phosphorylated on its N-terminal region (White et al., 2011) and also out of necessity, as we found unique restriction sites encompassing regions containing the above-mentioned number of consecutive sites. The synthesized DNA fragments containing the mutagenized phosphorylation sites were cloned into the pENTR vector that contained the full-length Mxc by replacing a number of wild-type sites with its corresponding set of mutated phosphorylation sites (12 or 22). In doing so, we generated two Mxc phosphorylation mutants that displayed 12 or 22 sites, changed to alanines, and named them AP12 and AP22, respectively (Figure 3.1). To first test for expression of these two Mxc phosphorylation mutants, we cloned them into pAVW, a *Drosophila* gateway vector carrying an N-terminal Venus (GFP modified) tag driven by the Actin promoter. We transfected pAVW::mxc^{AP12} and pAVW::mxc^{AP22} into *Drosophila*-cultured S2 cells and asked if both fusion proteins, GFP-Mxc^{AP12} and GFP-Mxc^{AP22}, were expressed, localized at the histone locus, and were recognized by MPM-2 antibody. Indeed, we observed that both GFP-Mxc^{AP12} and GFP-Mxc^{AP22} were expressed and concentrated at the histone

locus, as evidenced by the GFP (+) signal co-localizing with HLB markers, and also displayed MPM-2 (+) foci (data not shown). A caveat to this experiment is that we maintained endogenous Mxc present at all times.



	Mxc Accumulation in the HLB		MPM2 Focus
Protein name	wt Mxc present	wt Mxc absent	wt Mxc absent
Mxc	+	+	+
1172	+	ND	ND
G43	ND	+	+
G46	+	+	+
AP12	+	+	+
AP22	+	+	+
AP36	-	ND	ND
AP35	-	ND	ND
62-64	+	ND	ND
36-C6	-	ND	ND
35-C6	-	ND	ND
36-N7	-	ND	ND
35-N7	-	ND	ND
N7-ALA	+	ND	ND
C6-ALA	+	ND	ND

Figure 3.1. HLB accumulation and MPM-2 reactivity of Mxc's CycE/Cdk2 phosphorylation mutants. On the left, all Mxc CycE/Cdk2 phosphorylation mutants generated to date are listed. Full-length Mxc (1873 amino acids in length) is at the top of the figure. Yellow rectangle represents the LisH domain (aa 6-38). Purple rectangle represents the novel SIF domain (aa 39-185). Black rectangle represents the AT hook domain (aa 1523-1535). Three intrinsic restriction enzyme sites are labeled in green: MreI, SalI, and ScaI. One extrinsic restriction enzyme site is labeled in red: SpeI site generated by site directed mutagenesis (S.D.M) for cloning purposes. Black circles symbolize wild-type serines or threonines. Red circles symbolize serines or threonines mutagenized to alanines. On the right, a table describing name, accumulation in the HLB when endogenous Mxc is present or absent, and Mxc's reactivity to MPM-2 antibody in a null *mxc* background for each Mxc molecule depicted on the left. + sign = Mxc phosphorylation positive for accumulation at HLB or for MPM-2 reactivity in immunofluorescence assays. – sign = Mxc phosphorylation negative for accumulation at HLB or for MPM-2 reactivity in immunofluorescence assays. ND = not determined.

Therefore, and based on our most recent findings that Mxc is capable of self-interacting to concentrate at the histone locus (Terzo et al., 2015), and that endogenous Mxc is also recognized by MPM-2 (White et al., 2011), we realized that endogenous Mxc's presence would not allow us to generate conclusive results from these experiments. To circumvent this issue, we decided to generate transgenic mxc^{AP12} and mxc^{AP22} flies, so that we can cross them to null *mxc* flies and ask *in vivo*, in the absence of endogenous *mxc*, if mxc^{AP12} and mxc^{AP22} can concentrate at the histone locus and be recognized by MPM-2. When transgenic mxc^{AP12} and mxc^{AP22} flies were crossed to mxc^{G48} mutant flies (mxc^{G48} is a null mutation that confers lethality at larval stages) both mxc phosphorylation mutants were capable of rescuing lethality given by the null mutation. Hemizygous males for mxc^{G48} and either heterozygous or homozygous for mxc^{AP12} and mxc^{AP22} were fertile.

Interestingly, homozygous females for *mxc*^{G48} and heterozygous or homozygous for *mxc*^{AP12} were fertile, but females with *mxc*^{AP22} segregating in the null *mxc*^{G48} background were sterile. Homozygous *mxc*^{G48} females with either one or two copies of the *mxc*^{AP22} transgene not only failed to lay eggs (very few to no eggs), but they also showed in a

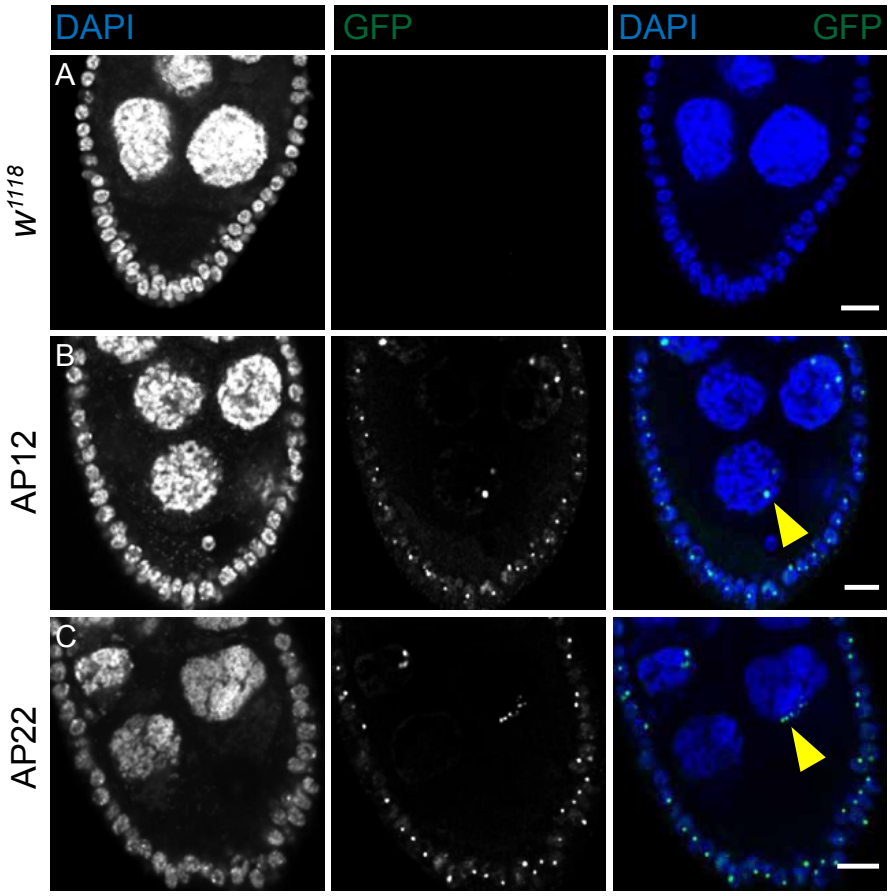


Figure 3.2. Both GFP-Mxc^{AP12} and GFP-Mxc^{Ap22} localize at the histone locus. Stage-8 egg chambers from *w*^{1118}, *mxc*$^{G48/G48}$; *mxc*$^{AP12/12}$ and *mxc*$^{G48/G48}$; *mxc*$^{A22/22}$ ovaries were stained with DAPI and anti-GFP antibody. (A) Wild-type (*w*^{1118}) egg chamber showing no GFP staining. (B) AP12 (GFP-Mxc^{AP12}) egg chamber showing focal localization (yellow arrowhead) typical of wild-type Mxc. (C) AP22 (GFP-Mxc^{AP22}) egg chambers showing a bigger number of GFP (+) foci in nurse cells (yellow arrowhead), when compared to AP12 (B) in the *mxc*^{G48} background. Bars: 10μM.

large number of ovarioles, an abrupt cessation of oogenesis at stage-8 egg chambers, a typical feature of the well-documented female-sterile phenotype in *Drosophila* (Gigliotti et al., 2000). We previously determined that Mxc's maternal load was completely degraded by the 2nd instar larval stage (Terzo et al., 2015). Based on the findings that mxc^{AP12} and mxc^{AP22} rescue lethality given by mxc^{G48}, as both reach adulthood and can be maintained as stocks in our laboratory and on the female-sterile phenotype observed in $mxc^{G48/G48}$; mxc^{AP22} ovaries, we decided to test for GFP-Mxc^{AP12} and GFP-Mxc^{AP22} concentration at the histone locus and for MPM-2 reactivity using adult ovaries. We performed immunofluorescence assays on w^{1118} (wild-type), $mxc^{G48/G48}$; $mxc^{AP12/12}$, and $mxc^{G48/G48}$; $mxc^{A22/22}$ stage-8 egg chambers using anti-GFP antibody to label the fusion proteins and found that both GFP-Mxc^{AP12} and GFP-Mxc^{AP22} displayed focal localization (Figure 3.2A-C). Interestingly, we also began to notice that GFP-Mxc^{AP22} consistently showed a bigger number of GFP (+) foci in nurse cells as GFP-Mxc^{AP12} did (Figure 3.2B and C). When we stained $mxc^{G48/G48}$; $mxc^{AP12/12}$ and $mxc^{G48/G48}$; $mxc^{AP22/22}$ stage-8 egg chambers with anti-FLASH and anti-Mxc antibodies, we found that both were capable of accumulating and co-localizing with FLASH at the histone locus, as evidenced by Mxc/FLASH (+) foci (Figure 3.3A-C). In fact, both GFP-Mxc^{AP12} and GFP-Mxc^{AP22} also co-localized with other known HLB factors, such as Coilin and Mute (data not shown). Ovary staining with MPM-2 showed that both GFP-Mxc^{AP12} and GFP-Mxc^{AP22} were reactive to the antibody, but the former consistently displayed noticeably brighter MPM-2 foci than latter (Figure 3.4A-C). An interesting observation made during our studies is the difference in the number of foci detected in GFP-Mxc^{AP22} ovary stainings when compared

to that of GFP-Mxc^{AP12} (Figures 3.2-4). This finding is consistently observed in ovary stainings and it is specific to GFP-Mxc^{AP22}. Perhaps, this is one factor contributing to the bigger number of Mxc, FLASH, and GFP signal and also to the dimmer MPM-2 signal observed for GFP-Mxc^{AP22} foci. It is worth mentioning that these immunofluorescence assays have been conducted only on adult ovary tissues and we focused our attention on

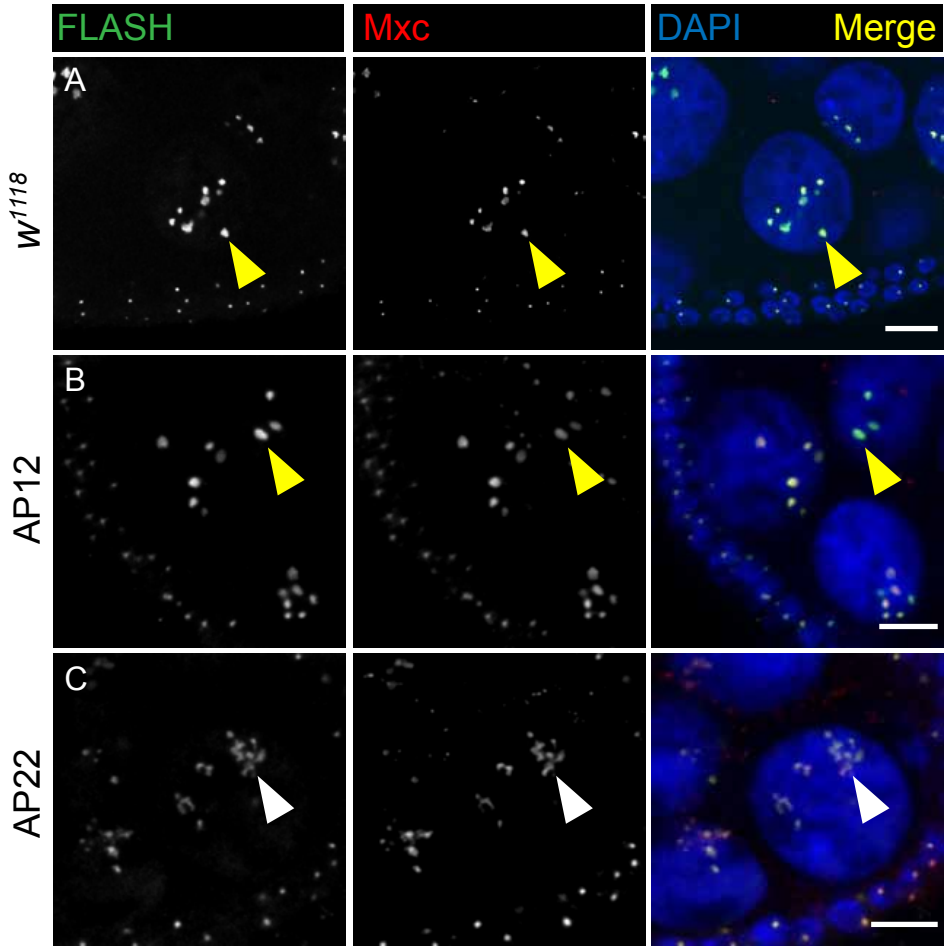


Figure 3.3. Both GFP-Mxc^{AP12} and GFP-Mxc^{Ap22} co-localize at the histone locus with FLASH. Stage-8 egg chambers from w^{1118}, $mxc^{G48/G48}$; $mxc^{AP12/12}$ and $mxc^{G48/G48}$; $mxc^{A22/22}$ ovaries were stained with DAPI, anti-FLASH and anti-Mxc antibodies. (A) Wild- type (w^{1118}) egg chamber showing FLASH and Mxc co-

localizing at the histone locus (yellow arrowhead). (B) AP12 (GFP-Mxc^{AP12}) egg chamber showing FLASH and Mxc co-localization at the histone locus (yellow arrowhead). (C) AP22 (GFP-Mxc^{AP22}) egg chambers showing a bigger number of smaller FLASH/Mxc co-localizing foci (white arrowhead). Bars: 10μM.

germline precursor-derived nurse cells. Nurse cells are known to be highly polyploid cells and it would be interesting to ask if this nuclear phenotype, likely as a result of genomic instability, can also be observed in cells of lower ploidy.

