SOMATIC SEX DETERMINATION IN *D. MELANOGASTER*: INSIGHTS IN THE ESTABLISHMENT TO MAINTENANCE TRANSITION

A Dissertation

by

ALEJANDRA NOEMI GONZALEZ ROJOS

Submitted to the Office of Graduate Studies of
Texas A&M University
in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

May 2012

Major Subject: Biology

Somatic Sex Determination in *D. melanogaster*: Insights in the Establishment to Maintenance Transition

SOMATIC SEX DETERMINATION IN *D. MELANOGASTER*: INSIGHTS IN THE ESTABLISHMENT TO MAINTENANCE TRANSITION

A Dissertation

by

ALEJANDRA NOEMI GONZALEZ ROJOS

Submitted to the Office of Graduate Studies of
Texas A&M University
in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

Approved by:

Chair of Committee,	James W. Erickson
Committee Members,	Bruce Riley
	Sarah Bondos
	Keith Maggert
Head of Department,	U.J. McMahan

May 2012

Major Subject: Biology

ABSTRACT

Somatic Sex Determination in *D. melanogaster*: Insights in the Establishment to Maintenance Transition.

(May 2012)

Alejandra Noemi González Rojos, B.S., Pontificia Universidad Católica de Chile

Chair of Advisory Committee: Dr. James W. Erickson

In *Drosophila melanogaster*, sex is determined at the pre-blastoderm stage via an X-chromosome counting mechanism. During this process embryos that carry two X chromosomes begin to develop as females while embryos with one X start the male developmental program. The X-linked genes involved in sex determination, also called X-signal elements (XSEs), are: *sisterlessA* (*sisA*), *sisterlessB* (*sisB*), *unpaired* (*upd*), and *runt*. These genes are responsible for the transcriptional activation of the master regulatory gene *Sex-lethal* (*Sxl*). Expression of *Sxl* is initially accomplished only in females through activation of the establishment promoter *SxlPe*. Later in development, *Sxl* is transcribed in both sexes through a maintenance promoter, *SxlPm*, but functional *Sxl* protein is only produced in female flies. Since *Sxl* is at the top of the sex determination cascade, understanding its regulation is key to comprehend the process of sex determination. The experiments in this dissertation were designed to better understand two aspects of the sex determination mechanism: How the protein encoded

by XSE element *sisA* interacts with *SxlPe*, and how the transition from regulation by *SxlPe* to regulation by *SxlPm* occurs.

The *sisA* protein (SisA), as part of the b-ZIP protein family, is thought to bind to its target as a dimer, but a dimerization partner has not yet been found. This work uses knockouts and germ-line clones to examine interaction between *sisA* and three SisA partner candidates, *atf4*, *CG16813*, and *CG16815*. Although the evidence described here suggest that none of the three SisA partner candidates genetically interact with *Sis*, we cannot rule out the possibility of redundancy between the different candidate proteins.

This research unravels the timing and regulation of *SxlPm* expression. I have shown, contrary to previous thought, that expression of *SxlPe* and *SxlPm* overlaps for a brief period. Several of the same proteins that are involved in the regulation of *SxlPe,* including the XSE *sisB,* also regulate *SxlPm.* This sex specific regulation leads to a sexually dimorphic pattern of activation and early expression of *SxlPm*. A common enhancer region was found to regulate *SxlPe* as well as *SxlPm*. These results highlight the importance of the transition between *SxlPe* and *SxlPm* for the proper establishment of sex determination and have implications for how the sex determination mechanism evolved.

DEDICATION

This dissertation is dedicated to my parents Cristina Rojos and Alfonso González, who taught me that life presents many obstacles, and the strongest are not those who never fall, but those who do, and learn how to get back up.

ACKNOWLEDGEMENTS

I would like to thank my committee chair, Dr. Erickson, and my committee members, Dr. Riley, Dr. Maggert, Dr. Bondos, for their guidance and support throughout the course of this research.

I would like to thank my parents for their enormous sacrifices for my education and made me the person I am today, as well as the love and connection to my country that my sister provided me with.

Thanks also go to my friends and colleagues for being with me every step of the way, and made me feel life was a lot more than just graduate school. I also want to extend my gratitude to the department faculty and staff for making my time at Texas A&M University a great experience.

Finally, thank you to my husband for his patience, company and unconditional love.

TABLE OF CONTENTS

CHAPTER Page

LIST OF FIGURES

Page

LIST OF TABLES

CHAPTER I

INTRODUCTION

Drosophila somatic sex determination and dosage compensation

Understanding how cells, genes or organisms can sense or "count" concentrations of regulators in their environment is critical to elucidate how choices in development are made. Many crucial biological processes (like quorum sensing in bacteria, *Drosophila* body axis development established by morphogen concentration, etc.) rely on counting mechanisms. In *Drosophila melanogaster*, sex determination is the mechanism in which X chromosome dose is counted at the promoter region of a master regulatory gene, sex lethal (*Sxl*), establishing the sexual fate of the organism. During this process, XX flies produce early SXL and begin a female developmental program. In contrast, XY flies lack Sxl, leading to male development by default. *Sxl* encodes for a protein that has two domains homologous to a RNA recognition motif, RMM domain, which is highly conserved among the mRNA binding protein family [1,2]. *Sxl* directs female development by binding to the 3' proximal splicing site in *tra* pre-mRNA and thus inducing female specific pre-mRNA splicing of the gene *transformer* (*tra*) (Fig. 1). Occupancy of this splicing site by SXL prevents binding of the splicing factor U2AF. As a consequence, splicing at the alternative 3' distal splicing site occurs, leading to the formation of functional TRA protein. In males, the absence of SXL allows binding of the

---

This dissertation follows the style of PLoS Biology.

U2AF factor to the 3' proximal splicing site in the *tra* pre-mRNA. Splicing at this site produced non-functional *tra* mRNAs resulting in lack of TRA protein in male flies. TRA being an RNA binding protein itself regulates splicing of *doublesex* (*dsx*) (Fig. 1) [3-8]. TRA splices *dsx* mRNA into a female mRNA form that is translated into a transcription factor called DSX^F. This protein controls female somatic differentiation by inducing feminizing genes and by repressing genes which function in male differentiation. DSX^F acts during early development in differentiation of the female secondary sexual characters (like differentiation of genital disk, ovaries, etc.). DSX^F also acts throughout the life of the fly to maintain yolk proteins and the female fat body tissue. In males, on the other hand, lack of TRA leads to the formation of a male specific form of DSX called DSX^M. This protein only differs on the C-terminal domain with DSX^F (the male protein is longer because it possess additional coding sequences at the C-terminal domain). DSX^M controls male somatic differentiation by activating genes that promote male secondary characteristics (like sex combs, male genitalia, etc) and repressing genes that activate female characteristics. Remarkably, lack of the *dsx* gene produced intersex flies where genes that control female and male development are de-repressed and the flies exhibit both secondary characteristics [2-4,9,10]. TRA also regulates splicing of the gene *fruitless* (*fru*). In female flies TRA protein binds near the 5' splicing site closest to the 5' end of the mRNA and promotes the splicing at this site, thus introducing a stop

codon near the N-terminus of FRU, preventing functionality [11]. In males, however lack of TRA allows the splicing to occur at the default 5' splicing site and the mRNAs produced are translated into functional FRU (Fig. 1). This protein is mainly expressed in the male central nervous system and controls important aspects of the male courtship behavior [12-17].

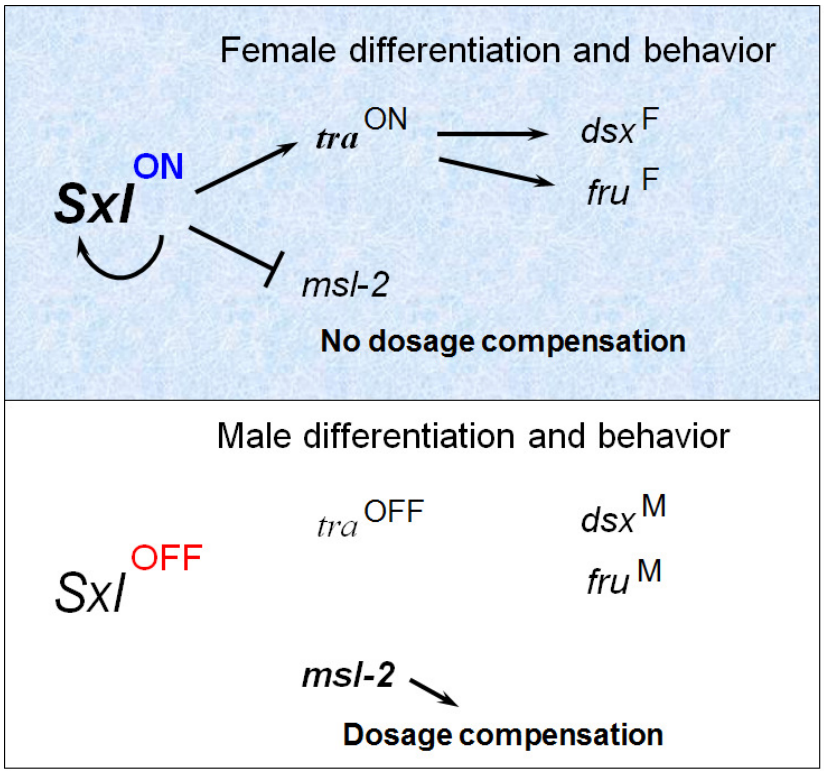


Fig. 1: Somatic sex determination cascade. In the top panel- female flies expressed SXL which can maintain itself through an auto-regulatory feedback loop. SXL blocks *msl2* mRNA translation. At the same time SXL promotes the formation of functional TRA protein. Transformer protein splices *dsx* and *fru* to the female mode. In the lower panel- Male flies do not express SXL. Because SXL is absent TRA is not produced and *dsx* and *fru* are spliced to the default male mode. At the same time MSL2 protein is produced.

Sxl is also the master regulatory gene of dosage compensation (Fig. 1) [18]. In many diploid organisms males have one X chromosome while their female counterparts have two X chromosomes. Dosage compensation is the mechanism which equalizes gene expression from the single male X chromosome with the female XX chromosomes (although evolutionary speaking, it probably evolved to compensate the single X chromosome expression so that the overall levels of expression were balanced with the diploid autosomal expression). In *D. melanogaster*, dosage compensation is achieved by increasing expression of the single male X chromosome. This is accomplished by the dosage compensation complex (DCC complex). DCC complex major components are MSL1, MSL2, MSL3, MLE (RNA helicase), MOF (histone acetyltransferase) and the rox noncoding RNAs [19-22]. Although the mechanisms that underline this process are very controversial, it has been shown that the DCC complex is first attracted to the male X chromosome by the hyper-transcription of the *rox* gene locus [22-25]. DCC then recognizes entry sites of sequences, sequences proven to be enriched in the male X chromosome [26]. It is still a mystery if the complex can then spread over the male X chromosome from the *rox* gene's locus [24,27], or if there are sequences around the main entry sites with lower affinity that attract the complex to assist in covering the entire X chromosome. As a result, increase in transcription of the X chromosome is achieved [22,26,28]. In females *Sxl* is at the top of this cascade of events and regulates this process by blocking translation of *msl2* [3,29,30]. There are binding sites for SXL at the 5' and 3' UTR of *msl2* mRNA. Previous studies demonstrated that the binding of

SXL to the 3' UTR prevented the assemble of the 43S ribosomal pre-initiation complex at the 5' of *msl2* mRNA, and that this mechanism was in part responsible for decreasing translation of *msl2* [8,31]. Activity of the SXL binding sites at the 5'UTR of *msl2* mRNA was only recently uncovered. SXL binding sites are near three upstream open reading frames located at the 5'UTR of *msl2* mRNA. Jan Medenbach et. al. demonstrated that SXL had the ability to increase initiation of ribosome scanning at the nearest upstream open reading frame, and as a result, a decrease in ribosomal scanning of the major *msl2* open reading frame was observed [32]. These two methods of hindering translation assure MSL2 absence in female flies. On the contrary, absence of SXL in males results in production of MSL-2 and the assembly of the dosage compensation complex on the male X chromosome. As a result, the rate of transcription of a male's single X is elevated two-fold, balancing its expression with autosomal genes. It is important to note that lack of SXL in females will cause overexpression of X linked genes and death of the fly. Equally, presence of SXL in males will prevent dosage compensation from occurring and expression from the single X chromosome will not be sufficient to equalize autosomal expression leading to death. This phenotype is what gave the name to the *Sex lethal* gene.

Sxl initiation and maintenance: when and how does the fly know its sex

Initially *D. melanogaster* embryos are syncytial. It is in this complicated environment, filled with maternally deposited factors and marked rapid nuclear divisions where sex determination takes place. Around nuclear division 8, the first zygotic genes

start to express, and with them the X-linked signaling elements (XSE) that signal the genetic information in response to the number of X chromosomes [3,8,33]. The fly first activates *Sxl* in nuclear division 12. At this time, only XX female flies carrying a double dose of the XSEs accumulate enough XSE proteins to trigger expression of a very specialized promoter, *SxlPe*. This step is called "establishment", and is the decisive step by which SXL early protein is formed, initiating the female sexual pathway. In contrast, the male fate is established because XY males carry only a single copy of the XSE elements and are unable to reach the concentration threshold necessary to activate *SxlPe* (Fig. 2). Lack of SXL early protein will promote the default male development [3,33,34].

Remarkably, the female pathway is driven by only a brief pulse of *SxlPe* promoter activity. *SxlPe* is activated in cycle 12 but is permanently shut down in the beginning of nuclear division 14 when the embryo starts cellularizing. Thereafter, continued *Sxl* expression relies on a process called "maintenance". The maintenance stage is defined by the presence of properly regulated transcripts from the maintenance promoter; *SxlPm*. *SxlPm* is expressed in both male and female flies. In males, transcripts from this promoter include a specific male exon 3. This exon contains a translational stop codon that prevents production of functional *Sxl* protein (Fig. 2) [3]. In females the male exon 3 is skipped and functional *Sxl* protein is made [35]. Skipping the male exon 3 in females is due to *Sxl* auto-regulating its own mRNA splicing. This is auto-regulatory feedback loop is what maintains proper *Sxl* expression for the remainder of the fly's life. [1,36,37]. The role of *SxlPe*, then, is to provide the brief pulse of early SXL that initiates

the autoregulatory loop. Once *SxlPm* transcripts are spliced in the productive, female, mode there is no longer a need for the early form of SXL as the properly spliced female *SxlPm* transcripts can provide a self-regulating source of *Sxl* protein.

The mechanism controlling *Sxl* splicing control has been an object of intense study. It was initially thought that SXL binding sites on its pre mRNA overlapped the Spliceosome recruitment site around exon 3. It was assumed that SXL binding around the male exon precluded binding and formation of the splicing machinery and with that inclusion of exon 3 in the mRNA (Fig. 2) [35]. However, it has been recently demonstrated that SXL binding sites are located 200 bp upstream and downstream of the male exon 3, and SXL binding to its mRNA does not prevent formation of the basic splicing machinery. SXL also interact with basic splicing factors like PPS, U1 snRNP and SNF. One of the mechanism previously proposed, that SXL could block binding of other splicing factors to the splicing recognition site or it could abolish Spliceosome catalytic activity [38,39].

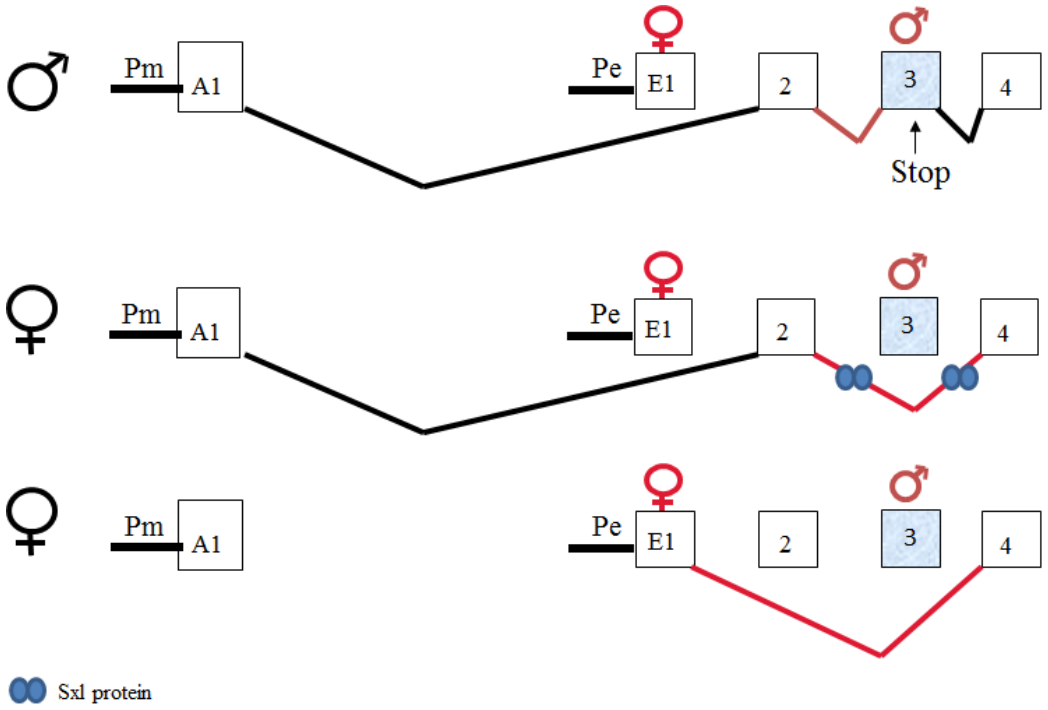


Fig. 2: Male and female splicing of *Sxl*. In males transcripts from *SxlPm* include the male exon 3 which has a stop codon that prevents formation of functional SXL protein. In female flies transcripts from *SxlPm* skip exon 3 and can produced functional SXL protein. This transcripts skip exon 3 with the aid of SXL that binds to either side of the male specific exon an prevent splicing from happening at this site. In females initial source of SXL is produced from the establishment promoter (*Pe*). Default splicing of the transcripts from *Pe* promoter do not include exon 3.