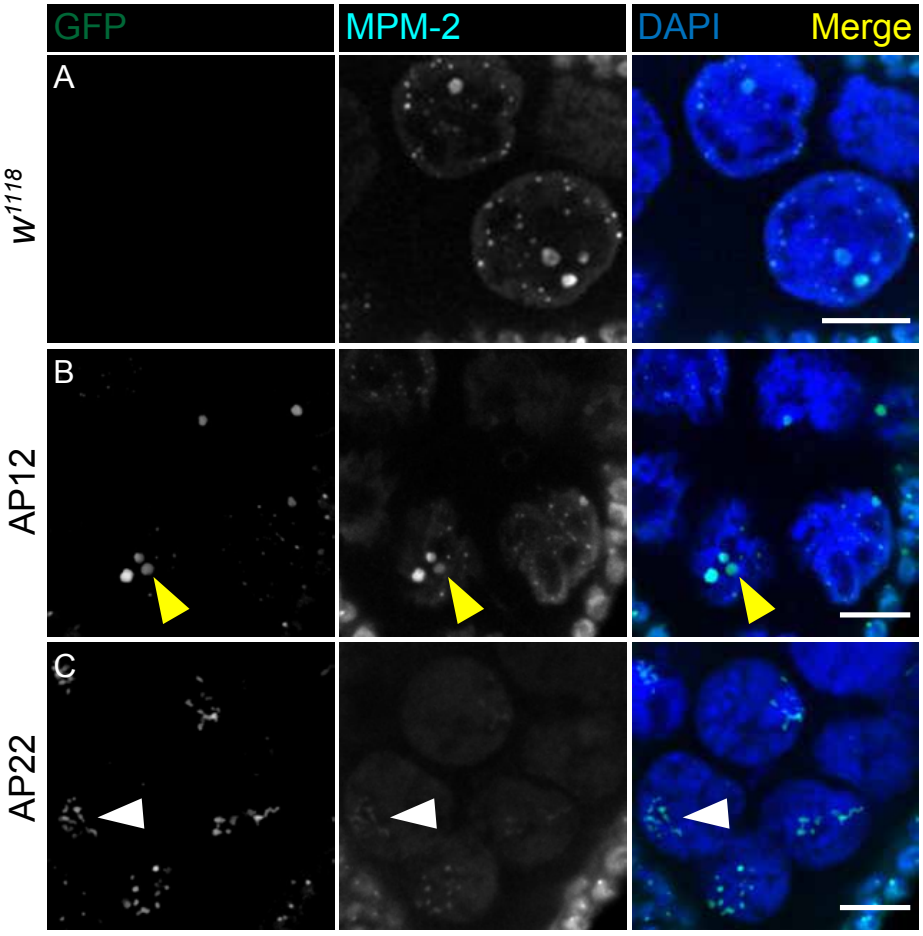


Figure 3.4. Both GFP-Mxc^{AP12} and GFP-Mxc^{Ap22} display MPM-2 foci. Stage-8 egg chambers from *w*^{1118}, *mxc*$^{G48/G48}$; *mxc*$^{AP12/12}$ and *mxc*$^{G48/G48}$; *mxc*$^{A22/22}$ ovaries were stained with DAPI, and anti-GFP and MPM-2 antibodies. (A) Wild- type (*w*^{1118}) egg chamber showing MPM-2 recognizing endogenous Mxc at the histone locus. (B) AP12 (GFP-Mxc^{AP12}) egg chamber showing GFP/MPM-2 co-localizing foci at the histone locus (yellow arrowhead). (C) AP22 (GFP-Mxc^{AP22}) egg chambers showing a bigger number of smaller GFP/MPM-2 co-localizing foci (white arrowhead). Bars: 10μM.

For instance, GFP-Mxc^{AP22} does not display multiple foci in stage-8 egg chamber follicle cells, whose polyploidy is given by a smaller DNA content than that of nurse cells. However, there is a possibility that somatic follicle cells can experience a milder version of genomic instability that cannot be perceived by the microscopy technology that we have available. Whether this is a true phenomenon associated with Mxc's phosphorylation state, which affects proper chromatin condensation, requires further investigation. Interestingly, it has been shown that a mutation of the *Drosophila* gene Nup154, which encodes a nucleoporin protein, induced the female-sterile phenotype manifested as oogenesis arrested at vitellogenic stages accompanied by clear defects in nurse cell chromatin organization (Gigliotti et al., 1998; Gigliotti et al., 2000).

To continue exploring where Mxc's CycE/Cdk2 phosphoepitopes lie, we generated a more complete set of Mxc phosphorylation mutants. The rational behind this strategy was to develop molecular tools that would allow us to conduct a more systematic analysis of the localization of CycE/Cdk2 phosphoepitopes of Mxc. We generated two Mxc phosphorylation mutants that contained either all SP/TP sites changed to alanines or all but serine 34 (S34) found in the LisH domain and called them AP36 and AP35, respectively (Figure 3.1). When GFP-Mxc^{AP36} and GFP-Mxc^{AP35} were transfected into S2

cells and stained for GFP and Mxc, both constructs showed a diffused nuclear localization compared to that of GFP-Mxc (Figure 3.5 A-C).

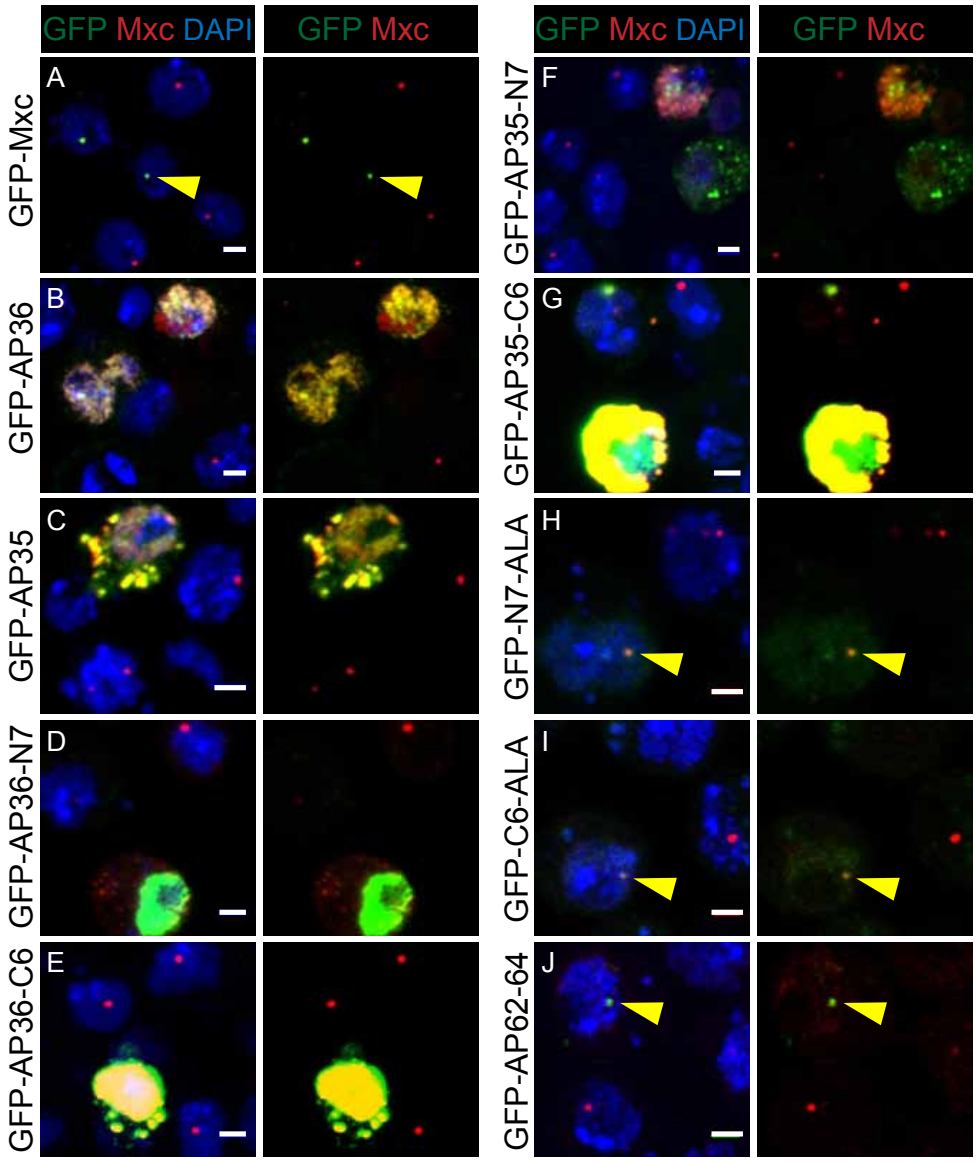


Figure 3.5. CycE/CDk2 phosphoepitopes affect Mxc's accumulation at the histone locus. Drosophila S2 cells were transfected with GFP-Mxc and corresponding GFP-Mxc phosphorylation mutants and subsequently stained with anti-GFP and anti-Mxc antibodies, and DAPI. (A) GFP-Mxc's concentration at

the histone locus is evidenced by GFP/Mxc co-localizing foci (yellow arrowheads). (B-G) GFP-Mxc phosphorylation mutants fail to concentrate at the histone locus. (H-J) GFP-Mxc phosphorylation mutants are capable of concentrating at the histone locus, as evidenced by GFP/Mxc co-localizing foci (yellow arrowheads). The faint Mxc signal in transfected cells, where co-localizing foci are observed, is likely a result of anti-Mxc antibody being outcompeted by the anti-GFP antibody. Bars: 2μM.

We reasoned that the observed diffused nuclear signal from both Mxc phosphorylation mutants could have been the result of either the lack of phosphoepitopes necessary for Mxc to effectively concentrate at the histone locus or that we mutagenized amino acids that are not necessarily target of phosphorylation, but rather that play a crucial role in the proper folding of Mxc.

Based on our previous observation that an Mxc phosphorylation mutant as GFP-Mxc^{AP22} is capable of concentrating at the histone locus in S2 cells and in adult ovaries (Figure 3.1-4) and that constructs such as GFP-Mxc^{AP36} and GFP-Mxc^{AP35} are not (Figures 3.1 and 3.5), we decided to more carefully study combinations of the last thirteen putative sites. We previously showed that shorter fragments of Mxc are capable of concentrating at the histone locus in the presence of endogenous Mxc (residues 1-1172; 1-1481; and 1-1642) and in the absent of it (residues 1-1481 and 1-1642) (Terzo et al., 2015). To test for localization and concentration at the histone locus, we generated four more constructs by swapping sets of seven or six, of the last thirteen putative phosphorylation sites on Mxc, with their respective wild-type serines or threonines using GFP-Mxc^{AP36} and GFP-Mxc^{AP35} constructs as templates. We named these four constructs GFP-Mxc$^{AP36\text{-}N7}$, GFP-Mxc$^{AP36\text{-}C6}$, GFP-Mxc$^{AP35\text{-}N7}$, and GFP-Mxc$^{AP35\text{-}C6}$ and reasoned that this new strategy could attenuate the potential effects of mutating important

phosphorylation sites on Mxc or of disrupting Mxc's proper folding. However, when we transfected S2 cells with these four constructs, we observed again a diffused nuclear GFP signal, instead of the typical focal localization at the histone locus displayed by GFP-Mxc (Figure 3.5A and D-G). In light of these results and of the challenges faced while exploring Mxc's putative CycE/Cdk2 phosphorylation sites, we reasoned that subtle changes in the number of changed putative sites could have a milder effect on Mxc's ability to concentrate at the histone locus and/or proper folding. As discussed above, changing clusters of 12 or 22 consecutive putative phosphorylation sites did not affect Mxc's ability to concentrate at the histone locus, or at least not at levels that can be perceived by the light microscopy technologies employed in the present study. With this in mind, we specifically mutagenized small subsets (seven or six) of the last thirteen putative phosphorylation sites while maintaining the rest of the thirty-six sites as wild type. These Mxc phosphorylation mutants carried 7 sites, GFP-Mxc$^{N7\text{-}ALA}$, or 6 sites GFP-Mxc$^{C6\text{-}ALA}$ changed to alanines. In addition, it has been recently proposed that Mxc's CycECdk2-dependent phosphorylation could be a required signal to release FLASH from Mxc. This step would allow for the FLASH/Lsm11 interaction to occur, which is required to ensure high efficiency of the U7-dependent cleavage and to therefore prevent histone mRNA misprocessing (Yang et al., 2014). To explore this possibility, we also generated a third phosphorylation mutant containing three consecutive putative CycE/Cdk2 sites changed to alanines that we named GFP-Mxc$^{AP62\text{-}64}$. GFP-Mxc$^{AP62\text{-}64}$ contains T1762, S1763, and T1764 changed to alanines. These three amino acids of Mxc have been found to be a conserved region in the protein sequence of its human homolog, NPAT. When we transfected S2 cells with these three mutants, we observed that all of

them were capable of concentrating at the histone locus (Figure 3.5H-J). These results demonstrate that GFP-Mxc$^{N7\text{-}ALA}$, GFP-Mxc$^{C6\text{-}ALA}$, and GFP-Mxc$^{AP62\text{-}64}$ carry intact phosphorylation sites, or a number of them, sufficient for these Mxc phosphorylation mutants to accumulate at the histone locus. As the entirety of these cellular assays were conducted maintaining endogenous Mxc present, we speculate that any abnormality in the ability of GFP-Mxc$^{N7\text{-}ALA}$, GFP-Mxc$^{C6\text{-}ALA}$, and GFP-Mxc$^{AP62\text{-}64}$ to accumulate at the histone locus and promote and maintain histone gene expression will only be appreciated in the absence of endogenous Mxc. To accomplish this, we are currently cloning GFP-Mxc$^{N7\text{-}ALA}$, GFP-Mxc$^{C6\text{-}ALA}$, and GFP-Mxc$^{AP62\text{-}64}$ into the pUbi-GFP vector, described above, to generate transgenic flies. We will test if these new Mxc phosphorylation mutants can accumulate at the histone locus, form a stable HLB, be detected by MPM-2 antibody, and promote and maintain histone gene expression at normal levels.

Role of Mxc's phosphorylation in HLB assembly and histone mRNA biosynthesis

To explore the role that Mxc's CycE/Cdk2 phosphorylation plays in HLB assembly, we have conducted immunofluorescence assays and asked if known HLB factors concentrate at the histone locus. So far, we have observed that many Mxc phosphorylation mutants have not been able to accumulate at the histone locus, but that others have and that the latter group can concentrate HLB factors at the histone locus (Figures 3.1 and 3.5). These findings made us wonder how CycE/Cdk2-dependent phosphorylation affects Mxc's ability to form a stable HLB that can promote and maintain histone gene expression during S phase. To explore the role that Mxc's phosphorylation plays in histone mRNA biosynthesis, we turned to biochemical assays to

assess histone transcript levels. Based on the data described above, we reasoned that the differences in ovary phenotype (arrested oogenesis at stage-8 egg chambers and subtle chromatin decondensation) between $mxc^{G48/G48}$; $mxc^{AP12/12}$ and $mxc^{G48/G48}$; $mxc^{AP22/22}$ flies could be due to a deleterious HLB assembly that could translate into detectable levels of defective histone mRNA biosynthesis.