SxlPe – the switch that initiates sex determination

SxlPe expression is sexually dimorphic – it is active in female, but not in male, embryos. *SxlPe* is one of the earliest developmental targets of the fly. By the end of nuclear division 12 *SxlPe* is homogeneously expressed in all the somatic cells of the embryo. This expression increases through cycle 13 and reaches a peak at the beginning of cycle 14. Thereafter expression rapidly decays, but accumulated *SxlPe*-derived mRNA products can be detected through gastrulation [33]. Female-specific activation of

SxlPe is accomplished by the diplo-X levels of the X-chromosome signal elements or XSEs. Among the XSEs are *sisA*, *sisB*, *upd,* and *runt* (Fig. 3). These X-linked factors, once called numerator elements (for more explanation see below) share many genetic characteristics: 1) Increasing the copy number of these genes results in male death due to inappropriate expression of *SxlPe* in males. 2) Decreased copy numbers of these genes result in female death because *SxlPe* is less effectively activated. 3) Female lethal effects should vanish when combined with a *Sxl* constitutive allele like Sxl^{M1} that activates the *Sxl* splicing cascade independent of the normal contribution from *SxlPe*. 4) Likewise, male lethal effects of increased XSE dose are fully suppressed by *Sxl* null allele like Sxl^{F1}. 5) All XSEs encode positive regulators of *SxlPe* consistent with their biological function of conveying the dose of X chromosomes to *Sxl*. [3,40]. Among the XSEs *sisA* and *sisB* are the strongest activators since they have a strong influence over *SxlPe* activity. These two factors are expressed before nuclear division 10 and are absolutely required to express *SxlPe*, characteristics that distinguish them from their XSEs counterparts. The two weaker XSEs, *runt*, and *upd*, share the same genetic properties but have either smaller, later, or more spatially limited effects on *SxlPe*. For these reasons, *upd* and *runt* are sometimes referred to as reinforcement factors [41].

The strong XSE *sisA* is located at the X chromosome in the cytological location 10B2 (Flybase) and it fulfills all the characteristics required to be an XSE [40]. Genetic interaction between *sisA* and *SxlPe* promoter has been previously shown [42] but the molecular basis of this interaction are still a mystery. The *sisA* gene product is a highly divergent member of the b-ZIP family of transcription factors. b-ZIP proteins encode a

coiled-coil domain that is involved in dimerization as well as functioning as the DNA binding domain [33]. The structure of *sisA* protein predicts that it cannot form homodimers, suggesting that it requires an as yet unknown dimerization partner to activate *SxlPe*. The search for this dimerization partner constitutes Chapter X of this dissertation.

sisb is located on the X chromosome cytological location 1A8. *sisb*'s involvement in sex determination has been linked to the T4 transcriptional unit from the Achaete-Scute complex, also called *scute* [43,44]. *In-situ* hybridization showed that *sisb* transcripts start to appear uniformly through the embryo by nuclear division 9. Its expression peaks around cycle 12 and suddenly decreases in early cycle 14 [33]. The *sisb* gene product belongs to the basic helix lop helix (bHLH) family of transcription factors. These proteins are characterized by having a basic DNA binding domain and a helix loop helix dimerization domain. SisB protein forms a hetero-dimer with the maternally deposited factor encoded by *daughterless* (*da*). This hetero-dimeric protein complex has the ability to bind *SxlPe* and activate its transcription [45].

The XSE gene *unp* encodes for a ligand of the JAK-STAT pathway. The pathway is activated by binding the ligand Ump to its receptor Domeless. Domeless is a receptor tyrosine kinase that transduces the signal by phosphorylating Hopscotch, the drosophila JAK homolog. Phosphorylated Hopscotch (Hop) produces a conformational change in the Dome homolog less receptor, which allows for binding of the maternally deposited transcription factor STAT92E. Binding of STAT92E to the Domeless receptor promotes STAT92E phosphorylation. Phosphorylated STAT92E has the capability to

enter the nuclei and directly activate promoter expression [41,46]. This pathway's involvement in sex determination has been demonstrated genetically. Avila F.W. et al. showed that not only *ump* but also *hop* and *Stat92E* genetically interacted with *SxlPe***.** Curiously mutations in the JAK- STAT pathway seemed to alter *SxlPe* expression only in nuclear division 14 and it affected only the central part of the embryo. This fact agreed with *upd* expression starting only in nuclear division 13. Because *SxlPe* can be activated, but not fully maintained, in the absence of the JAK-STAT the signaling pathway is viewed as functioning to reinforce *SxlPe* activity [41].

The XSE *runt* was discovered in *Drosophila melanogaster* because of its involvement in segmentation. As a pair rule gene, Runt regulates transcription of a number of other segmentation genes [47,48]. Runt belongs to the RUNX family of DNA binding proteins. This relatively new family of proteins has a DNA binding domain homologous to the murine PEBP2αB protein domain [49]. Runt can directly bind and activate the promoter of its target genes [48] or repress target genes. As a repressor, its activity depends upon its WRPW domain, which interacts with the co-repressor Groucho [50]. Runt has been described as an XSE, consequently it has the ability to activate *SxlPe* [40]. Mutations in *runt* cause a decrease in *SxlPe* gene expression beginning in nuclear division 13 in the central region of the embryo suggesting that it, like *upd*, may not be a primary activator of *SxlPe*. Instead it appears, through an unknown mechanism, to maintain *SxlPe* activity once the promoter has been activated. Although there is a published paper claiming that Runt can directly bind to *SxlPe* promoter, the experiments published only showed weak binding that is more consistent with non-specific DNA

binding [51] (James Erickson unpublished data), thus the mechanism by which Runt regulates *SxlPe* remains unknown.

Negative regulation of *SxlPe* promoter is accomplished through one autosomal gene, *deadpan* (*dpn*), and at least three maternally deposited factors, *extramachrochetae* (*emc*), *groucho* (*gro*) and *hey* (Fig. 3) [52-54]. *Dpn* belongs to the bHLH family of transcription factors which binds to *SxlPe* promoter. Once bounded to *SxlPe* promoter, Dpn recruits the maternally deposited factor Gro. These factors together decrease *SxlPe* expression. Mutations of *dpn* have only a modest influence in *SxlPe* expression when compare to Dpn binding site mutants or a decrease of the *gro* maternally supplied product. This together with the fact that *dpn* only began expressing by nuclear division 12 [53] suggest that another protein is probably binding to Dpn binding sites and inhibiting *SxlPe* expression. One of these proteins is Hey, which also belongs to the bHLH family of transcription factors [52].

Emc is an HLH protein that lacks the basic binding domain. Mutation in the *emc* gene produces only a weak reduction in *SxlPe* activity. This repressor has been thought to operate by capturing bHLH activators like SisB [55]

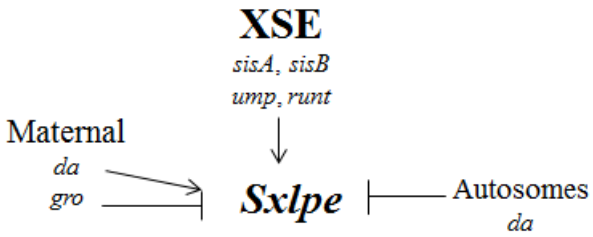


Fig. 3: *SxlPe* regulators. XSEs, together with the maternal factor *da*, activate *SxlPe* while the maternal factor *gro*, *emc* together with the autosomal gene *dpn* repress *SxlPe* expression.

SxlPe promoter structure

The *SxlPe* minimal regulatory sequences have been mapped to a 1.4 kb region, upstream of the transcription initiation site (+1) in the early exon 1, E1, of *Sxl*. Expression of a *lac Z* reporter gene driven by the –1.4 kb of the *SxlPe* promoter (–1.4 *SxlPe* promoter) showed that –1.4 kb is the minimal promoter region necessary to resemble endogenous *SxlPe* expression. The –1.4 *SxlPe* promoter can also be subdivided in two functionally different parts: 1) the proximal 400 bp, necessary to regulate sex specificity and 2) the distal 1000 bp that defines promoter strength. To study the influence of *SxlPe*'s activators in the expression of the –400 bp region, a transgene including a *lac Z* reporter driven by the –400 bp promoter region (*Pe*-400 *lacZ*) was introduced in *sisB* heterozygous mutant flies. *lac Z* expression was reduced in the *sisB* heterozygous mutant background when compared to the *lac* Z expression of the same construct in a *WT* organism. These results demonstrated that *sisB* positively interacted with the -400 bp region [42]. Later, six SISB/Da binding sites within the –400 bp region

and five binding sites in the region between –0.8 kb to –1.1 kb of the promoter were found [45]. Surprisingly, only 3 of the binding sites codified for a canonical E-box consensus sequence while the other eight binding sites contained functional, yet non-canonical SisB binding sites sequences [45].

The same transgenic *Pe*-400 *lacZ* reporter described above was used to study the relationship between the -400 bp region and *sisA*. The transgene was introduced in flies that were heterozygous for a *sisA* mutant allele. Just like in *sisB* mutants, *lacZ* expression was reduced in a *sisA* heterozygous background when compared to the same trans-gene in a *WT* background. This result indicates that *sisA* is a positive regulator of the -400 bp of *SxlPe* [42]. Although *sisA* and *SxlPe* genetically interact, it is unknown whether this interaction is direct.

There are three predicted STAT92E binding sites in *SxlPe* -1.4 region, located at -253, -393, and -428 bp. Mutation of these sites causes a decreased in *SxlPe* activity in the central region of the embryo [41].

Four functional Dpn binding sites have also been found in *SxlPe*. Two of these binding sites, −110 and −121 bp, were found to be canonical Dpn binding sites, while the other two sites −160 and −330 bp were non-canonical Dpn binding sites. One Dpn binding site (-110) overlaps with the 3’ most SisB/Da binding site. This arrangement of overlapping activator and repressors binding sites suggests that competency between these protein complexes could be the mechanism used for repression [52].

Insights into the counting mechanism

The idea that in *Drosophila melanogaster* sex is determined by the X chromosome to Autosomal ratio (X/A) was first proposed in 1921 by Bridges [56] and it has been the dominate hypothesis until recently. This idea rose from the observation that female flies developed from individuals that carried an X/A ratio of 1.0 (ex. 2X/2A, 3X/3A) and male flies from those with a X/A ratio of 0.5 or less (XY/2A; XY/3A). Most importantly, flies carrying 2X and 3A (triploid flies with 2 X chromosomes) with the intermediate X/A of 0.67 developed as intersexes. Through the years, our understanding of the sex determination process has grown enormously. The molecular identity of the sex determining factors (activators and repressor) has been uncovered, and with that the speculation about the molecular mechanisms that drive the sex determination process. Initially it was proposed that both (feminizing) positive and (masculinizing) negative regulators bind directly to *SxlPe* and that the effects of the different transcription factors was interpreted by the promoter, causing an all or none effect in *SxlPe* activation [42]. The second model invokes titration. In this model the negative regulators physically interact with the activator proteins. Consequently, when the levels of repressor proteins are equal or greater than the levels of activators, as it occurs in males, repressors titrate out the activators preventing them from binding to the promoter and rendering *SxlPe* inactive. On the contrary in female flies, activators are in greater numbers than repressors (more X chromosomes thus more activators) and some activators remain free to bind and activate *SxlPe* [57]. Although these models seem to explain the basic X/A ratio observations, they did not fit with the molecular identity and function of X linked

activators and autosomal repressors. The main missing link was the nature of the autosomal repressor Dpn. Dpn is a weak repressor that does not have enough strength to balance the strength of the XSEs. The other known repressors were maternally deposited and consequently in equal amount in female and male flies. In addition these repressors bind *SxlPe* and execute its activity directly rather than by capturing the XSEs. Taking all of these factors in to consideration, it was clear that a reassessment of the X/A ratio hypothesis was necessary. An X counting mechanism seemed to better fit the modern understanding of sex determination, but how could the observed phenotypes of triploid and haploid organisms and cells be explained with the X counting mechanism? The answer developed from a molecular characterization of haploid and diploid organisms. In this analysis it was discovered that haploid organisms expressed Sxl because they have an extra nuclear division before the embryo reached cellularization. This resulted in a greater time window during which XSEs could accumulate, permitting them to activate *SxlPe* and become females. Triploids showed the reciprocal effect. Triploid embryos cellularize a nuclear division earlier. As a result, the time necessary for the accumulation of the XSEs decreases, leaving 2X/3A animals with insufficient *SxlPe*-mRNA and early Sxl protein, to reliably trigger Sxl autoregulation [8,52].

To explain how activators and repressors can lead to an all or none response in *SxlPe* activity a new model rose. This model proposed that changes at the promoter are translated as chromatin modifications that amplify male/female differential response to XSEs. In female flies, a double dose of XSEs activate *SxlPe* changing the chromatin environment to an active state, which increases the activation potential of XSEs (Fig. 4

A). Contrarily in males, the limited amount of XSEs prevents *SxlPe* activation keeping the chromatin in an inactive state and amplifying the repressive potential of repressors (Fig. 4 B) [52].

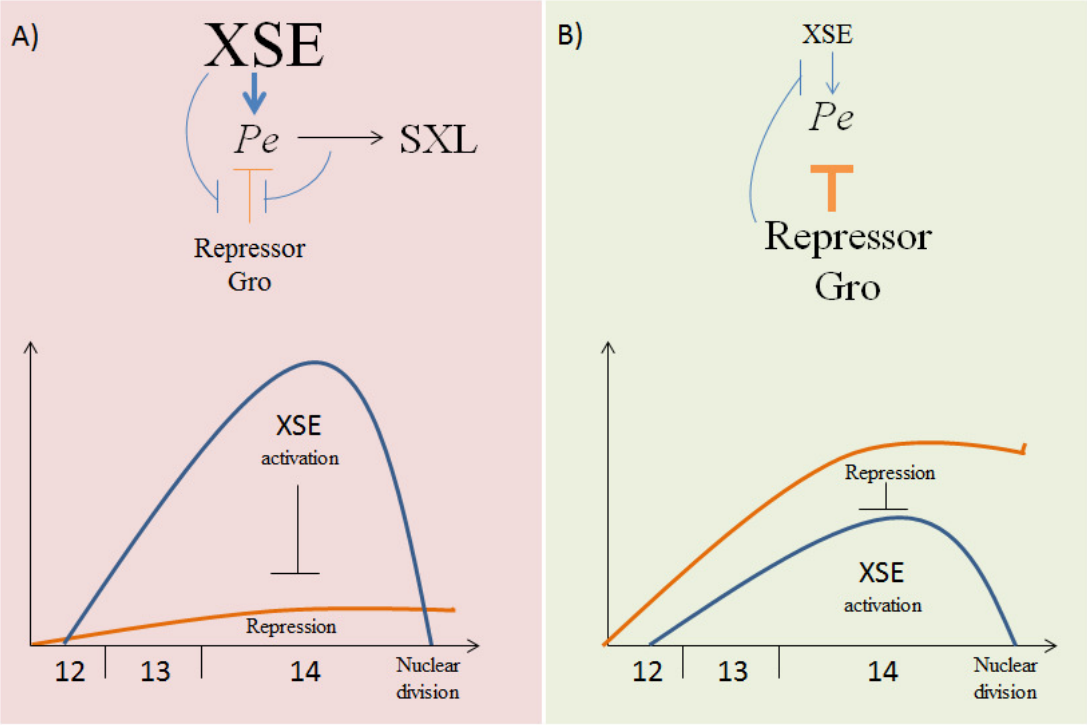


Fig. 4: Modern model of the molecular sex determination mechanism. A) Female flies have a double amount of XSEs which is enough to overcome the repression threshold and activate *SxlPe*. Active chromatin is acetylated, countering repression, and further amplifying activation. Also, XSEs like Runt could interact with repressors like Gro to further increase the activating response. B) In males, only single amount of XSEs is produced and this is not enough to overcome the repression, rendering the *SxlPe* promoter inactive. Inactive chromatin gets methylated which amplifies the repression and dampen the activation down.

CHAPTER II

IDENTIFICATION OF SISTERLESS A PARTNER PROTEIN

SisA is a strong regulator of *SxlPe* and a key XSE factor in sex determination. Mutations in *sisA* result in *SxlPe* inactivity and a decrease in female viability. Increase in the copy number of *sisA* ectopically activates *SxlPe* in males, reducing male viability [34]. *sisA* is known to regulate *SxlPe*. As explained earlier, when a *lacZ* reporter driven by *SxlPe* is introduced in a *sisA* mutant background, the activity of *SxlPe* decreases indicating *SxlPe-sisA* interaction [42]. Likewise, *sisA* mutants show female lethal interactions with mutations that specifically affect the early functions of *Sxl*, but not with mutations affecting only Sxl late functions [34]. Because its preponderant role in sex determination it is crucial to understand the molecular nature and the mechanisms by which *sisA* regulates *SxlPe*.

SisA belongs to the basic leucine zipper (b-ZIP) transcription factor family. The b-ZIP proteins have an α-helical structure, known as the basic leucine zipper composed of a short basic DNA binding region followed by a leucine zipper dimerization domain. The dimerization domain is characterized by the presence of 4 or 5 repetitive arrays of a seven amino acids motif (‘a’, ‘b’, ‘c’, ‘d’, ‘e’, ‘f’, and ‘g’) (Fig. 5). The amino acids in the ‘a’ and ‘d’ positions create an hydrophobic core on one face of the helix which stabilizes to the dimer [58]. Although SisA retains the main features of the b-ZIP family, it adopts a non-canonical dimerization domain. This b-ZIP domain has charged amino acids in the ‘a’ position. SisA has an arginine at the ‘a’ position in the second heptad and

a glutamic acid in the ‘a’ position in the third heptad. This modification has been proposed to prevent homodimerization and eventually distinguishes SisA as a unique member of the b-ZIP family [58].

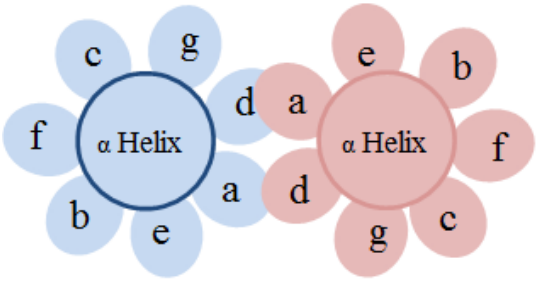


Fig. 5: b-ZIP protein dimerization topology. The topology adopted by the seven amino acids within one loop of the protein’s coil-coiled domain is shown. The main interacting amino acids are at the position ‘a’ and ‘d’.