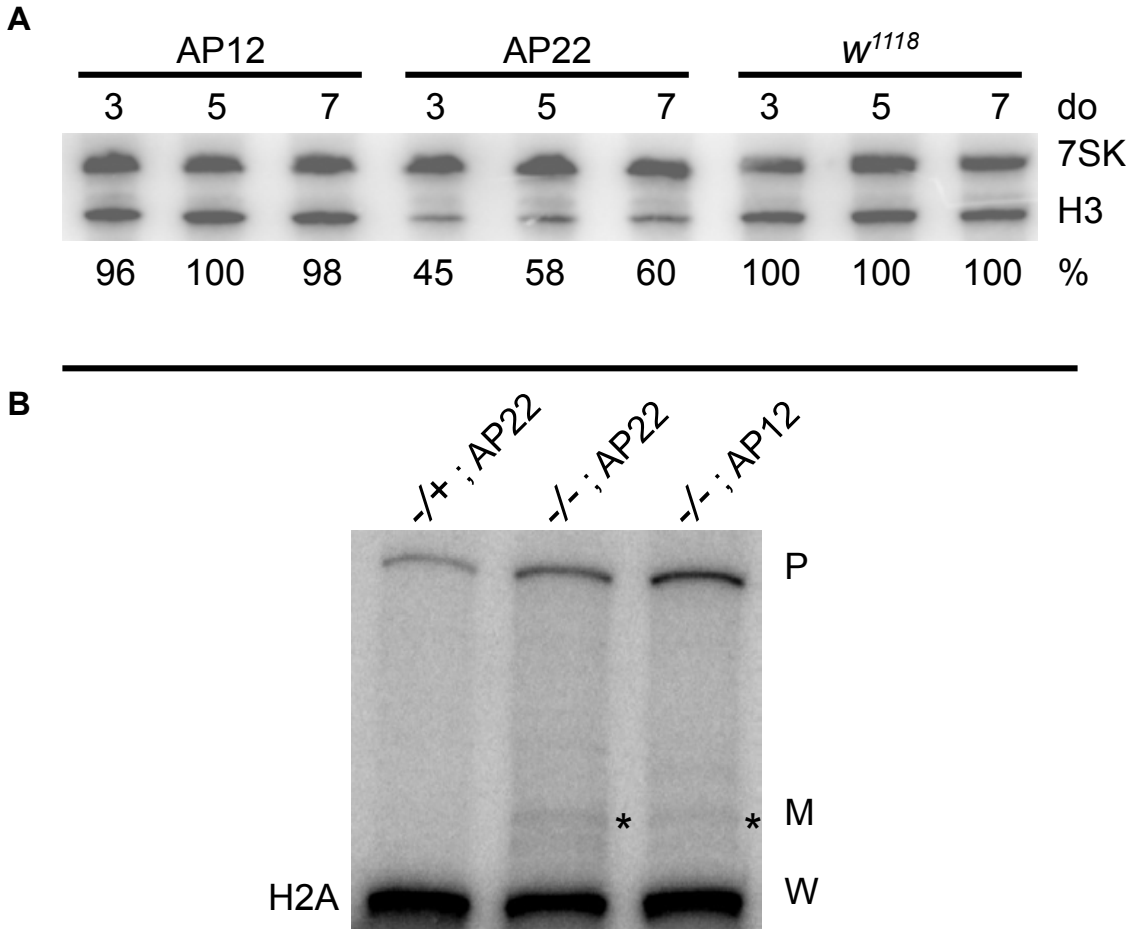


Figure 3.6. Both GFP-Mxc^{AP12} and GFP-Mxc^{Ap22} display defective histone mRNA biosyhnthesis. Cellular RNA was extracted with Trizol Reagent (Invitrogen) from ovaries of 3-7-day old female flies and used for Northern blots and S1 Nuclease protection assay. (A) Northern blot analysis of histone H3

transcript levels from ovaries of $mxc^{G48/G48}$; $mxc^{AP12/12}$, w^{1118}, and $mxc^{G48/G48}$; $mxc^{A22/22}$ female flies. 1 µg of total RNA per well was run on a 6% acrylamide 8M Urea denaturing gel. 7SK RNA was used as a loading control on both gels. Numbers below each lane represent the averaged percentage of histone H3 transcript levels obtained from two independent experiments. do = day old. (B) 5 µg of total RNA were used for each S1 nuclease protection reaction. P = probe; M = misprocessed histone H2A mRNA; and W = normally processed histone H2A mRNA. (*) = misprocessed histone H2A mRNA species.

To further interrogate this view, we conducted Northern Blots and S1 Nuclease protection assay using probes for histone H3 and H2A, respectively. These two biochemical assays were performed using RNA from ovaries of female flies that were 3-7-day old. Northern Blot analysis revealed a dramatic decrease in histone H3 transcript levels in ovaries from $mxc^{G48/G48}$; $mxc^{AP22/22}$ female flies when compared to their wild-type (w^{1118}) and $mxc^{G48/G48}$; $mxc^{AP12/12}$ counterparts (Figure 3.6A). When we looked for potential processing defects using a probe for histone H2A, we did not detect decreased levels of normally processed histone mRNA. However, we were able to detect faint bands of misprocessed histone H2A in ovaries from both $mxc^{G48/G48}$; $mxc^{AP22/22}$ and $mxc^{G48/G48}$; $mxc^{AP12/12}$ flies that were not seen in ovaries from wild-type siblings (GFP-Mxc^{AP22} segregating in endogenous Mxc background) (Figure 3.6B). Our preliminary data suggest that CycE/Cdk2-dependent phosphorylation plays an important role in the ability of Mxc to accumulate at the histone locus and to subsequently form a stable, mature HLB.

Discussion and future work

To date, several post-translational modifications (PTMs) are known to be involved in tuning the interactions driving NB assembly and disassembly (Zhu and

Brangwynne, 2015). For instance, SUMOylation of PML is thought to play an important role in PML body assembly (Shen et al., 2006). Phosphorylation is another important modification for regulating the biding of low complexity domains to cytoplasmic or nuclear bodies. For instance, *in vivo* work demonstrated that by changing normally phosphorylatable serine residues in the low complexity domain of the splicing factor (SRSF2) to glycine, CLK1 was not able to phosphorylate the mutated domains, giving rise to a hypophosphorylated form of the protein that could not be released from nucleoli (Kwon et al., 2014). This strongly suggests that phosphorylation plays an important role in the regulation of engagement and disengagement of factors with the nucleoli. In addition, a recent work demonstrated that the assembly and disassembly of membrane-less, cytoplasmic P granules are regulated by the phosphorylation of low complexity domains found in two of its components, MEG-1 and MEG-3. MEG (maternal-effect germline defective) proteins are germ plasm components that are required redundantly for fertility in *C. elegans*. Phosphorylation of MEGs promotes granule disassembly and dephosphorylation induces granule assembly (Wang et al., 2014).

Our laboratory has recently shown that an approximately 300-nucleotide sequence containing the H3-H4 bidirectional promoter is sufficient for mediating the concentration of HLB factors and for the expression of all of the canonical histone genes in *Drosophila* (Salzler et al., 2013). The biochemical data presented in this dissertation strongly suggest that Mxc's CycE/Cdk2-dependent phosphorylation plays a role in its ability to properly accumulate at the histone locus and thus form a stable HLB capable of controlling the expression of all histone genes (as histone H3 and H2A genes do not share the same bidirectional promoter). In a null *mxc* background, both GFP-Mxc^{AP12} and GFP-Mxc^{AP22}

are capable of accumulating at the histone locus and concentrating HLB factors, but the subtle differences identified by immunofluorescence techniques between these two *mxc* phosphorylation mutants (Figures 3.2-3.4) become more apparent when histone mRNA biosynthesis is explored employing biochemical assays (Figure 3.6). It is tempting to think that GFP-Mxc^{AP22}, when segregating in a null *mxc* background, fails to effectively recruit certain HLB factor(s) at the histone locus, due to the lack of one, or more, phosphorylation site that is present on the GFP-Mxc^{AP12} phosphorylation mutant. Although only GFP-Mxc^{AP22} shows a dramatic decrease of histone H3 transcript levels in a null background (Figure 3.6A), both GFP-Mxc^{AP22} and GFP-Mxc^{AP22} show subtle processing defects as observed for histone H2A (Figure 3.6B). All in all, these results along with the observed subtle phenotypes, suggest that Mxc's phosphorylation by CycE/Cdk2 is important for the formation of a stable HLB. Our data also suggest that Mxc does not harbor all of the CycE/Cdk phosphoepitopes in the region encompassing the putative sites 2-23 (see Figure 3.1). It is likely that Mxc carries more than one CycE/Cdk2 epitope and that, if so, these could be in regions of the protein that have yet to be mutagenized and tested and that they could also be involved in different aspects of Mxc's function at the HLB. In support of this notion, human NPAT has been shown to possess five CycE/Cdk2 phosphorylation sites and when all five are mutagenized to alanines NPAT 's ability to activate an H2B reporter construct in human cells was dramatically reduced (Ma et al., 2000). Ma and colleagues' results not only suggest that CycE/Cdk2 phosphorylation of NPAT substantially contributes to NPAT-mediated histone H2B transcriptional activation, but also that this aspect of NPAT's function can be mediated by one or more CycE/Cdk2 phosphoepitopes. We hypothesize that Mxc

harbors multiple CycE/Cdk2 phosphoepitopes, which confer it with the ability to accumulate at the histone locus and to assemble a stable, mature HLB. By changing the number of phosphorylated sites on Mxc, the cell cycle controls Mxc's ability to remain at the histone locus at an optimal concentration, to maintain a fully assembled HLB, and to promote active histone transcription during S phase. We propose that dephosphorylation of Mxc begins as the S phase approaches its end, which gives rise to Mxc's gradual disengagement from the HLB (reducing its concentration, at the histone locus, below optimal levels for HLB assembly). This phenomenon could conclude with the inactivation of histone gene expression triggered by HLB destabilization and subsequent removal of some histone mRNA biosynthetic factors from the histone locus at the end of S phase, a point from which unphosphorylated Mxc could remain chromosome associated at the histone locus until the onset of the next S phase. Perhaps, an alternative pathway that helps initiate HLB disassembly could be triggered by the Mxc's hyperphosphorylation (multiple phosphorylated sites on Mxc while approaching the end of S phase), which could signal for phosphatases to dephosphorylate Mxc and consequently begin the disengagement of it so that only unphosphorylated Mxc remains chromosome bound at the histone locus.

Future work will benefit from concentrating and purifying full-length Mxc, a task that has proven daunting in our hands, to subject the protein to Mass Spectrometry (MS) to perform a more precise mapping of phosphorylation sites. In recent years, MS has revolutionized the analysis of signaling by allowing rapid identification of phosphorylation sites with high precision and sensitivity (Dephoure et al., 2013). In terms of defining more precisely Mxc's ability to concentrate HLB factors at the histone locus,

one should consider implementing new high-resolution microscopy techniques that our laboratory will have soon access to at UNC. This type of microscope is not only capable of acquiring images, from fixed tissues and in real time, with great resolution and depth, but it can also perform a quantitative assessment of the biological material being analyzed. We also propose to perform, along with the biochemical assays mentioned above (Northern Blot and S1 Nuclease protection), Fluorescence In Situ Hybridization (FISH) using RNA probes to cytologically explore how Mxc phosphorylation mutants, forming an unstable, and thus not as affective, HLB, affect histone gene expression throughout *Drosophila* development.

Materials and Methods

Drosophila strains

mxc allele (mxc^{G48}) was previously described (Santamaria and Randsholt, 1995; Terzo et al., 2015; White et al., 2011). *mxc* transgenic flies ($gfp\text{-}mxc^{AP12}$ and $gfp\text{-}mxc^{AP22}$) were generated by injecting y^1w^{1118};PBac[y^+-attP-3B]VK00033 embryos (BestGene, Inc.) with a ϕC31-compatible vector containing an N-terminal GFP tag and expression of the fusion protein was driven by the *ubiquitin* promoter in pUGW.

Immunofluorescence

For ovaries, the following primary antibodies were used: monoclonal mouse MPM-2, (1:2000, Millipore); chicken anti-GFP (1:1000, Millipore); affinity-purified polyclonal rabbit anti-FLASH (1:2000), and affinity purified rabbit and guinea pig anti-Mxc (1:2000) (White et al., 2011; Yang et al., 2009). For S2 cells, immunostaining was

performed as described (Terzo et al., 2015; White et al., 2011). The secondary antibodies used (1:2000) in all experiments were: goat anti-rabbit IgG labeled with Alexa fluor 488 (Abcam) or Cy5 (Jackson); goat anti-mouse IgG Cy3 (Jackson); donkey anti-chicken Cy2 (Jackson); and goat anti-guinea pig IgG Cy3 or Cy5 (Jackson). DNA was detected by incubating tissue in 1ug/ml DAPI (DAKO Corporation, Carpinteria, CA) for 1 min.

Imaging

Confocal images for ovary immunostaining were obtained at a zoom of 1.0-5.0 with a 63x Plan Achromat (NA 1.4) objective using the ZEN data acquisition software on a laser-scanning confocal microscope (710; Carl Zeiss). Confocal images for *Drosophila* S2 cells were taken at a zoom of 4.0-8.0 with a 63x Plan Achromat (NA 1.4) objective on a laser-scanning confocal microscope (SP5; Leica, Exton, PA).

Molecular biology

Mxc protein fragments used for immunostaining experiments were expressed in *Drosophila* cultured S2 cells or as transgenes in Gateway-compatible vectors (Carnegie Institution) as previously described (Terzo et al., 2015; White et al., 2011).

An 882-nucleotide long DNA fragment of the *mxc* gene containing three putative CycE/Cdk2 phosphorylation sites mutagenized (T1762/A; S1763/A; and T1764/A) was synthesized by GenScript (Piscataway, NJ, USA). The synthesized DNA fragment was cloned into the pUC-57 vector, cut with SpeI and AscI restriction enzymes and cloned into pENTR Gateway vector.

Northern Blotting

Northern blotting was performed using a 6% 7M urea acrylamide gel to resolve histone mRNAs and 7SK RNA (Nguyen et al., 2012) as previously described (Mullen and Marzluff, 2008).

S1 Nuclease Protection Assay

Probe was created by 5' end labeling BspEII cut H2A DNA with a-^{32}P-dCTP using the Klenow fragment of DNA pol I (New England Biolabs). The DNA probe was gel purified and hybridized to total ovary RNA or control yeast tRNA followed by digestion with S1 nuclease (Lanzotti et al., 2002). Protected fragments were resolved on a 6% polyacrylamide-7M urea gel and visualized by autoradiography.

REFERENCES

Calvi, B.R., M.A. Lilly, and A.C. Spradling. 1998. Cell cycle control of chorion gene amplification. *Genes Dev*. 12:734-744.

Courchaine, E., and K.M. Neugebauer. 2015. Paraspeckles: Paragons of functional aggregation. *J Cell Biol*. 210:527-528.

Davis, F.M., T.Y. Tsao, S.K. Fowler, and P.N. Rao. 1983. Monoclonal antibodies to mitotic cells. *Proc Natl Acad Sci U S A*. 80:2926-2930.

Dephoure, N., K.L. Gould, S.P. Gygi, and D.R. Kellogg. 2013. Mapping and analysis of phosphorylation sites: a quick guide for cell biologists. *Mol Biol Cell*. 24:535-542.

Gigliotti, S., G. Callaini, S. Andone, M.G. Riparbelli, R. Pernas-Alonso, G. Hoffmann, F. Graziani, and C. Malva. 1998. Nup154, a new Drosophila gene essential for male and female gametogenesis is related to the nup155 vertebrate nucleoporin gene. *J Cell Biol*. 142:1195-1207.

Gigliotti, S., D. Rotoli, A. Manzi, F. Graziani, and C. Malva. 2000. Female sterile mutations and egg chamber development in Drosophila melanogaster. *Int J Dev Biol*. 44:581-589.

Han, T.W., M. Kato, S. Xie, L.C. Wu, H. Mirzaei, J. Pei, M. Chen, Y. Xie, J. Allen, G. Xiao, and S.L. McKnight. 2012. Cell-free formation of RNA granules: bound RNAs identify features and components of cellular assemblies. *Cell*. 149:768-779.

Hennig, S., G. Kong, T. Mannen, A. Sadowska, S. Kobelke, A. Blythe, G.J. Knott, K.S. Iyer, D. Ho, E.A. Newcombe, K. Hosoki, N. Goshima, T. Kawaguchi, D. Hatters, L. Trinkle-Mulcahy, T. Hirose, C.S. Bond, and A.H. Fox. 2015. Prion-like domains in RNA binding proteins are essential for building subnuclear paraspeckles. *J Cell Biol*. 210:529-539.

Kwon, I., M. Kato, S. Xiang, L. Wu, P. Theodoropoulos, H. Mirzaei, T. Han, S. Xie, J.L. Corden, and S.L. McKnight. 2013. Phosphorylation-regulated binding of RNA polymerase II to fibrous polymers of low-complexity domains. *Cell*. 155:1049-1060.

Kwon, I., S. Xiang, M. Kato, L. Wu, P. Theodoropoulos, T. Wang, J. Kim, J. Yun, Y. Xie, and S.L. McKnight. 2014. Poly-dipeptides encoded by the C9orf72 repeats bind nucleoli, impede RNA biogenesis, and kill cells. *Science*. 345:1139-1145.

Lanzotti, D.J., H. Kaygun, X. Yang, R.J. Duronio, and W.F. Marzluff. 2002. Developmental control of histone mRNA and dSLBP synthesis during Drosophila embryogenesis and the role of dSLBP in histone mRNA 3' end processing in vivo. *Mol Cell Biol*. 22:2267-2282.

Ma, T., B.A. Van Tine, Y. Wei, M.D. Garrett, D. Nelson, P.D. Adams, J. Wang, J. Qin, L.T. Chow, and J.W. Harper. 2000. Cell cycle-regulated phosphorylation of p220(NPAT) by cyclin E/Cdk2 in Cajal bodies promotes histone gene transcription. *Genes Dev*. 14:2298-2313.

Mullen, T.E., and W.F. Marzluff. 2008. Degradation of histone mRNA requires oligouridylation followed by decapping and simultaneous degradation of the mRNA both 5' to 3' and 3' to 5'. *Genes Dev*. 22:50-65.

Nguyen, D., B.J. Krueger, S.C. Sedore, J.E. Brogie, J.T. Rogers, T.K. Rajendra, A. Saunders, A.G. Matera, J.T. Lis, P. Uguen, and D.H. Price. 2012. The Drosophila 7SK snRNP and the essential role of dHEXIM in development. *Nucleic acids research*. 40:5283-5297.

Perkins, J.R., I. Diboun, B.H. Dessailly, J.G. Lees, and C. Orengo. 2010. Transient protein-protein interactions: structural, functional, and network properties. *Structure*. 18:1233-1243.

Salzler, H.R., D.C. Tatomer, P.Y. Malek, S.L. McDaniel, A.N. Orlando, W.F. Marzluff, and R.J. Duronio. 2013. A sequence in the Drosophila H3-H4 Promoter triggers histone locus body assembly and biosynthesis of replication-coupled histone mRNAs. *Dev Cell*. 24:623-634.

Santamaria, P., and N.B. Randsholt. 1995. Characterization of a region of the X chromosome of Drosophila including multi sex combs (mxc), a Polycomb group gene which also functions as a tumour suppressor. *Mol Gen Genet*. 246:282-290.

Seet, B.T., I. Dikic, M.M. Zhou, and T. Pawson. 2006. Reading protein modifications with interaction domains. *Nat Rev Mol Cell Biol*. 7:473-483.

Shen, T.H., H.K. Lin, P.P. Scaglioni, T.M. Yung, and P.P. Pandolfi. 2006. The mechanisms of PML-nuclear body formation. *Mol Cell*. 24:331-339.

Terzo, E.A., S.M. Lyons, J.S. Poulton, B.R. Temple, W.F. Marzluff, and R.J. Duronio. 2015. Distinct self-interaction domains promote Multi Sex Combs accumulation in and formation of the Drosophila histone locus body. *Mol Biol Cell*.

Ubersax, J.A., and J.E. Ferrell, Jr. 2007. Mechanisms of specificity in protein phosphorylation. *Nat Rev Mol Cell Biol*. 8:530-541.

Uversky, V.N. 2013. Intrinsic disorder-based protein interactions and their modulators. *Curr Pharm Des*. 19:4191-4213.

van der Lee, R., M. Buljan, B. Lang, R.J. Weatheritt, G.W. Daughdrill, A.K. Dunker, M. Fuxreiter, J. Gough, J. Gsponer, D.T. Jones, P.M. Kim, R.W. Kriwacki, C.J. Oldfield, R.V. Pappu, P. Tompa, V.N. Uversky, P.E. Wright, and M.M. Babu. 2014. Classification of intrinsically disordered regions and proteins. *Chem Rev*. 114:6589-6631.

Vucetic, S., H. Xie, L.M. Iakoucheva, C.J. Oldfield, A.K. Dunker, Z. Obradovic, and V.N. Uversky. 2007. Functional anthology of intrinsic disorder. 2. Cellular components, domains, technical terms, developmental processes, and coding sequence diversities correlated with long disordered regions. *J Proteome Res*. 6:1899-1916.

Wang, J.T., J. Smith, B.C. Chen, H. Schmidt, D. Rasoloson, A. Paix, B.G. Lambrus, D. Calidas, E. Betzig, and G. Seydoux. 2014. Regulation of RNA granule dynamics by phosphorylation of serine-rich, intrinsically disordered proteins in C. elegans. *Elife*. 3:e04591.

White, A.E., B.D. Burch, X.C. Yang, P.Y. Gasdaska, Z. Dominski, W.F. Marzluff, and R.J. Duronio. 2011. Drosophila histone locus bodies form by hierarchical recruitment of components. *J Cell Biol*. 193:677-694.

White, A.E., M.E. Leslie, B.R. Calvi, W.F. Marzluff, and R.J. Duronio. 2007. Developmental and cell cycle regulation of the Drosophila histone locus body. *Mol Biol Cell*. 18:2491-2502.

Yang, X.C., B.D. Burch, Y. Yan, W.F. Marzluff, and Z. Dominski. 2009. FLASH, a proapoptotic protein involved in activation of caspase-8, is essential for 3' end processing of histone pre-mRNAs. *Mol Cell*. 36:267-278.

Yang, X.C., I. Sabath, L. Kunduru, A.J. van Wijnen, W.F. Marzluff, and Z. Dominski. 2014. A conserved interaction that is essential for the biogenesis of histone locus bodies. *J Biol Chem*. 289:33767-33782.

Zhu, L., and C.P. Brangwynne. 2015. Nuclear bodies: the emerging biophysics of nucleoplasmic phases. *Curr Opin Cell Biol*. 34:23-30.

CHAPTER IV

DISCUSSION AND FUTURE DIRECTIONS

Gaining insight into the molecular mechanisms orchestrating gene expression will advance our general understanding of cellular homeostasis under normal and pathological conditions. The genome controls gene expression by utilizing a myriad of regulatory components some of which are discrete membrane-less, but at the same time defined nuclear suborganelles known as NBs (Dundr, 2012; Dundr and Misteli, 2010; Misteli, 2007). NBs compartmentalize the nucleus and create distinct environments where they can carry out their specific functions (Dundr and Misteli, 2010; Matera et al., 2009; Morimoto and Boerkoel, 2013; Sleeman and Trinkle-Mulcahy, 2014). The work contained in this dissertation was conducted to explore the connection between HLB assembly and histone mRNA biosynthesis. In particular, this research was focused on the molecular mechanisms underlying the dynamic nature of HLB assembly throughout the cell cycle. We utilized the *Drosophila* HLB as a model for NB formation and function. Our laboratory previously demonstrated that Mxc is crucial for HLB assembly and histone mRNA biosynthesis (White et al., 2011). Here, we demonstrated that Mxc plays an integral role in concentrating factors at the histone locus that are required for histone mRNA biosynthesis. We showed that Mxc utilizes multiple domains to accumulate at the histone locus and to concentrate histone mRNA biosynthetic factors required to assemble

a fully functional HLB. We found that two N-terminal domains of Mxc, the LisH domain and the novel SIF domain, are required for Mxc's self-interaction, which is essential for its concentration at the histone locus. We also demonstrated that Mxc harbors a region between amino acids 721-1481 that is necessary for HLB assembly independent of the LisH and SIF domains, adding another structural tier to the 3-dimenssional lattice that supports the HLB. The last 195 amino acids of Mxc were shown to be required for the recruitment of FLASH, a fundamental component of the histone pre-mRNA processing machinery, to the histone locus.

Mxc was shown to directly react with MPM-2, which strongly suggests that it harbors CycE/Cdk2 phosphoepitopes and therefore that the cell cycle is involved in regulating HLB assembly and histone mRNA biosynthesis (White et al., 2011). The present dissertation also contains the work that is currently being conducted to identify these putative CycE/Cdk2 sites on Mxc and to explore their contribution to HLB assembly and histone gene expression. Based on the data herein presented, we hypothesize that Mxc harbors multiple CycE/Cdk2 sites and that these are required for Mxc's accumulation at the histone locus and for the assembly of a stable, mature HLB.

HLB assembly is dynamic and it involves multiple protein-protein interactions

We previously showed that Mxc and FLASH are first associated with the histone locus during cell cycle 10 of Drosophila embryogenesis, one cycle before histone gene transcription begins (White et al., 2011). Our laboratory has also shown that Mxc is crucial for HLB assembly and histone gene expression (White et al., 2011). Here, we demonstrate that Mxc employs distinct domains to increase its local concentration at the

histone locus in order to assemble a fully formed and thus functional HLB. We propose that Mxc is a multivalent protein with the capacity of locally concentrating at the histone

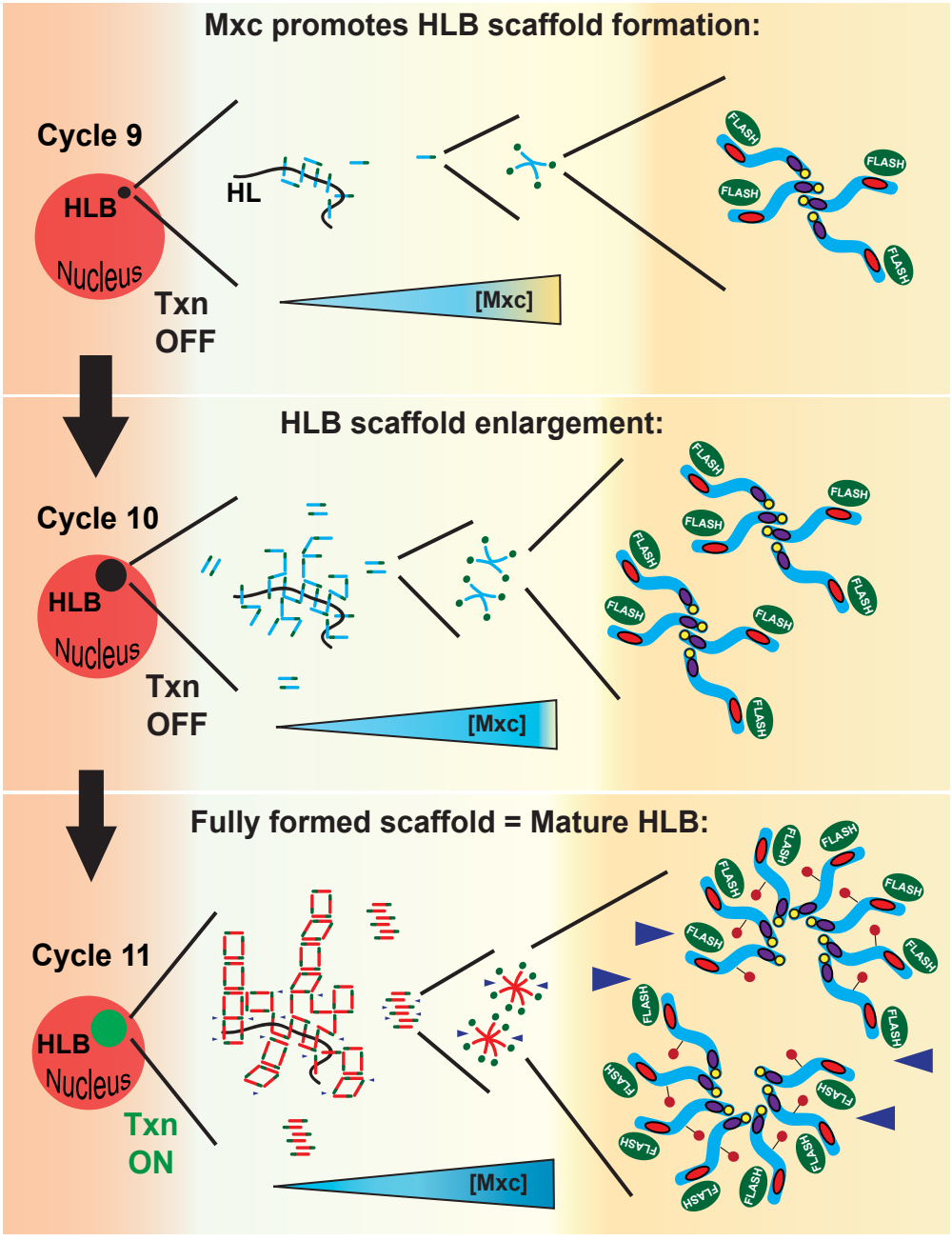


Figure 4.1. HLB assembly is dynamic and requires an Mxc scaffold and multiple protein-protein interactions. Mxc (light blue lines) and FLASH (green circles) can self-organize and increase their local

concentration (represented by blue inside the horizontal triangle) at the histone locus (HL) to begin the formation of a molecular scaffold during cell cycle 9 of *Drosophila* embryogenesis. During cell cycle 10, an even larger increment in Mxc's concentration at the histone locus gives rise to a significant enlargement of the HLB scaffold. At the onset of S phase of cycle 11, CycE/Cdk2-dependent phosphorylation of Mxc (represented by red lines and red lollipops in the magnification) takes place and, along with the initiation of histone gene expression (green HLB), the recruitment of other HLB factors that are signal dependent (purple triangles) begins. At cell cycle 11, multiple protein-protein interactions facilitate the formation of a fully formed molecular scaffold containing optimal concentrations to assemble a mature HLB capable of maintaining histone gene expression throughout S phase. In the magnification on the far right, Mxc displays three domains: LisH domain (yellow circle), SIF domain (purple oval), and 721-1481 region (red oval). For more details on Mxc's domains, read chapter II and see Figure 2.10.

locus throughout the early cell cycles of *Drosophila* embryogenesis, via self-interaction and the physical binding to other HLB factors, which renders Mxc, once it reaches a determined focal stoichiometry, capable of assembling a 3-dimensional lattice, the HLB (Figure 4.1). During cell cycle 11, the 3-dimensional scaffold, dependent on the optimal local concentration of all the required HLB factors, along with Mxc's CycE/Cdk2-dependent phosphorylation, and the initiation of histone mRNA transcription give rise to the dynamic formation of a mature HLB that requires multiple protein-protein interactions.