Consistent with structural predictions that SisA cannot form homodimers [58], our laboratory has been unable to footprint or gel-shift *SxlPe* with purified SisA protein or with *in-vitro* translated protein (D. Yang and J. Erickson unpublished). In addition, both of our experimental results and theoretical predictions ([58], C. Vinson pers. communication) suggest SisA most likely has a novel dimerization partner. I attempted to identify that partner by analyzing the most plausible candidate proteins derived from two-hybrid analysis and by carrying out a search for protein partners using a *Drosophila* cell culture expression system. Most of the remainder of this chapter is organized into a detailed description of my genetic analysis of the candidate genes identified in two-hybrid screens. Because those results were ultimately inconclusive, I conclude this

chapter with a description of two additional experimental approaches I used to identify SisA partners so as to provide background and advice for future attempts to identify how SisA binds to DNA.

Candidates from a two-hybrid screen

To find SisA partner proteins, my laboratory first looked for proteins that physically interacted with SisA. The technique used for this purpose was the Gal 4-based two-hybrid assay [59]. To pull out most of SisA's interacting proteins, both the SisA coiled–coil domain and the full-length *sisA* proteins were fused to the DNA binding domain of yeast Gal4. A collection of cDNA library clones from embryos of 0 to 6 hours old was fused to DNA encoding the activation domain of Gal4. Many proteins interacted with SisA and its coil-coil domain. To eliminate false positives in the experiment described above, proteins with similar characteristics to *Drosophila* SisA (*sisA* from *D. pseudoobscura* and *D. virilis*) were used as positive controls. Different b-ZIP proteins (Giant and CNC) were used as negatives controls. The use of these controls helped to reduce the number of candidate genes from eighteen to eight. Within the eight interacting proteins two were predicted bHLH proteins, two encoded Zn-fingers, three were predicted b-ZIP proteins, and the other was a gene with no known function. Since the SisA partner is expected to be a b-ZIP transcription factor I selected the interacting b-ZIP proteins as the most likely candidates.

The three b-ZIP proteins were Atf-4, a gene and protein previously identified as functioning in pupal development [60], and two related, previously unidentified genes, *CG16815* and *CG16813,* neither of which have identifiable homologs in other species.

It is important to note that both ATF-4 and CG16815 were isolated multiple times and strongly interacted with SisA b-ZIP domain and full-length SisA. CG16813, in contrast, was isolated only once and the color of the β- Galactosidase reaction was much weaker than the binding detected with CG16815, both of which indicate a weaker interaction.

As explained above, SisA has a non-canonical dimerization domain. It was predicted that proteins with a non-canonical dimerization domain structure can easily interact with other proteins that share a non-canonical dimerization domain [58]. All three of the identified proteins have b-ZIP sequences that are predicted as non-canonical dimerization domains making them plausible candidates to interact with SisA ([58] and C. Vinson peers communication).

An important characteristic to consider while looking for SisA partner was its time of action. *SxlPe* is active between nuclear cycles 12 and early 14. Transcription of *sisA* starts earlier, at nuclear cycle 8, and it is expressed everywhere in the embryo during the period when *SxlPe* is active. Thereafter *sisA* expression is limited to the yolk nuclei, until the yolk nuclei decay late in embryogenesis [33]. Accordingly, SisA's partner needs either to be expressed in the precellular embryo, or be maternally deposited in the egg. Expression of *atf4* is known to be a maternally [60] consistent with

a possible function in sex determination. Since the expression patterns of CG16815 and CG16813 were unknown, I determined it by using *in-situ* hybridization (Fig. 6).

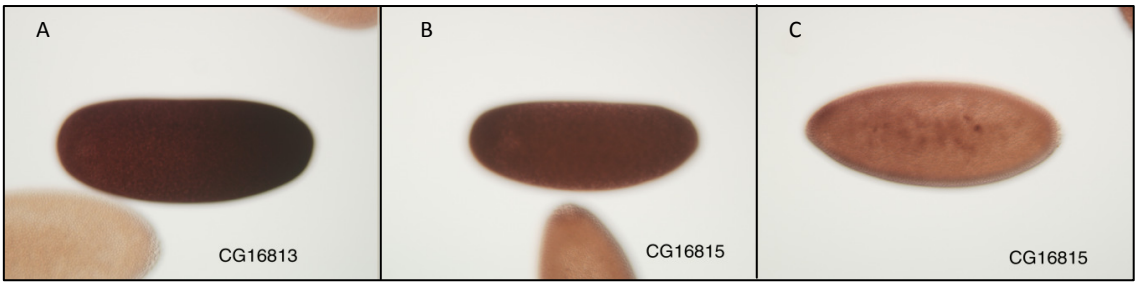


Fig. 6: *In-situ* hybridization of Sis A's candidates genes. A) *CG16813* maternal product. B) *CG16815* maternal product. C) *CG16815* zygotic expression attached to yolk, 15-20 min in nuclear cycle 14.

As shown in Fig. 6, the *CG16813* gene is maternally deposited and distributed throughout the egg. *CG16815* is also a maternally deposited mRNA, initially distributed throughout the embryo, but it is expressed in the yolk nuclei by cycle 14. This later expression of *CG16815* mRNA is strikingly similar to the later distribution to *sisA* mRNA which is also limited to the yolk nuclei after the beginning of cycle 14. *CG16815* mRNA has a somewhat similar distribution to *sisA* mRNA which is present predominately in the yolk nuclei after the beginning of cycle 14. *atf4* expression did not need to be examined because it was known to be a maternally deposited mRNA.

The maternally contributed dimerization partner of the X-signal element Scute is Daughterless (Da). The *da* gene exhibits a dominant female-specific lethal genetic interaction with zygotic *scute* mutations raising the likelihood that a similar dominant genetic interaction might occur between *sisA* and its maternally-supplied partner. To

assess genetic interaction between the candidates and *sisA* I needed deletions or loss of function mutations in *CG16815, CG16813,* and *atf4*. The only mutation available for *CG16815* and *CG16813* also included 26 other genes [61] and did not show any evidence of genetic interaction with *sis A*. There were many mutants available for *atf4*. I chose *atf4R6* which deleted the entire DNA region encodes for the dimerization domain of ATF4. The deletion was produced by imprecise excision of a P-element that removed all exons of *atf4*, except for 1 to 4. There were many mutants available for *sisA* mutant, I chose $sisA^{1}$ [3,33,42,62]. $sisA^{1}$ allele changes an amino acid in the binding domain of the protein, precluding SisA from activating *SxlPe* [62]. Since there were only *atf4* and *sisA* mutants available I decided to study interaction between these two genes first. The cross was performed as follows: $sisA^{1}$/Y male flies were crossed to *atf4R6*/ CyO female flies. Unfortunately, no decrease in $sisA^{1}$/+; *atf4R6*/ + female viability was found, indicating that there was no dominant genetic interaction (Table 1). One possibility is that in heterozygous animals sufficient amounts of the products of these genes are produced by the embryo or deposited by the mother in the egg, to compensate for the loss of one copy of *sisA*. To elucidate if there was a genetic interaction between SisA and *CG16813, CG16813* and *atf4*, I needed to first generate a knockout of *CG16813, CG16813* and then eliminate the maternal contribution of these three partner genes. This is generally accomplished by abolishing the expression of those genes in the maternal germ-line. The dominant female-sterile technique [63] can be used to produce fully viable heterozygous mothers that carry a germ-line that is homozygous for a deletion of the candidate genes. Thus, the maternal contribution of the partner genes is eliminated in the

great majority of the embryos. The dominant female-sterile technique is accomplished by using the dominant female-sterile allele *OvoD1* and FLP- FRT recombination as explained in Fig. 7.

Table 1: *atf4* does not show dominant genetic interaction with *sisA*

sisA^{1}/Y**;** +/+ X +/+**;** *atf4R6*/CyO

	Second chromosome phenotype	
First chromosome phenotype	***atf4R6*/+ Number of progeny**	**+/CyO Number of progeny**
sisA^{1}/+	88	92
+/Y	90	89