Our results raise important questions regarding Mxc's ability to dynamically localize and concentrate at the histone locus throughout the cell cycle. For instance, an HLB formed by truncated Mxc^{G46} cannot be considered functionally effective as it fails to recruit FLASH at the histone locus, giving rise to slightly lower levels of normally processed histone mRNA and notably a local accumulation of unprocessed histone

mRNA species (Rajendra et al., 2011; Terzo et al., 2015). However, this 3-dimensional lattice, with truncated Mxc^{G46} as the building block, remains functionally active as hypomorphic mxc^{G46} embryos reach adulthood and therefore are capable of supporting histone mRNA biosynthesis (Landais et al., 2014; Rajendra et al., 2011; Terzo et al., 2015). Does truncated Mxc^{G46} fail to retain FLASH inside the 3-dimensional lattice (HLB) in close proximity to and at the appropriate concentration at the histone locus? Although we (Terzo et al., 2015) and others (Landais et al., 2014; Rajendra et al., 2011) have been able to show that FLASH's local concentration is decreased by performing immunofluorescence assays, this type of analysis lacks the analytical aspect needed to assess the amounts of FLASH (Mxc and all known HLB factors) required for effectively maintaining histone mRNA biosynthesis throughout S phase. This fundamental aspect of Mxc's function in HLB assembly requires further investigation that can benefit with the advent of powerful high-resolution microscopy technologies capable of making such assessment not only with a higher quality, but also, and more importantly, quantitatively. In conjunction with new and powerful genome-editing technologies such as CRISPR (Clustered Regularly Interspaced Short Palindromic Repeats) (Jinek et al., 2012) now widely used in numerous model organisms, including *Drosophila*, we can insert fluorescent tags in the genes encoding HLB factors. By doing so, I propose to quantitate the optimal concentration (number of molecules) of each endogenous HLB factor, at the histone locus, either individually or in various combinations to also assess their interactions and dynamic behaviors. High-resolution microscopy technologies will allow us to more closely explore the complex structure that the HLB represents throughout the entire cell cycle, and at different stages of development, and therefore to widen our

understanding of how this intricate and dynamic membrane-less NB functions and how its individual components interact with each other.

The dynamism that characterizes the behavior of NBs (Deryusheva and Gall, 2004; Dundr et al., 2004; Dundr and Misteli, 2010; Handwerger et al., 2005; Handwerger and Gall, 2006; Matera et al., 2009), and some cytoplasmic bodies (Brangwynne et al., 2009), and the fluidity with which components of these molecular compartments can traverse across them have been well documented (Brangwynne, 2013; Dundr et al., 2004; Dundr and Misteli, 2003; Hyman et al., 2014). Several lines of research have shown that factors focally accumulated at these subcellular compartments have a high tendency to interact with themselves and with each other and are crucial to give rise to these higher-order assemblages (Conduit et al., 2014; Dundr and Misteli, 2010; Hebert and Matera, 2000; Kawaguchi et al., 2015; Matera et al., 2009; Miele et al., 2005; Toretsky and Wright, 2014; Voronina et al., 2011). Together, these findings strongly suggest that multivalent proteins can act as scaffolds that mediate the local concentration of factors and that oligomerization is a common feature to the assembly of these compartments (Good et al., 2011; Matera et al., 2009; Zaidi et al., 2007). We propose that the *Drosophila* Mxc is capable of forming a higher-order assemblage by an oligogmerization dependent on its ability to self-interact. This large Mxc oligomer involves numerous Mxc proteins recruited from the nucleoplasm to the histone locus to give rise to the HLB. Furthermore, we speculate that the increase in the local concentration of Mxc could act as an attracting force to recruit other HLB factors in order to reach critical concentrations required to form a mature HLB ready to initiate and maintain histone mRNA biosynthesis. Supporting the role of Mxc (and its mammalian ortholog, NPAT) in HLB

assembly, by recruiting histone mRNA biosynthetic factors to the histone locus, critical molecular interactions involved in HLB formation have recently been shown. For instance, the C-terminus of NPAT (human ortholog of Mxc) was shown to interact with the C-termini of both FLASH and YARP (mammalian ortholog of *Drosophila* HLB component Mute) and that these interactions are essential for concentrating FLASH and YARP at the histone locus in HeLa cells (Yang et al., 2014). We recently demonstrated that Mxc harbors two N-terminal domains (LisH and SIF) involved in its self-interaction, which is crucial for Mxc's ability to concentrate at the histone locus and to assemble the HLB (Terzo et al., 2015). In addition, while the C-termini of FLASH and Mxc play a crucial role in concentrating FLASH and U7 snRNP at the histone locus, FLASH's N-terminus has been shown to be required for histone pre-mRNA processing and also for interacting with U7 snRNP through the N-terminus of Lsm11. The latter FLASH/Lsm11 interaction has been shown to be fundamental for histone pre-mRNA processing, but not for HLB assembly (Burch et al., 2011). Combined, these findings demonstrate that HLB assembly involves several protein-protein interactions between distinct HLB factors, and themselves, and strongly suggest that the phenomenon of assembly brings HLB components close together so that they can physically bind via weak interactions in a reversible and highly dynamic fashion.

HLB assembly conforms to the "seed and grow" model

How are NBs first assembled? What signal is required for these complex structures to come about under diverse physiological conditions or types of stress? To date, an emergent principle in the field of nuclear organization states that certain

subnuclear domains are associated with gene loci and that the nature of such association is highly dynamic and responsive to distinct cellular signals (Matera et al., 2009). Many of the known NBs are not constitutively associated with a specific locus on the genome. However, other NBs are known to associate with particular loci. For instance, the nucleolus is known to intimately associate with the genes encoding the 35S preribosomal RNA. As the morphology of the nucleolus has been shown to directly correlate with the relative transcriptional activity of the rRNA genes (Haaf et al., 1991; Scheer et al., 1984), it is now evident that nucleoli assemble as a consequence of rRNA transcription and processing, and also the assembly of ribosomal subunits (Matera et al., 2009). Another example of a chromosome-associated NB is the HLB, which forms at the histone locus where histone mRNA biosynthetic factors are concentrated (Ghule et al., 2008; Liu et al., 2006; White et al., 2007). CBs and PML bodies are examples of bodies known to associate transiently with particular loci, but they are not thought to remain constitutively associated at these sites (Matera et al., 2009). Whether NBs constitutively or transiently associate with specific loci on the genome, or are able to assemble and remain chromosome dissociated freely floating in the nucleoplasm, all NBs require an initial signal that triggers their assembly.

In recent years, several research groups have attempted to shed light on the mechanisms underlying NB assembly and more specifically on how NB formation is initiated. Two current notions of NB formation are considered. In one scenario, nuclear compartments can follow an ordered assembly pathway in which all of their components are assembled in a hierarchical, or predefined, sequence. On the other hand, NBs can follow a stochastic self-organization where the assembly of their components occurs in a

randomized fashion (Dundr and Misteli, 2010; Machyna et al., 2013; Matera et al., 2009). Our laboratory demonstrated that the *Drosophila* HLB is assembled in a hierarchical manner and that Mxc and FLASH, first in localizing at the histone locus, together serve as a seeding platform for the subsequent steps in HLB assembly (White et al., 2011). Our laboratory has also demonstrated that Mxc and FLASH seed the HLB formation by self-organization properties that allow them to form a proto-HLB, even in the absence of histone genes, and that a transcription-dependent ordered recruitment of other factors is required for the formation of a mature HLB (Salzler et al., 2013; White et al., 2011).

Supporting evidence for the stochastic assembly model comes from a cell culture study in which individual CB components were tethered to an engineered random site integrated into the genome (Kaiser et al., 2008). With this approach, the authors demonstrated that any CB component is capable of initiating the formation of the entire NB, suggesting that, in line with self-organization, the order of components in NB assembly is not important for CB formation. In addition, two subsequent studies shed light on the role of RNA in NB formation, indicating that coding and non-coding RNAs can act as structural elements in their formation. These two studies utilized different technical approaches to provide evidence for the *de novo* nucleation of HLB with the associated CB, nuclear speckle, paraspeckle, and nuclear stress bodies at sites where different RNAs were tethered (Shevtsov and Dundr, 2011) and that paraspeckles can be visualized undergoing *de novo* formation by inducing transcription of NEAT1 (Mao et al., 2011). These experiments in which the initiation of NB assembly can be triggered by any component clearly supports the notion that RNAs or proteins can support NB formation. However, the Lac repressor tethering system is creating an artificial platform

or scaffold upon which any component specific to a certain NB can give rise to such nuclear compartment. This system cannot faithfully address what component is truly seeding, or the first one in triggering the initiation of, NB assembly.

Recently, a new hybrid model (seed and grow) that reconciles the stochastic and the hierarchical evidence above mentioned proposes that NB assembly can be triggered by an initial seeding event which is non-random and driven by a biological process, such as transcription. Subsequently, specific proteins may bind to the seed to begin the formation of a scaffold primarily driven by random and self-organized mechanism (Dundr, 2011). Combining our previous work and the data contained in this dissertation, we favor the "seed and grow" model as the mechanism underpinning HLB assembly. Although we have shown that Mxc and FLASH can stochastically form a proto-HLB by self-organization properties and that histone mRNA transcription does not initiate HLB assembly, but is rather required for its maturation, we propose that Mxc and FLASH are nucleated at the histone locus by a non-random, and still unknown, signal and together initiate the seeding that precedes the recruitment of all required histone mRNA biosynthetic factors. Furthermore, the multiple protein-protein interactions between distinct HLB components along with histone mRNA transcription, and post-transcriptional modifications (PTMs) such as Mxc's phosphorylation, are the driving forces that locally nucleate these factors at critical concentrations required for stabilizing the HLB and effectively maintaining histone gene expression throughout the entire S phase. After the seeding takes place, we also propose that some protein-protein interactions follow an ordered pattern of assembly (e.g., Spt6's local concentration triggered by specific signals during S phase) while others occur in a randomized fashion.

HLB behavior and histone mRNA biosynthesis during *Drosophila* embryogenesis

To explore *in vivo* the behavior of Mxc and how dynamic HLB assembly is in real time, we turned to the *Drosophila* embryo, as it is known to be an excellent system to perform live imaging experiments. In *Drosophila*, once fertilization has taken place, the nuclei inside the egg undergo 13 rounds of highly synchronized multiplications and divisions in one shared cytoplasm called a syncytium. The first cell cycles are very rapid, lack gap phases, and cells divide approximately every 8.5 minutes, but later on they begin to slow down as interphases progressively increase their length from about 9 minutes in cycle 10 to 14 minutes in cycle 13 (Foe and Alberts, 1983). During cell cycle 14, not only S phase is much longer than that of any other previous cell cycle, but it also exhibits the first G2 gap phase. At fertilization, the genome of the embryo is nearly quiescent and it depends on maternally supplied gene products to control development during the first hours of embryogenesis. As the embryo develops, zygotic gene products become required for fundamental biological events such as cell cycle progression (Farrell and O'Farrell, 2014).

Zygotic histone gene transcription is known to be initiated during nuclear cycle 11, while, in general, most other genes begin to be transcribed in nuclear cycle 14 (Edgar and Schubiger, 1986). By staining fixed embryos, our laboratory previously showed that both Mxc and FLASH accumulate at the histone locus during cell cycle 10, one cell cycle before zygotic histone gene expression begins (White et al., 2011). While reviewing the latter results, we were able to observe that Mxc/FLASH foci consistently become brighter in all of the nuclei as the embryo undergoes subsequent cell cycles (see figure 10 of

White et al., 2011 for more details). We reasoned that perhaps Mxc and FLASH are increasing their presence at the histone locus during the cell cycles previous to cycle 11, in preparation for the initiation of histone mRNA transcription. Interestingly, it has also been shown that all genes acquire competence for transcriptional activation during nuclear cycle 10 of *Drosophila* embryogenesis (Edgar and Schubiger, 1986). In addition, research recently conducted by several groups suggests that transcription can and does happen at, and even before, cycle 10 in *Drosophila* (Farrell and O'Farrell, 2014). In light of these findings, we wondered how dynamic the HLB assembly is during these early cell cycles and why HLB formation (Mxc/FLASH foci) begins a cell cycle before histone gene expression is initiated? Therefore, we sought to gain a better understanding of how HLB assembly behaves during early *Drosophila* development as it unfolds. To do this, we employed live-imaging microscopy, a highly sensitive technology that does not disrupt biological tissues, as fixation of tissues does. To capture in real time how the HLB assembles, we followed GFP-tagged full-length Mxc during early cell cycles. We demonstrated that Mxc is first visualized accumulating at the histone locus as early as cell cycle 9 of *Drosophila* embryogenesis. Based on these findings, we favored the idea that Mxc increasingly accumulated at the histone locus, during the cycles previous to the firing of the histone genes, in order to reach a critical concentration that allows it to more effectively concentrate all HLB factors, also at their required concentrations, to promote histone mRNA biosynthesis during cycle 11. In support of this notion, we were able to demonstrate that GFP-Mxc foci undergo a significant size enlargement as they progress from nuclear cycle 9 to 11. In addition, we demonstrated that this HLB expansion requires Mxc's self-interaction as the driving force mediated by both the LisH and SIF

domains. Combining our previous work with the results herein presented, we demonstrate that Mxc is crucial for the assembly of a mature HLB and that it plays an integral role in concentrating all known histone mRNA biosynthetic factors at the histone locus, which is essential to support normal development (Terzo et al., 2015; White et al., 2011). Interestingly, it has been recently shown in *Drosophila* that 24 transgenic copies of the 5-kilobase histone gene cluster, in a histone deletion background, generate the same amount of protein and mRNA as the 200 copies of the endogenous histone genes (McKay et al., 2015). In the same study, the authors demonstrate that 24 copies of histone transgene can also assemble an HLB and normally process histone transcripts. However, the authors observed that an HLB formed at the transgene locus is not fully assembled, as evidenced by the lowered signal (diffused) of Mute (*muscle wasted*), a known HLB factor, when compared to that of HLBs formed in wild-type animals. It is therefore tempting to speculate that although a small subset of histone gene copies is sufficient for normal, active transcription to happen during development, a larger number of these genes may be necessary to act as a scaffold that can serve as a platform to properly concentrate all factors required to assemble a mature HLB, which is crucial for maintaining efficient histone mRNA transcription throughout the entire development of the animal. Lastly, the need of such high histone gene copy number may play a crucial role in chromatin architecture of the actively transcribing histone genes by maintaining their compaction in a relaxed state, which may vary depending upon a particular stage of development, type of tissue, or a combination of both.

Can phase transition mediate HLB Assembly?

In our proposed HLB assembly model (discussed in more detail in chapter II), we postulate that Mxc together with FLASH "seed" the formation of the HLB. Furthermore, we propose that Mxc employs distinct domains to form a molecular lattice that acts as a platform to concentrate all the required histone mRNA biosynthetic factors at the histone locus. We favor the notion that Mxc acts as a scaffolding protein orchestrating the HLB assembly by self-interacting and through interactions with other HLB factors, and also by permitting interactions among these factors to ultimately facilitate the downstream reactions involved in histone mRNA biosynthesis. Moreover, we propose that Mxc can do so not only by employing structured domains, such as the LisH and SIF, but also low complexity regions of disordered structure and uncharacterized functions. Despite important advances recently made in the field of nuclear organization, the biophysical rules dictating the assembly of NBs remains poorly understood. To further explore the possibility that Mxc acts as a scaffold that orchestrates HLB assembly to promote histone gene expression, we conducted SMART (Simple Modular Architecture Research Tool) software analysis on the amino acid sequence of Mxc. Aside from the structured domains LisH and AT-hook (previously characterized on other proteins) and the novel SIF domain, we found ten regions of on Mxc lacking any known function annotation, which were defined as low complexity domains. In general, proteins are made up of a single or multiple domains that can have distinct molecular functions. These domains are referred to as structured domains and they often fold independently, give rise to precise tertiary contacts, and adopt specific 3-dimensional structures that carry their function. However, a

large fraction of the proteome of any organism consists of polypeptide segments that are not likely to form a defined 3-dimensional structure, but are nevertheless functional (van der Lee et al., 2014). The latter protein segments are referred to as intrinsically disordered regions (IDRs) and it is now well established that these domains actively participate in diverse functions mediated by proteins. For instance, IDRs are frequently subjected to post-translational modifications (PTMs) that increase the functional states in which a protein can exist in the cell (Collins et al., 2008). IDRs are also known to display short motifs of 3-10 residues that permit interaction with structured domains in other proteins (Davey et al., 2012). These two features in isolation or combined allow for the interaction and recruitment of diverse proteins in space and time thus facilitating the regulation of virtually all cellular processes (van der Lee et al., 2014). In addition, 28 separate functions were proposed to be distinguished for IDRs, based on the literature analysis of 150 proteins containing disordered regions of 30 residues or longer (Dunker et al., 2002). These IDR functionalities have been summarized and classified into three broad functional categories: (1) facilitated regulation via diverse post-translational modifications, (2) scaffolding and recruitment of different binding partners, and (3) conformational viability and adaptability (Tompa, 2005; van der Lee et al., 2014). It is important to mention that a protein can consist of several disordered regions that belong to different functional categories (Uversky, 2013a; Uversky, 2013c). Based on these functional categories of IDRs, we think that Mxc could be classified within the second group, as we propose that it acts as a scaffold utilizing structured (folded) domains to self-interact and to concentrate HLB factors at the histone locus. Furthermore, we propose that Mxc's low complexity domains are crucial for HLB assembly, perhaps as

display domains for PTMs, such as phosphorylation, which our data strongly suggest plays an important role in Mxc's accumulation at the histone locus and in conferring stability to the HLB.

In recent years, several research groups have shown that proteins containing IDRs (or low complexity domains) can undergo large-scale association through homotypic or heterotypic multivalent interactions (van der Lee et al., 2014). These proteins harboring low complexity domains have been shown to undergo phase transition, which leads to separated liquid droplets, hydrogels, and protein aggregates or fibrils (Toretsky and Wright, 2014). Many cellular processes are conducted inside organelles that are enclosed within lipid membranes. However, other functions depend upon assemblies of proteins and nucleic acids that are membrane-less. This category of biological macromolecules, which includes assemblages such as NBs, can give rise to distinct compartments in the nucleoplasm or cytoplasm via phase separation. Consistent with the hypothesis that postulates that NBs assemble through intracellular phase separation, NBs often behave like liquid droplets of RNA and proteins (Toretsky and Wright, 2014). For instance, it is well known that somatic cell nucleoli are capable of fusing with one another. By using the large and numerous extrachromosomal nucleoli in the nucleus of large *Xenopus laevis* oocytes, the dynamics of coalescence were shown to be quantitatively consistent with coalescence dynamics of simple liquids, such as two oil droplets fusing in water (Brangwynne et al., 2011). Interestingly, the same study also revealed ATP-dependent dynamics, suggesting a possible form of 'active diffusion' within the nucleolus itself. When authors depleted ATP, nucleoli exhibited approximately 10-fold increase in apparent viscosity almost to the point of "freezing". A study conducted in *Drosophila*

oocytes has shown that *de novo* NBs can be induced to nucleate and grow from the nucleoplasm by mechanical perturbations of the egg chamber, which likely changes the nucleoplasmic salt concentration and therefore the strength of molecular interactions (Singer and Gall, 2011). These NBs are highly spherical and can be observed to undergo striking liquid-like coalescence events. Recently, further quantitative support for the role of phase separation in the formation of nuclear structures was provided by a study demonstrating that Ddx4, an essential component for the assembly and maintenance of the related nuage in ammals, P-granules in worms, and pole plasm and polar granules in flies (Liang et al., 1994), assembles *in vitro* into liquid phase droplets, which depends on protein concentration, salt concentration, and temperature in a manner that can be quantitatively modeled as a phase transition (Nott et al., 2015). The authors also demonstrate that the disordered N-terminus of human Ddx4 can spontaneously self-associate both in HeLa cells and *in vitro*. Thus, the latter study provides strong evidence that phase separation can drive the formation of synthetic bodies that are qualitatively similar to those naturally observed in the nucleus.

Our work demonstrates that Mxc contains structured (LisH and SIF) and low complexity (disordered) domains. In light of our findings and taking into consideration the mounting evidence supporting phase transition as the assembly force for several NBs, it is tempting to speculate that this biophysical phenomenon can also dictate HLB assembly. In recent years, several proteins containing multiple protein-protein, protein-RNA interaction domains, intrinsically disordered regions (IDRs), or a combination of folded and disordered domains have been shown to play a crucial role in organizing molecular structures in both the nucleus and the cytoplasm (Conduit et al., 2014; Elbaum-

Garfinkle et al., 2015; Hebert and Matera, 2000; Hennig et al., 2015; Kwon et al., 2014; Li et al., 2012; Nott et al., 2015; Rajgor et al., 2014; Weber and Brangwynne, 2012; Yang et al., 2014). Recently, it has been shown that the *Drosophila* Centrosomin (Cnn) protein, involved in the regulation of centrosome size asymmetry in neuroblasts, by differential control of its incorporation into the pericentriolar material (PCM), is phosphorylated at the centrosomes where it acts as a scaffold around the centrioles (Conduit et al., 2014; Conduit and Raff, 2010). The same authors demonstrated that Cnn's Polo/Plk1-dependent phosphorylation initiates the assembly of a Cnn scaffold that is essential for efficient centrosome maturation in *Drosophila*. In this study, the authors unambiguously demonstrate that the Cnn scaffold is in constant flux by performing photoconversion experiments. They utilized fluorescence recovery after photobleaching (FRAP) to examine the spatiotemporal dynamics of GFP-Cnn incorporation into mitotic centrosomes and showed that Cnn molecules incorporate only into the center of the PCM and then move slowly outward. By performing 3-dimensional-structured illumination super-resolution microscopy (3D-SIM), they examined the architecture of the GFP-Cnn molecules inside the PCM in real time and demonstrated that Cnn forms an extended structure that appeared to emanate from the centrioles. The authors also showed that Cnn is phosphorylated by Polo-Plk1 in a domain they called PReM (phosphoregulated-multimerization), which allows it to assemble into a scaffold around the centrioles. Moreover, they performed size exclusion chromatography multi-angle light-scattering (SEC-MALS) analyses to show that purified MBP-Cnn existed predominantly as dimers, which they went on to show that is independent of phosphorylation and rather dependent on a leucine zipper (LZ) domain (Conduit et al., 2014). Lastly, their centrosome assembly

model suggests that arrangements of hydrophobic and hydrophilic residues within the LZ domain could allow for multimeric structures. The authors speculate, based on their yeast two hybrid (Y2H) data, that multiple regions of Cnn can self-interact and could therefore participate in multimerization (Conduit et al., 2014). In this dissertation, we have presented data demonstrating that Mxc plays an essential role in the formation of a 3-dimensional structure that gives rise to the HLB. We have shown that Mxc is capable of self-interacting, which is one of the driving forces of HLB assembly. Our live-imaging experiments unambiguously show that the HLB (GFP-Mxc foci) undergoes a significant enlargement in size throughout the initial cell cycles of *Drosophila* embryogenesis. Lastly, we have presented ample evidence strongly suggesting that Mxc's CycE/Cdk2-dependent phosphorylation plays a fundamental role in Mxc's ability to assemble a stable HLB. The work conducted by Conduit and colleagues (2014) has been enlightening to us due to the remarkable similarities observed between Centrosomin (Cnn) and Mxc. These two *Drosophila* proteins, although in different cellular compartments (cytoplasm and nucleus, respectively) are capable of assembling molecular scaffolds or lattices that are integral for centrosome and HLB to remain stable and functional. Both, Cnn and Mxc are phosphorylated and have been proposed to have the capacity to form higher-order multimers, by utilizing multiple domains, that are essential for these scaffolds to assemble and expand outward or enlarge throughout the subsequent cell cycles, as is the case for centrosomes and HLBs, respectively. Based on the parallels above mentioned between Cnn and Mxc, we reason that we can perform many of the assays that Conduit et al (2014) conducted to further explore the different aspects of Mxc's role as a scaffold required for HLB formation. Can Mxc form higher-order multimers as it was proposed

for Cnn? We propose to follow up on our most recent work by performing size exclusion chromatography multi-angle light-scattering (SEC-MALS) analysis to more precisely interrogate the mechanistic details of Mxc's aggregation (multimerization). In addition, we propose to perform 3-dimensional-structured illumination super-resolution microscopy (3D-SIM) and fluorescence recovery after photobleaching (FRAP) to more closely and precisely examine Mxc's scaffold (lattice) structure in real time and the dynamics of Mxc's incorporation into the HLB. Combined, these three powerful technologies will help us shed light on the role that Mxc plays in HLB assembly and also on the spatiotemporal aspect of the molecular mechanisms and biophysical forces involved in forming such complex structure. Interestingly, a recent work developed a physical description of centrosome growth in which it is proposed that centrosome material occurs in a soluble form and a form that tends to give rise to droplets by phase separation. The authors went on to show that an autocatalytic chemical transition between these forms (formation of droplet phase in the cytosol) accounts for key features of centrosome growth in *C. elegans* (Zwicker et al., 2014). Could therefore phase transition also be the biophysical force that governs HLB assembly?

It is becoming increasingly clear that both cytoplasm and nucleoplasm can exist in a variety of distinct phases and that weak and reversible binding between proteins and proteins and RNA govern a wide range of interactions that give rise to liquid-like ribonucleoprotein (RNP) droplets (Brangwynne, 2013). Three characteristics define a liquid compartment: (1) compartments should fuse after touching, (2) compartments should be spherical, which is driven by surface tension, and (3) their components should undergo rapid internal rearrangement (Hyman et al., 2014). Therefore, to explore if there

is a constant mixing of HLB components, such as Mxc, within a liquid-like droplet, our work can benefit, in addition to the assays and technologies above mentioned, from the implementation of a technique known as half-bleaching. In this method, half a structure is bleached, and then the distribution of the fluorescence within the photomanipulated structure is determined over time thus giving a spatiotemporal assessment of the dynamic flux of a particular component(s) (Brangwynne et al., 2009; Hyman et al., 2014; Patel et al., 2015).

Taking into consideration the data herein presented and the accruing evidence supporting phase transition as a biophysical mechanism underlying RNP bodies, such as NBs, we hypothesize that Mxc employs distinct domains and several CycE/Cdk2 sites to coordinate the formation of a scaffold that is promoted by phase transition to locally concentrate optimal levels of histone mRNA biosynthetic factors at the histone locus during S phase (Figure 4.2). Furthermore, we propose that Mxc by establishing several protein-protein interactions stabilizes the HLB and its CycE/Cdk2 epitopes are also involved in the relay of signals to turn specific steps of histone gene expression on and off, whose efficiency of reaction is increased by maintaining HLB factors in close proximity. Lastly, we propose that histone mRNA transcription and CycE/Cdk2 epitopes on Mxc, shown to be essential for recruiting determined HLB factors at the histone locus (Salzler et al., 2013; White et al., 2011), are required for locally modulating thermodynamic parameters that help mature an HLB without altering the passive phase separation mechanism, as it has been shown for the nucleolus in *C. elegans* (Berry et al., 2015).