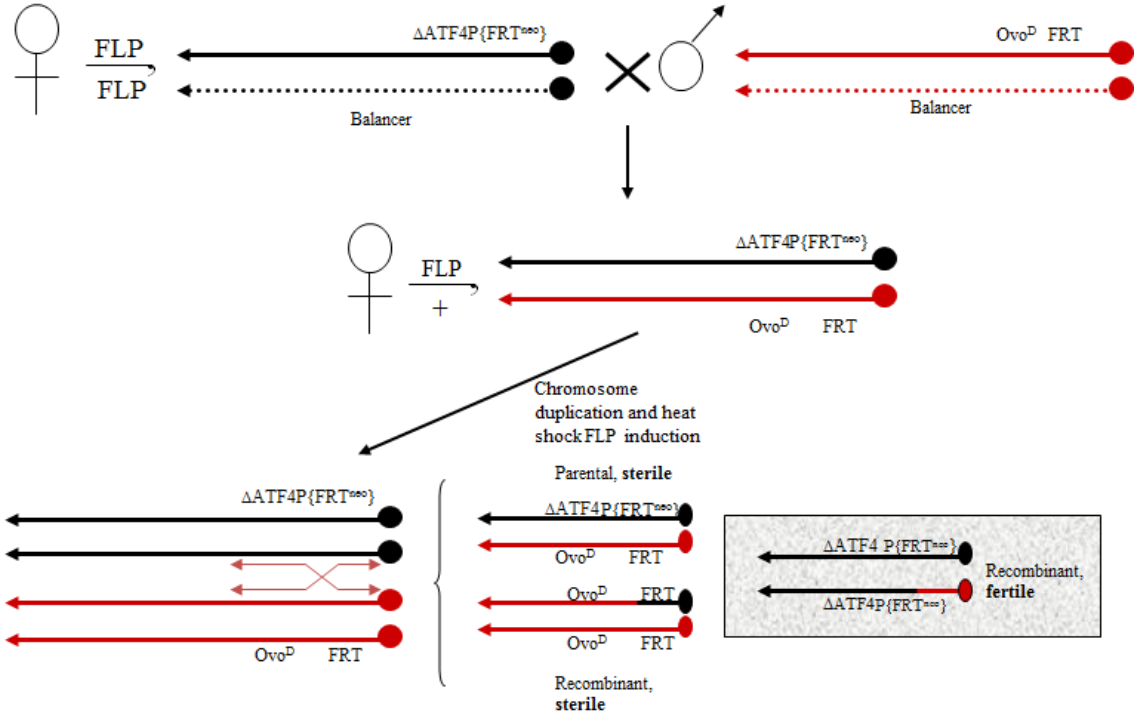


Fig. 7: Dominant female-sterile technique. Female flies carrying the ΔATF4/ Ovo^D second chromosome are heat shocked to induce FLP expression. FLP enzyme catalyzes FRT specific recombination. Germ-line cells containing the Ovo^D mutation will die while ΔATF4 homozygous cells will survive.

Because of the unique nature and the chromosomal location of the three candidate genes I took different approaches to generate the germ-line clones in each case. These approaches will be described in the sections below.

atf4 germ-line clones

atf4 is also called *cryptocephal* (*crc*). Its cytological location is 39C2-4, at the base of chromosome 2L. The *crc* mRNA is spliced to produce three different proteins: CRC-A, CRC-B, and CRC protein. CRC-A and CRC-B represent 85% and 10%,

respectively, of the total CRC protein in the adult fly head, and are the only isoforms of the protein that have a b-ZIP domain. Mutation or deletion of *aft4* altering the expression of the CRC-A or CRC-B protein causes defects in larval molting, and metamorphosis in the fly [60].

To eliminate the maternal contribution of *atf4* using the dominant female-sterile technique, the *atf4R6* deletion and the FRT genetic element needed to be on the same chromosome arm. Unfortunately, *atf4* (at 39C) is located immediately adjacent to the available FRT (40A), making it nearly impossible to generate the recombinant flies by standard recombination. Instead these recombinants were made using P-induced male recombination (in Drosophila, recombination does not usually occur in males, but it can be induced artificially using an exogenous recombination function). In this technique, a recombination event is induced in males at the P-element, by supplying the P-transposase (brought by the stable Δ2,3 insertion) [64]. A drawback of using P-induced male recombination is that processes like chromosome inversion, deletions and duplications can also occur. The greatest obstacle to isolating the required recombinants was that both the *atf4R6* deletion and the FRT (40A) element carry all or part of the original P elements that were used to generate the *atf4* deletion [60] and 40A insertion (see fly base). This means that recombination or associated chromosomal modifications could be generated at either the *atf4* locus or the 40A insertion site. To maximize my chances of recovering the wanted recombinants while avoiding rearranged or deleted chromosomes, I selected for the presence of the *atf4R6* deletion and its associated P element and for the FRT (40A) containing P element. This was done by choosing flies

with dominant *w+* eye marker contained in the *atf-4* P insertion, and for the neomycin resistant (neo^R) marker encoded by the 40A FRT (40A) insertion. Presence of this the neo^R gene allows flies to grow in food supplemented with G418 antibiotic (Fig. 8). I also selected against the dominant bristles marker *Scutoid* (*Sco*) located distally from the FRT element on chromosome 2L (cytological location 35B2). This gene has a phenotype that can be easily spotted by short or missing scutelar bristles in the flies' back. Consequently flies that can grow in G418 but do not have short or missing scutelar bristles must be recombinant. If they also have the red eyes, they might be carrying the desired recombinant chromosome (Fig. 8)

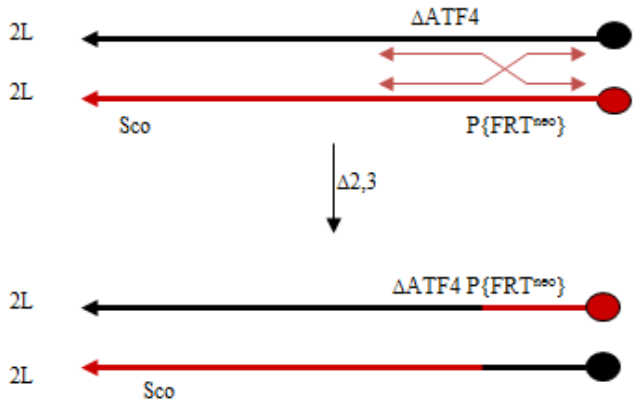


Fig. 8: P element male induced recombination. In presence of the transposase (Δ2,3) recombination will be induced at the P element or its proximity. Top recombinant chromosome can be distinguished because of the loss of the *Sco* marker and resistance to G418 antibiotic.

Although the crosses and screens were done so as to maximize the chances of identifying the proper recombinant, small deletions or rearrangements that damaged but

did not remove the FRT(40A) element or the *atf-4* locus would have passed the initial screen. It was thus necessary to corroborate the presence of the *atf4R6* deletion along with the FRT element in the recombinants. The FRT (40A) element was confirmed by PCR. The 40A FRT insertion has many repetitive regions which made the task of amplifying across the FRT repeats difficult (Fig. 9). To overcame this limitation I decided to amplify the 5' and 3' regions of the insertion as described in Fig. 9 below.

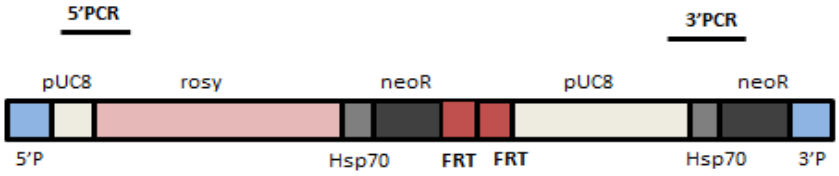


Fig. 9: 40A FRT insertion. Molecular map shows repetitive sequences given by the multiple copies of neo^R gene, Hsp70, P elements and the central FRT duplication.

The existence of *atf4R6* deletion in the recombinants was confirmed through complementation as well as single embryo PCR. To asses complementation the recombinant lines were crossed to the original *atf4R6* fly line, as well as a mutant fly line for *atf4* called crc^1. This mutant has a single base change that replaces Q 171 to R [60] and it is hypomorphic. Complementation data disclosed by Hewes et. al.[60] showed that only 2% of the crc^1/ crc^1 flies survive and that there are not crc^1/ *atf4R6* survival. I also tested for the presence of ATF4 b-ZIP domain. A single embryo PCR was performed to detect the region that encodes for ATF4 b-ZIP domain. In the *atf4R6* FRT/CyO stock ¼ of the embryos are expected to be *atf4R6* FRT/*atf4R6* FRT, these embryos lack the b-

ZIP domain and should result in absence of a PCR band. As a control, a PCR from the original *atf4R6* line in addition to W1118 (wild type for *atf4*) embryos were analyzed.

In a small scale experiment I analyzed 519 neomycinresistant flies. Among them, 27 lines looked like recombinants. They lost *Sco* and had red eyes that marked the *atf4R6* deletion. After PCR and genetic analysis described above none of the 27 lines had the recombinant chromosome I was looking for and in fact, all seem to have lost the *atf4R6* deletion, because they complemented the original deletion as well as the crc^{1} mutant. Single embryo PCR of these stocks showed a positive b-ZIP domain band for 100% of the tested embryos (around 80 per line and controls) (Table 2).

One of the factors that could lower the chances of finding the *atf4R6* FRT recombinant chromosome is the fact that the *atf4R6* deletion carries the original P insertion. Thus, recombination at the P element from *atf4R6* could be more efficient than recombination at the P element from the FRT insertion. We should also consider the chromosomal environment where this gene is located. *atf4* is located by the centric heterochromatin which is has a tightly packed chromatin structure [65]. This packed chromatin might make the distal P element (by the FRT element) less accessible to the transposase enzyme.

To overcome the low efficiency of the *atf4R6* FRT recombination I scaled up the same experiment described above, and also attempted to recombine the crc^{1} mutation with a FRT into the same chromosome. Since both recombination attempts lead to different outcomes I will first describe the *atf4R6* FRT recombination results.