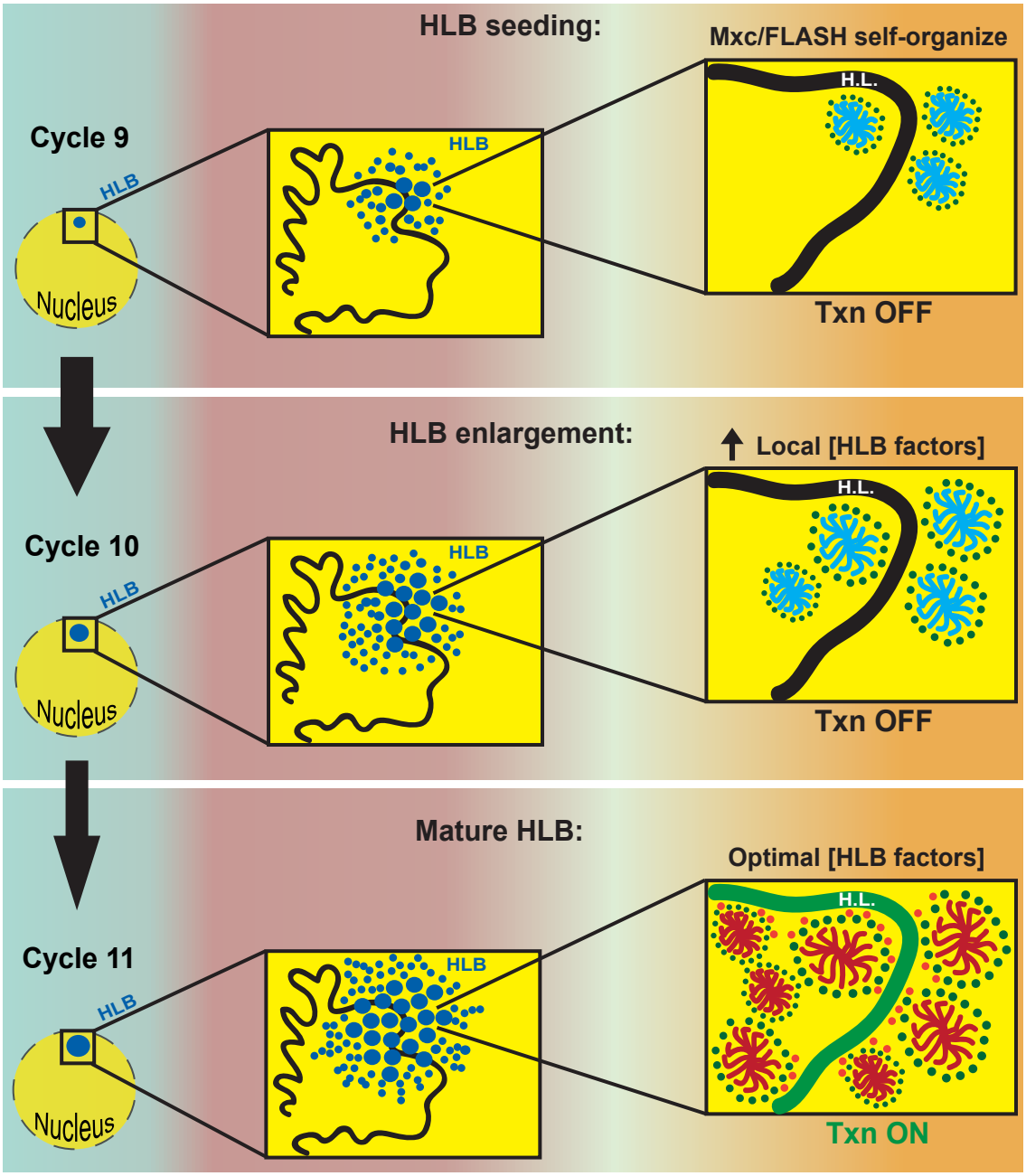


Figure 4.2. Phase separation-dependent HLB assembly model. Mxc and FLASH self-organize, which allows them to focally concentrate even in the absence of the histone locus. Under normal physiological conditions and in the presence of the histone locus (black H.L = transcription OFF), Mxc and FLASH increase their local concentration near the histone locus. This stochastic self-organization step allows Mxc

and FLASH to demix from the nucleoplasm by forming liquid-like droplets, here represented by light blue lines (Mxc) surrounded by green circles (FLASH), that are nucleated in close proximity to the histone locus by an unknown signal. This gives rise to the seeding of the HLB during cell cycle 9 of *Drosophila* embryogenesis. We do not exclude the possibility that HLB assembly can be facilitated by phase transition even earlier in development. During cell cycle 10 and by a physical property that defines liquids, two or more Mxc/FLASH liquid droplets found close to one another begin to fuse into bigger droplets. This continuous and increasing local concentration ([HLB factors]) is also facilitated by Mxc's ability to self-interact, which gives rise to a 3-dimensional structure/scaffold that will serve as a platform for histone gene expression. At the onset of cell cycle 11, histone mRNA transcription begins as a result of the Mxc scaffold that initiated the recruitment of required histone mRNA biosynthetic factors, which along with the firing of transcription and Mxc's CycE/Cdk2 phosphorylation (red lines), all together signal for the nucleation of optimal levels of all factors required for effectively assembling a mature HLB and maintaining histone gene expression throughout S phase. Orange circles in liquid droplets of cell cycle 11 represent HLB factors that are only recruited and concentrated at optimal levels when Mxc is phosphorylated by the CycE/CDk2 complex and transcription has begun (green H.L = transcription ON).

Phase separation under normal and pathological conditions

What are the consequences of phase separation as a result of concentrating proteins both in cytoplasm and nucleoplasm? What happens when these proteins engage or assemble into liquid droplets that fail to dissolve or disassemble? It is now well known that when certain proteins concentrate above physiological levels, they tend to trigger aggregation leading to solid gels or even crystals, which does not create the appropriate environment for chemical reactions to happen (Hyman et al., 2014). Cells have developed a mechanism to cope with protein aggregation known as disaggregases, which are chaperones that forcibly untangle protein aggregates (Doyle et al., 2013). In addition, cells can also regulate the dynamics of the aggregated compartments by, for instance, the

process of phosphorylation or dephosphorylation (Wippich et al., 2013). On the other hand, under conditions such as metabolic syndrome or in the presence of mutant proteins, which tend to aggregate faster than their normal counterparts, cells do not have the capacity to respond by dissolving these aggregates.

Illustrating this pernicious aggregation phenomenon, several diseases of the brain are characterized by the development of toxic aggregates, such as amyloid formations in Alzheimer's disease (Brundin et al., 2010), synuclein plaques in Parkinson's disease (Shulman et al., 2011), or plaques seen in amyotrophic lateral sclerosis (Robberecht and Philips, 2013). These proteins are normally capable of forming liquid-like phases, but when diseases develop from mutations of specific proteins, they give rise to aggregates with more solid-like properties, which cannot dissociate like liquids do. Supporting this notion, several groups have shown that mutations in FUS (fused in sarcoma) are associated with amyotrophic lateral sclerosis (ALS) and rare forms of frontotemporal lobar degeneration (FTLD) (Deng et al., 2014; Woulfe et al., 2010). In addition, some groups have reported that the prion-like low complexity domains of FUS can polymerize into fribrous amyloid-like assemblies in a cell-free system (Han et al., 2012; Kato et al., 2012; Kwon et al., 2013; Kwon et al., 2014). Until recently, the relationship between amyloid-like fibers (hydrogels) that form *in vitro* and the *in vivo* function of the proteins that give rise to these structures was difficult to understand. However, a recent study explored the dynamics of FUS in living cells and its relationship with the onset of disease. The authors demonstrate that both *in vivo*, and under physiological conditions *in vitro*, FUS is capable of forming liquid-like droplets that are dependent on the prion-like low complexity domain and can convert into solid aggregates, which can be exacerbated

by disease-associated mutations in the prion-like domain (Patel et al., 2015). Moreover, a recent work from Roy Parker's laboratory strongly suggest that liquid-like phase separation could contribute to RNA granule formation through synergistic, multivalent interactions of defined RNA-binding domains, IDRs, and RNA (Lin et al., 2015). The authors also demonstrate that based on salt resistance, morphology, and FRAP analysis these phase-separated liquid droplets mature over time to more stable and less dynamic assemblies that could potentially become amyloid-like fibers. Therefore, they propose a unified model in which the progression from dynamic liquid droplets to stable fibers could be regulated in cells by factors that enhance or inhibit fiber nucleation and growth, or that actively dissolve fibers, such asVCP/Cdc48 and Hsp70/Hsp40, known disaggregase complexes involved in controlling RNA granule lifetimes and turnover (Buchan et al., 2013; Walters et al., 2015). These structures can be formed with varying physical properties and chemical compositions depending upon distinct biological needs. When this fine-tuned regulation is deleterious, due to mutations that exacerbate fiber formation or impair disaggregase activity, these may lead to excess fiber formation and aberrant granules, which could explain amyloid-like fibers' contribution to neurodegenerative diseases.

Combined, these findings further underscore the relevance of understanding the biophysical aspects underpinning molecular assemblages, such as NBs. It is clear that the field of nuclear organization is already greatly benefitting from the emerging technologies that allow for more thorough evaluation of entire RNP assemblages. Imaging technologies are advancing, but probe and protein labeling remain a limitation (Huang et al., 2010). Improvement in designing probes, imaging technologies, and

intracellular biophysical measurements will provide us with a greater understanding of how these assemblages form and their roles/functions in diverse biological processes under normal and pathological conditions.

A fully assembled HLB is essential for histone mRNA biosynthesis efficiency

This dissertation contains the work exploring the mechanisms underlying histone locus body (HLB) assembly, the role of the cell cycle in its formation and stabilization, and how this nuclear assemblage contributes to histone gene expression. We employed the *Drosophila* HLB as a model to gain insight into NB formation and function. The data presented in this dissertation in conjunction with our previous work and the research conducted by others support the conclusion that Mxc plays a crucial role in the formation of a 3-dimenssional lattice that is fundamental for the assembly of a mature HLB, which is required for efficiently maintaining histone mRNA biosynthesis during S phase. In addition, our preliminary data indicate that a cell cycle input, via CycE/Cdk2 phosphorylation, governs the capacity of Mxc to effectively accumulate at the histone locus and to assemble a stable, mature HLB. We propose that Mxc harbors multiple CycE/Cdk2 phosphoepitopes and that these are part of a cell cycle-controlled mechanism to regulate HLB assembly perhaps by regulating the relay of distinct signals that are therefore required for histone gene expression during S phase. Our studies cannot exclude the possibility that Mxc's CycE/Cdk2 phosphoepitopes play a different role other than Mxc's focal concentration at the histone locus. Mxc plays an integral role in the formation of a molecular scaffold that serves as a platform for HLB assembly greatly

contributing to maintaining normal histone gene expression, which is essential for DNA replication and therefore for cellular homeostasis.

REFERENCES

Berry, J., S.C. Weber, N. Vaidya, M. Haataja, and C.P. Brangwynne. 2015. RNA transcription modulates phase transition-driven nuclear body assembly. *Proc Natl Acad Sci U S A*. 112:E5237-5245.

Brangwynne, C.P. 2013. Phase transitions and size scaling of membrane-less organelles. *J Cell Biol*. 203:875-881.

Brangwynne, C.P., C.R. Eckmann, D.S. Courson, A. Rybarska, C. Hoege, J. Gharakhani, F. Julicher, and A.A. Hyman. 2009. Germline P granules are liquid droplets that localize by controlled dissolution/condensation. *Science*. 324:1729-1732.

Brangwynne, C.P., T.J. Mitchison, and A.A. Hyman. 2011. Active liquid-like behavior of nucleoli determines their size and shape in Xenopus laevis oocytes. *Proc Natl Acad Sci U S A*. 108:4334-4339.

Brundin, P., R. Melki, and R. Kopito. 2010. Prion-like transmission of protein aggregates in neurodegenerative diseases. *Nat Rev Mol Cell Biol*. 11:301-307.

Buchan, J.R., R.M. Kolaitis, J.P. Taylor, and R. Parker. 2013. Eukaryotic stress granules are cleared by autophagy and Cdc48/VCP function. *Cell*. 153:1461-1474.

Burch, B.D., A.C. Godfrey, P.Y. Gasdaska, H.R. Salzler, R.J. Duronio, W.F. Marzluff, and Z. Dominski. 2011. Interaction between FLASH and Lsm11 is essential for histone pre-mRNA processing in vivo in Drosophila. *RNA*. 17:1132-1147.

Collins, M.O., L. Yu, I. Campuzano, S.G. Grant, and J.S. Choudhary. 2008. Phosphoproteomic analysis of the mouse brain cytosol reveals a predominance of protein phosphorylation in regions of intrinsic sequence disorder. *Mol Cell Proteomics*. 7:1331-1348.

Conduit, P.T., Z. Feng, J.H. Richens, J. Baumbach, A. Wainman, S.D. Bakshi, J. Dobbelaere, S. Johnson, S.M. Lea, and J.W. Raff. 2014. The centrosome-specific phosphorylation of Cnn by Polo/Plk1 drives Cnn scaffold assembly and centrosome maturation. *Dev Cell*. 28:659-669.

Conduit, P.T., and J.W. Raff. 2010. Cnn dynamics drive centrosome size asymmetry to ensure daughter centriole retention in Drosophila neuroblasts. *Curr Biol*. 20:2187-2192.

Davey, N.E., K. Van Roey, R.J. Weatheritt, G. Toedt, B. Uyar, B. Altenberg, A. Budd, F. Diella, H. Dinkel, and T.J. Gibson. 2012. Attributes of short linear motifs. *Mol Biosyst*. 8:268-281.

Deng, H., K. Gao, and J. Jankovic. 2014. The role of FUS gene variants in neurodegenerative diseases. *Nat Rev Neurol*. 10:337-348.

Deryusheva, S., and J.G. Gall. 2004. Dynamics of coilin in Cajal bodies of the Xenopus germinal vesicle. *Proc Natl Acad Sci U S A*. 101:4810-4814.

Doyle, S.M., O. Genest, and S. Wickner. 2013. Protein rescue from aggregates by powerful molecular chaperone machines. *Nat Rev Mol Cell Biol*. 14:617-629.

Dundr, M. 2011. Seed and grow: a two-step model for nuclear body biogenesis. *J Cell Biol*. 193:605-606.

Dundr, M. 2012. Nuclear bodies: multifunctional companions of the genome. *Curr Opin Cell Biol*. 24:415-422.

Dundr, M., M.D. Hebert, T.S. Karpova, D. Stanek, H. Xu, K.B. Shpargel, U.T. Meier, K.M. Neugebauer, A.G. Matera, and T. Misteli. 2004. In vivo kinetics of Cajal body components. *J Cell Biol*. 164:831-842.

Dundr, M., and T. Misteli. 2003. Measuring dynamics of nuclear proteins by photobleaching. *Curr Protoc Cell Biol*. Chapter 13:Unit 13 15.

Dundr, M., and T. Misteli. 2010. Biogenesis of nuclear bodies. *Cold Spring Harb Perspect Biol*. 2:a000711.

Dunker, A.K., C.J. Brown, J.D. Lawson, L.M. Iakoucheva, and Z. Obradovic. 2002. Intrinsic disorder and protein function. *Biochemistry*. 41:6573-6582.