To recombine *atf4R6* with the 40A FRT element I induced the recombination described above in Fig. 8. I analyzed 5,309 neomycin resistant flies (which should carry the FRT element). Among them only 153 flies presented the genetic characteristics supplied by the desired recombinant chromosome, loss of the *Sco* marker and presence of the *atf4R6* marked by the *w+* gene (in addition to neo^{R}). Of the 153 recombinant candidates only 143 were fertile. I molecularly tested for the presence of the FRT element in the 143 fly lines by using single fly PCR described above (Fig. 9). Results from the PCR indicated that only eight lines had the FRT element (they had the 5'and 3' PCR bands described in Fig. 9). This narrowed down the positive recombinants from 143 to only eight lines. To re-test for the presence of the *atf4R6* deletion genetic complementation was assayed by crossing to the original *atf4R6* and crc^{1}. This time none of the *atf4R6* FRT lines complemented crc^{1}or *atf4R6*. In addition I molecularly tested for the *atf4R6* deletion by assessing the absence of ATF4's b-ZIP domain. As I explained above a PCR that detects the b-ZIP region (deleted in *atf4R6*) should be negative in ¼ of the embryos (homozygous *atf4R6*) from the recombinant lines. The PCR showed absence of the b-ZIP domain in a quarter of the embryos as expected. As a conclusion I had obtained eight potential *atf4* FRT recombinant lines based on genetics studies and PCR (Table 2).

Table 2: *atf4R6* FRT recombination test and results summary

		Attempt 1	Attempt 2
To detect	**Test applied**	**Number of flies or fly lines positive for the test**	**Number of flies or fly lines positive for the test**
Presence of FRT	neo^R	519	5,309
Loss of Sco	Wild type scutelar bristles	27	153
	Fertility	27	143
Presence of FRT	3' and 5' PCR of 40A FRT insertion	4	8
Presence of *atf4R6*	Cross to crc^1 and look for lack of complementation	0	8
Presence of *atf4R6*	Single embryo PCR to detect lack of b-ZIP domain	0	8

Once the *atf4R6* FRT recombinants were made, they were used to eliminate *atf4* from the maternal contribution. This was accomplished by using the dominant female-sterile technique (Fig. 7). In this technique, the FLP enzyme catalyzes a recombination reaction between the FRT elements from homologous chromosomes. This recombination allows some of the cells to be *atf4R6* FRT homozygous (Fig. 7). Because the *OvoD1* mutation is a dominant female sterile mutation, most of the *OvoD1* homozygous cells, as well as the heterozygous *OvoD1* cells, cannot produce progenitor cells. Consequently, most of the germ-line is populated by cells that are homozygous for the deletion. To perform the dominant female-sterile technique, I introduced a chromosome that carried the gene encoding for the FLP enzyme into the *atf4R6* FRT/CyO stock. Next I brought the *OvoD1* allele by crossing *OvoD1* males to hsFLP/hsFLP; *atf4R6* FRT/CyO females

flies (Fig. 7). Larva from this cross was heat shocked to activate FLP expression at the development stage when the germ-line formed. After the heat shock, I collected hsFLP/+; *atf4R6* FRT/*OvoD1* females flies that emerged, and crossed them to wild type males to obtain the germ-line clone embryos. As a control, I collected the CyO siblings and crossed them to wild type males (Fig. 7). Among the CyO flies, hsFLP/+; *atf4R6* FRT/CyO flies should be fully viable. All the *atf4R6* FRT recombinants were examined by this procedure. From the eight lines tested, two of them did not produce either clones nor CyO progeny. This could be an indication that the line carrying the recombinant chromosome has some extra sensitivity to heat shock. Another line produced almost no clones but their CyO siblings were fully viable. The remaining 5 lines behave as expected; the CyO flies were fully viable but fewer than expected germ-line clone embryos were laid. Female and male progeny of non-virgin germ-line hsFLP/+; *atf4R6* FRT/*OvoD1* female were counted. The results are described below in Table 3

Female flies accounted for nearly 50% of the progeny, indicating that there is not a decrease in female viability in the clones. This survivor rate indicated that maternally deposited *atf4* is not necessary for sex determination or that *SxlPe* expression has not changed enough to decrease female viability.

Table 3: Maternally deposited ATF4 is not needed for female survival

atf4R6 FRT line number	Percentage of Females flies (#)	Percentage of Males flies (#)
59	75 (24)	25 (8)
4	43 (19)	57 (25)
151	51 (42)	49 (40)
95	51 (31)	49(30)
150	49 (67)	51(69)

Progeny of no-virgin heat shocked hsFLP/+; *atf4R6* FRT/*OvoD1* females. 100% is considered as the total amount of flies that emerged.

Analysis of the *crc*^{1} mutation

As a complementary experiment and back up, I recombined the hypomorphic *crc*^{1} mutation with the 40A FRT insertion (following the basic outline showed in Fig. 8). In this experiment I analyzed 11,227 neomycin resistant flies. Among these flies, only 198 looked like recombinants (they lost the dominant bristle marker Sco). I then tested for the presence of the FRT element in the recombinant flies using the single fly PCR described above to amplify the 3' and the 5' of the 40A FRT insertion (Fig. 9). The PCR results narrow down the positive lines from 198 to only 5 lines that presented a positive PCR reaction for both sides of the 40A FRT insertion. To confirm the presence of the *crc*^{1} mutation on the recombinant chromosome I designed primers that could distinguish the Wild type *crc* gene from the *crc*^{1}mutant allele. The *crc*^{1} mutant has an A to G replacement in the wild type gene. I made a primer that complemented the G from the mutant at the 3' end of the primer. After a limited number of amplifications I could

clearly differentiate between the wild type and the *crc*^{1} mutant chromosome. Only one of the five recombinants presented a positive *crc*^{1} PCR band. I tested for complementation by crossing the five *crc*^{1} FRT recombinant lines to the original *crc*^{1}mutation. The same recombinant line that was positive for the *crc*^{1} mutation (on the PCR assay) was the only line that did not complement the original *crc*^{1} mutant. Surprisingly, the *crc*^{1}FRT/*crc*^{1}survival rate of this cross was 19% instead of the 2% expected [60]. This result could indicate the presence of an unknown suppressor in the recombinant's background. To corroborate the presence of *crc*^{1} in the recombinant stock, I tested for *atf4R6* complementation and found a 0% *crc*^{1}FRT/ *atf4R6* survival rate as expected [60]. In spite of the increased *crc*^{1}FRT/ *crc*^{1}survival, I used this line to carry out the germ-line experiment.

The *crc*^{1} FRT recombinants were mated to flies carrying the gene that encodes for the FLP enzyme under the regulation of a heat shock promoter. As a control I crossed the resultant hsFLP/hsFLP; *crc*^{1}FRT/CyO flies to the original *crc*^{1} mutant line recovering 19% *crc*^{1}FRT/ *crc*^{1} survival. At the same time hsFLP/hsFLP; *crc*^{1}FRT/CyO female flies were made to *OvoD1*/CyO males. The larva from this cross was heat shocked and hsFLP/+; *crc*^{1}FRT/*OvoD1* female flies, as well as their CyO sisters, were collected. The females were crossed to wild type males to analyze their germ-line. Progeny emerged only from the CyO control cross. An occasional egg was laid by the hsFLP/+; *crc*^{1}FRT/*OvoD1* females but not enough to collect and analyze.

The 19% *crc*^{1}FRT/ *crc*^{1} survival were obtained as a progeny of the cross described above provided another means of analyzing crc1 for an effect on sex

determination. Because they are homozygous *crc*^{1} mutants, they lack zygotic *atf4,* and their progeny are predicted to be deficient in maternally deposited *atf4*. I counted female and male progeny from non-virgin *crc*^{1}FRT/ *crc*^{1} females. 56% of the progeny were females (50% of the flies should be females in a wild type stock) indicating that *atf4* is not involved in sex determination (Table 4).

Table 4: Germ-line clones from the *crc*^{1} mutant do not show decrease of female survival

Percentage of Females flies(#)	Percentage of Male flies(#)
56 (185)	44 (147)

Overall the results *crc*^{1} strongly suggest that maternally contributed *atf4* is not necessary for female survival, indicating that *atf4* is probably not involved in sex determination. It remains a formal possibility that the *crc*^{1} hypomorph is not sufficiently defective in maternal function to directly affect the viability of flies, but could still affect the *SxlPe* switch.

In summary, neither *atf4R6* nor *crc*^{1} seem to show any indication that maternally contributed *atf4* is involved in sex determination. Flies emerging from the apparent germ-line clones of the null mutant looked completely normal and viable. These findings indicate that atf-4 is not the exclusive SisA dimerization partner; however it remains possible that it is one of several redundant dimerization partners.

CG16815 and CG16 813

CG16815 and *CG16813* are two annotated genes with sequences similar to the bZIP family of transcription factors [60]. Because these genes are located near each other in the chromosome I chose to delete both of them together for my analysis.

CG16815 and *CG16813* are maternally deposited in the egg as described above (Fig. 6). Therefore, to assess their genetic interaction with *sisA* and their effect on *SxlPe* expression I needed to generate germ-line clones. As explained earlier, the technique requires a FRT element on the same chromosome arm as the mutations in the genes of interest (Fig. 7). The only available deletion that included *CG16813* or *CG16815* also contains 26 other genes [61], which suggests that any germ-line clones carrying it would likely be cell lethal and thus useless for analysis. To avoid this problem, I used the ends-out technique described below to generate a single deletion that eliminates both *CG16815* and *CG16813* genes but no other loci (Fig. 10) [66,67]

The ends-out technique requires creating a mutated targeting transgene (the donor) that can be excised from the genome as a linear fragment. This linear extra-chromosomal DNA induces site-specific recombination at the endogenous, targeted locus, because the linear DNA induces homologous double-strand break repair. [66,67]. Recombination between both ends of the linear fragment and chromosome resulted in the replacement of the endogenous loci with the donor fragment.