Edgar, B.A., and G. Schubiger. 1986. Parameters controlling transcriptional activation during early Drosophila development. *Cell*. 44:871-877.

Elbaum-Garfinkle, S., Y. Kim, K. Szczepaniak, C.C. Chen, C.R. Eckmann, S. Myong, and C.P. Brangwynne. 2015. The disordered P granule protein LAF-1 drives phase

separation into droplets with tunable viscosity and dynamics. *Proc Natl Acad Sci U S A*. 112:7189-7194.

Farrell, J.A., and P.H. O'Farrell. 2014. From egg to gastrula: how the cell cycle is remodeled during the Drosophila mid-blastula transition. *Annu Rev Genet*. 48:269-294.

Foe, V.E., and B.M. Alberts. 1983. Studies of nuclear and cytoplasmic behaviour during the five mitotic cycles that precede gastrulation in Drosophila embryogenesis. *J Cell Sci*. 61:31-70.

Ghule, P.N., Z. Dominski, X.C. Yang, W.F. Marzluff, K.A. Becker, J.W. Harper, J.B. Lian, J.L. Stein, A.J. van Wijnen, and G.S. Stein. 2008. Staged assembly of histone gene expression machinery at subnuclear foci in the abbreviated cell cycle of human embryonic stem cells. *Proc Natl Acad Sci U S A*. 105:16964-16969.

Good, M.C., J.G. Zalatan, and W.A. Lim. 2011. Scaffold proteins: hubs for controlling the flow of cellular information. *Science*. 332:680-686.

Haaf, T., D.L. Hayman, and M. Schmid. 1991. Quantitative determination of rDNA transcription units in vertebrate cells. *Exp Cell Res*. 193:78-86.

Han, T.W., M. Kato, S. Xie, L.C. Wu, H. Mirzaei, J. Pei, M. Chen, Y. Xie, J. Allen, G. Xiao, and S.L. McKnight. 2012. Cell-free formation of RNA granules: bound RNAs identify features and components of cellular assemblies. *Cell*. 149:768-779.

Handwerger, K.E., J.A. Cordero, and J.G. Gall. 2005. Cajal bodies, nucleoli, and speckles in the Xenopus oocyte nucleus have a low-density, sponge-like structure. *Mol Biol Cell*. 16:202-211.

Handwerger, K.E., and J.G. Gall. 2006. Subnuclear organelles: new insights into form and function. *Trends Cell Biol*. 16:19-26.

Hebert, M.D., and A.G. Matera. 2000. Self-association of coilin reveals a common theme in nuclear body localization. *Mol Biol Cell*. 11:4159-4171.

Hennig, S., G. Kong, T. Mannen, A. Sadowska, S. Kobelke, A. Blythe, G.J. Knott, K.S. Iyer, D. Ho, E.A. Newcombe, K. Hosoki, N. Goshima, T. Kawaguchi, D. Hatters, L. Trinkle-Mulcahy, T. Hirose, C.S. Bond, and A.H. Fox. 2015. Prion-like domains in RNA

binding proteins are essential for building subnuclear paraspeckles. *J Cell Biol*. 210:529-539.

Huang, B., H. Babcock, and X. Zhuang. 2010. Breaking the diffraction barrier: super-resolution imaging of cells. *Cell*. 143:1047-1058.

Hyman, A.A., C.A. Weber, and F. Julicher. 2014. Liquid-liquid phase separation in biology. *Annu Rev Cell Dev Biol*. 30:39-58.

Jinek, M., K. Chylinski, I. Fonfara, M. Hauer, J.A. Doudna, and E. Charpentier. 2012. A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity. *Science*. 337:816-821.

Kaiser, T.E., R.V. Intine, and M. Dundr. 2008. De novo formation of a subnuclear body. *Science*. 322:1713-1717.

Kato, M., T.W. Han, S. Xie, K. Shi, X. Du, L.C. Wu, H. Mirzaei, E.J. Goldsmith, J. Longgood, J. Pei, N.V. Grishin, D.E. Frantz, J.W. Schneider, S. Chen, L. Li, M.R. Sawaya, D. Eisenberg, R. Tycko, and S.L. McKnight. 2012. Cell-free formation of RNA granules: low complexity sequence domains form dynamic fibers within hydrogels. *Cell*. 149:753-767.

Kawaguchi, T., A. Tanigawa, T. Naganuma, Y. Ohkawa, S. Souquere, G. Pierron, and T. Hirose. 2015. SWI/SNF chromatin-remodeling complexes function in noncoding RNA-dependent assembly of nuclear bodies. *Proc Natl Acad Sci U S A*. 112:4304-4309.

Kwon, I., M. Kato, S. Xiang, L. Wu, P. Theodoropoulos, H. Mirzaei, T. Han, S. Xie, J.L. Corden, and S.L. McKnight. 2013. Phosphorylation-regulated binding of RNA polymerase II to fibrous polymers of low-complexity domains. *Cell*. 155:1049-1060.

Kwon, I., S. Xiang, M. Kato, L. Wu, P. Theodoropoulos, T. Wang, J. Kim, J. Yun, Y. Xie, and S.L. McKnight. 2014. Poly-dipeptides encoded by the C9orf72 repeats bind nucleoli, impede RNA biogenesis, and kill cells. *Science*. 345:1139-1145.

Landais, S., C. D'Alterio, and D.L. Jones. 2014. Persistent replicative stress alters polycomb phenotypes and tissue homeostasis in Drosophila melanogaster. *Cell reports*. 7:859-870.

Li, P., S. Banjade, H.C. Cheng, S. Kim, B. Chen, L. Guo, M. Llaguno, J.V. Hollingsworth, D.S. King, S.F. Banani, P.S. Russo, Q.X. Jiang, B.T. Nixon, and M.K. Rosen. 2012. Phase transitions in the assembly of multivalent signalling proteins. *Nature*. 483:336-340.

Liang, L., W. Diehl-Jones, and P. Lasko. 1994. Localization of vasa protein to the Drosophila pole plasm is independent of its RNA-binding and helicase activities. *Development*. 120:1201-1211.

Lin, Y., D.S. Protter, M.K. Rosen, and R. Parker. 2015. Formation and Maturation of Phase-Separated Liquid Droplets by RNA-Binding Proteins. *Mol Cell*. 60:208-219.

Liu, J.L., C. Murphy, M. Buszczak, S. Clatterbuck, R. Goodman, and J.G. Gall. 2006. The Drosophila melanogaster Cajal body. *J Cell Biol*. 172:875-884.

Machyna, M., P. Heyn, and K.M. Neugebauer. 2013. Cajal bodies: where form meets function. *Wiley Interdiscip Rev RNA*. 4:17-34.

Mao, Y.S., H. Sunwoo, B. Zhang, and D.L. Spector. 2011. Direct visualization of the co-transcriptional assembly of a nuclear body by noncoding RNAs. *Nat Cell Biol*. 13:95-101.

Matera, A.G., M. Izaguire-Sierra, K. Praveen, and T.K. Rajendra. 2009. Nuclear bodies: random aggregates of sticky proteins or crucibles of macromolecular assembly? *Dev Cell*. 17:639-647.

McKay, D.J., S. Klusza, T.J. Penke, M.P. Meers, K.P. Curry, S.L. McDaniel, P.Y. Malek, S.W. Cooper, D.C. Tatomer, J.D. Lieb, B.D. Strahl, R.J. Duronio, and A.G. Matera. 2015. Interrogating the function of metazoan histones using engineered gene clusters. *Dev Cell*. 32:373-386.

Miele, A., C.D. Braastad, W.F. Holmes, P. Mitra, R. Medina, R. Xie, S.K. Zaidi, X. Ye, Y. Wei, J.W. Harper, A.J. van Wijnen, J.L. Stein, and G.S. Stein. 2005. HiNF-P directly links the cyclin E/CDK2/p220NPAT pathway to histone H4 gene regulation at the G1/S phase cell cycle transition. *Mol Cell Biol*. 25:6140-6153.

Misteli, T. 2007. Beyond the sequence: cellular organization of genome function. *Cell*. 128:787-800.

Morimoto, M., and C.F. Boerkoel. 2013. The role of nuclear bodies in gene expression and disease. *Biology (Basel)*. 2:976-1033.

Nott, T.J., E. Petsalaki, P. Farber, D. Jervis, E. Fussner, A. Plochowietz, T.D. Craggs, D.P. Bazett-Jones, T. Pawson, J.D. Forman-Kay, and A.J. Baldwin. 2015. Phase transition of a disordered nuage protein generates environmentally responsive membraneless organelles. *Mol Cell*. 57:936-947.

Patel, A., H.O. Lee, L. Jawerth, S. Maharana, M. Jahnel, M.Y. Hein, S. Stoynov, J. Mahamid, S. Saha, T.M. Franzmann, A. Pozniakovski, I. Poser, N. Maghelli, L.A. Royer, M. Weigert, E.W. Myers, S. Grill, D. Drechsel, A.A. Hyman, and S. Alberti. 2015. A Liquid-to-Solid Phase Transition of the ALS Protein FUS Accelerated by Disease Mutation. *Cell*. 162:1066-1077.

Rajendra, T.K., K. Praveen, and A.G. Matera. 2011. Genetic Analysis of Nuclear Bodies: From Nondeterministic Chaos to Deterministic Order. *Cold Spring Harb Symp Quant Biol*. in press.

Rajgor, D., J.A. Mellad, D. Soong, J.B. Rattner, M.J. Fritzler, and C.M. Shanahan. 2014. Mammalian microtubule P-body dynamics are mediated by nesprin-1. *J Cell Biol*. 205:457-475.

Robberecht, W., and T. Philips. 2013. The changing scene of amyotrophic lateral sclerosis. *Nat Rev Neurosci*. 14:248-264.

Salzler, H.R., D.C. Tatomer, P.Y. Malek, S.L. McDaniel, A.N. Orlando, W.F. Marzluff, and R.J. Duronio. 2013. A sequence in the Drosophila H3-H4 Promoter triggers histone locus body assembly and biosynthesis of replication-coupled histone mRNAs. *Dev Cell*. 24:623-634.

Scheer, U., B. Hugle, R. Hazan, and K.M. Rose. 1984. Drug-induced dispersal of transcribed rRNA genes and transcriptional products: immunolocalization and silver staining of different nucleolar components in rat cells treated with 5,6-dichloro-beta-D-ribofuranosylbenzimidazole. *J Cell Biol*. 99:672-679.

Shevtsov, S.P., and M. Dundr. 2011. Nucleation of nuclear bodies by RNA. *Nat Cell Biol*. 13:167-173.

Shulman, J.M., P.L. De Jager, and M.B. Feany. 2011. Parkinson's disease: genetics and pathogenesis. *Annu Rev Pathol*. 6:193-222.

Singer, A.B., and J.G. Gall. 2011. An inducible nuclear body in the Drosophila germinal vesicle. *Nucleus*. 2:403-409.

Sleeman, J.E., and L. Trinkle-Mulcahy. 2014. Nuclear bodies: new insights into assembly/dynamics and disease relevance. *Curr Opin Cell Biol*. 28:76-83.

Terzo, E.A., S.M. Lyons, J.S. Poulton, B.R. Temple, W.F. Marzluff, and R.J. Duronio. 2015. Distinct self-interaction domains promote Multi Sex Combs accumulation in and formation of the Drosophila histone locus body. *Mol Biol Cell*.

Tompa, P. 2005. The interplay between structure and function in intrinsically unstructured proteins. *FEBS Lett*. 579:3346-3354.

Toretsky, J.A., and P.E. Wright. 2014. Assemblages: functional units formed by cellular phase separation. *J Cell Biol*. 206:579-588.

Uversky, V.N. 2013a. A decade and a half of protein intrinsic disorder: biology still waits for physics. *Protein Sci*. 22:693-724.

Uversky, V.N. 2013b. Under-folded proteins: Conformational ensembles and their roles in protein folding, function, and pathogenesis. *Biopolymers*. 99:870-887.

van der Lee, R., M. Buljan, B. Lang, R.J. Weatheritt, G.W. Daughdrill, A.K. Dunker, M. Fuxreiter, J. Gough, J. Gsponer, D.T. Jones, P.M. Kim, R.W. Kriwacki, C.J. Oldfield, R.V. Pappu, P. Tompa, V.N. Uversky, P.E. Wright, and M.M. Babu. 2014. Classification of intrinsically disordered regions and proteins. *Chem Rev*. 114:6589-6631.

Voronina, E., G. Seydoux, P. Sassone-Corsi, and I. Nagamori. 2011. RNA granules in germ cells. *Cold Spring Harb Perspect Biol*. 3.

Walters, R.W., D. Muhlrad, J. Garcia, and R. Parker. 2015. Differential effects of Ydj1 and Sis1 on Hsp70-mediated clearance of stress granules in Saccharomyces cerevisiae. *RNA*. 21:1660-1671.

Weber, S.C., and C.P. Brangwynne. 2012. Getting RNA and protein in phase. *Cell*. 149:1188-1191.

White, A.E., B.D. Burch, X.C. Yang, P.Y. Gasdaska, Z. Dominski, W.F. Marzluff, and R.J. Duronio. 2011. Drosophila histone locus bodies form by hierarchical recruitment of components. *J Cell Biol*. 193:677-694.

White, A.E., M.E. Leslie, B.R. Calvi, W.F. Marzluff, and R.J. Duronio. 2007. Developmental and cell cycle regulation of the Drosophila histone locus body. *Mol Biol Cell*. 18:2491-2502.

Wippich, F., B. Bodenmiller, M.G. Trajkovska, S. Wanka, R. Aebersold, and L. Pelkmans. 2013. Dual specificity kinase DYRK3 couples stress granule condensation/dissolution to mTORC1 signaling. *Cell*. 152:791-805.

Woulfe, J., D.A. Gray, and I.R. Mackenzie. 2010. FUS-immunoreactive intranuclear inclusions in neurodegenerative disease. *Brain Pathol*. 20:589-597.

Yang, X.C., I. Sabath, L. Kunduru, A.J. van Wijnen, W.F. Marzluff, and Z. Dominski. 2014. A conserved interaction that is essential for the biogenesis of histone locus bodies. *J Biol Chem*. 289:33767-33782.

Zaidi, S.K., D.W. Young, A. Javed, J. Pratap, M. Montecino, A. van Wijnen, J.B. Lian, J.L. Stein, and G.S. Stein. 2007. Nuclear microenvironments in biological control and cancer. *Nature reviews. Cancer*. 7:454-463.

Zwicker, D., M. Decker, S. Jaensch, A.A. Hyman, and F. Julicher. 2014. Centrosomes are autocatalytic droplets of pericentriolar material organized by centrioles. *Proc Natl Acad Sci U S A*. 111:E2636-2645.