As outlined in Fig. 10, I created a donor construct in which the *CG16815* and *CG16813* coding regions were deleted and replaced with the *w+* marker. The donor was then introduced on the X chromosome by P-element mediated transformation. To excise

the donor fragment, both FLP recombinase and the I-Sci site-specific endonuclease were induced via heat shock. The FLP recombinase results in excision of a circular donor molecule that is then linearized by the I-SceI enzyme to create the recombinogenic extra-chromosomal DNA fragment. The fragment is guided to the endogenous locations of the *CG16815* and *CG16813* genes because it carries 5.2 kb of DNA upstream and 5.6 kb of DNA downstream, of the deleted segment.

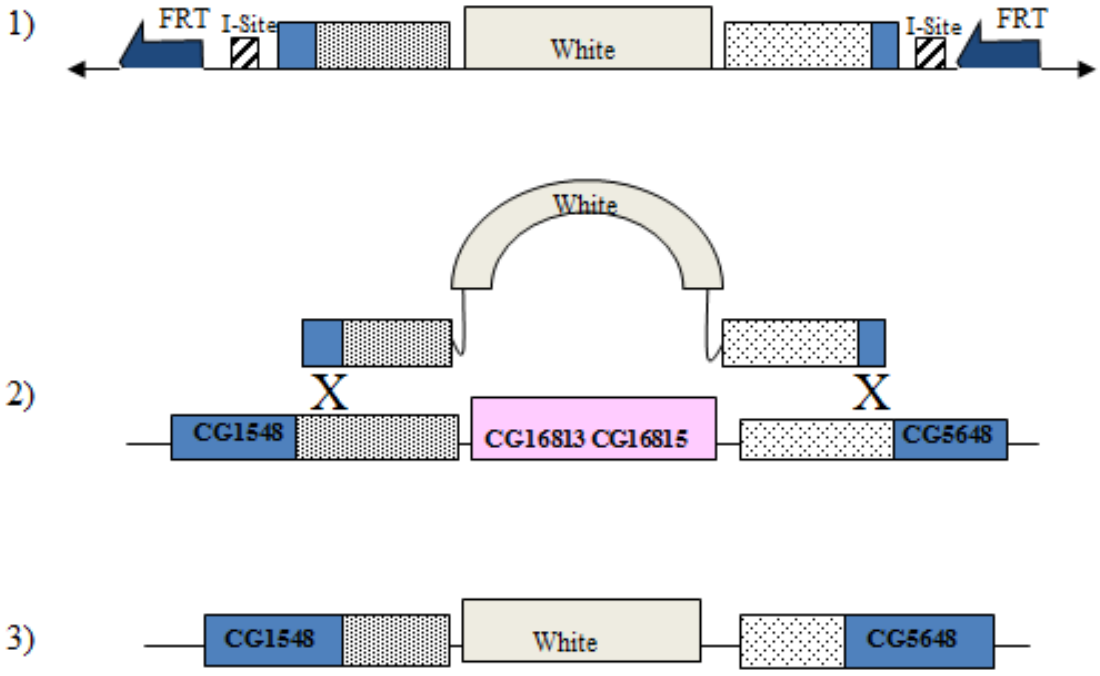


Fig. 10: Ends out gene conversion. 1) Gene conversion construct inserted in the chromosome. Note the FRT sites in the same direction will produce a circular extra-chromosomal excision when exposed to flip recombinase. I-sites will be recognized and cut with the I-SceI enzyme generating recombinogenic double stranded breaks. 2) Homologous recombination will occur at desire site. 3) Gene replacement has taken place introducing the white gene in the desired chromosomal location.

I first attempted to isolate deletions from crosses between 350 female flies that carried the donor insertion to males carrying the *FLP*- I-*SceI* genes on the third chromosome. This resulted in 15 candidate replacement flies that could be identified by their having uniform red eye color (*w*+) in the presence of the FLP recombinase. Unfortunately I found that 11 of the lines retained a modified donor fragment on the X chromosome and four lines had mobilized the donor to the 3rd chromosome by a non-homologous method.

I therefore scaled up the experiment by performing 1150 crosses between donor females and males carrying the *FLP*- I-*SceI* genes. I recovered 36 independent lines with solid red eye color, and 24 of these retained their solid eye color when exposed to the FLP enzyme. Of the 24 lines, 20 still had an insertion that mapped to the X chromosome, 3 mapped to the third, but the final line was found to be at second chromosome and was thus a candidate for a deletion of the *CG16815* and *CG16813* loci.

To confirm that the single line had the proper deletion, PCR was done using primers within the white locus and in the flanking DNA, just beyond what was included in the donor transgene. A positive signal from both upstream and downstream primers was obtained, indicating a correctly located insertion since neither the endogenous locus nor the donor fragment could be amplified by these primers (Fig. 11). Because the *CG16815* and *CG16813* double deletion proved homozygous viable, I confirmed that the double mutant lacked any PCR amplifiable material from the coding segments of either *CG16815* or *CG16813.* In controls the same primers amplified DNA from both wild type controls and deletion heterozygote

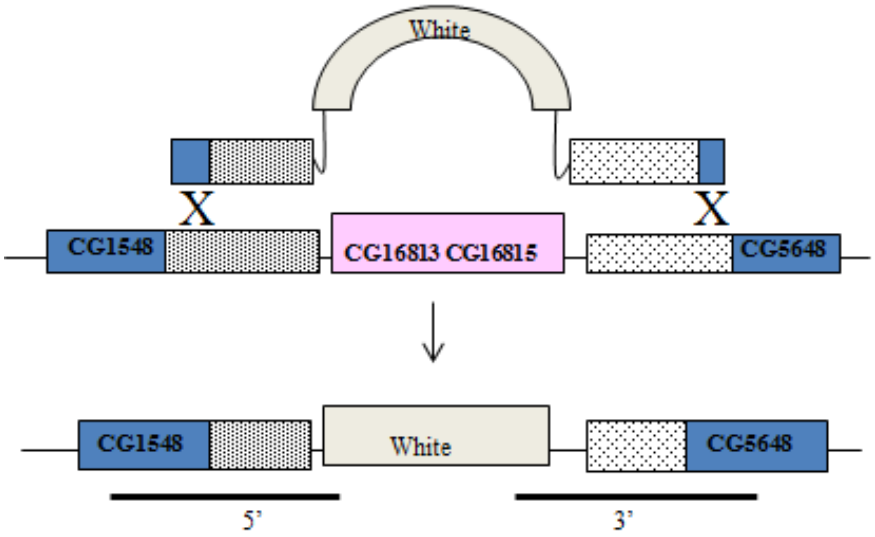


Fig. 11: Confirmation of the gene replacement. After gene replacement has taken place the location of the replacement can be corroborated by PCR through amplification of the 5' and 3' fragments.

Genetic analysis of Δ(*CG16815*, *CG16813*)

Our expectation was that flies defective in SisA's sex determination partner should exhibit a phenotype complementary to that of *sisA*. Specifically, if the partner is exclusively maternal, Δ(*CG16815*, *CG16813*) mothers should produce few, if any, daughters, and if zygotic (or both maternal and zygotic), then the female Δ(*CG16815*, *CG16813*) of Δ(*CG16815*, *CG16813*) mothers should also be unviable. My finding that the homozygous Δ(*CG16815*, *CG16813*) flies were viable and fertile and exhibited no sex bias among their progeny over multiple generations strongly suggested that neither of these two genes is SisA's sole partner in sex determination. Similar logic applies to *sisA*'s later function in the yolk nuclei. If either or both, *CG16815*, *CG16813* encoded SisA's essential partner(s) for yolk function, we would have expected complete

embryonic lethality for Δ(*CG16815*, *CG16813*) progeny of Δ(*CG16815*, *CG16813*) mothers.

While neither (CG16815, CG16813) can be the exclusive dimerization partner of SisA, it is possible that SisA might have multiple and redundant dimerization partners. If that were the case, one would predict that loss of one set of partners, i.e., (CG16815 and CG16813) would sensitize the flies to reductions in either SisA or the other partner proteins. To test this idea, I asked if decreasing the *sisA* gene dose by half in the homozygous Δ(*CG16815*, *CG16813*) knockout decreases female viability. I assessed the zygotic contribution of *CG16815* and *CG16813* by crossing *sisA1*/FM7; Δ(*CG16815*, *CG16813*)/CyO females to FM7/Y; Δ (*CG16815*, *CG16813*)/CyO males and counting the progeny. The results are shown in the Table 5. Viability of the $sisA^{1}$/FM7; Δ (*CG16815*, *CG16813*)/Δ (*CG16815*, *CG16813*) females was 95% when compared to brothers of the same genotype. This indicates that zygotic *CG16815* and *CG16813* do not interact with *sisA* to generate female-specific lethality and consequently are not involved in sex determination. To assess any possible maternal contribution of *CG16815* and *CG16813*, I carried out a series of crosses in which the mothers and fathers were completely deficient for *CG16815 and CG16813*, and where the experimental female offspring were heterozygous for *sisA1*. (Table. 6 a, b) In neither series of crosses was there any evidence for female-lethality, as *sisA1*/+; Δ (*CG16815*, *CG16813*) females and *sisA1*/Y; Δ(*CG16815*, *CG16813*) males were equally viable. Taken together, the genetic analysis of Δ(*CG16815*, *CG16813*) demonstrate that neither individual gene, nor both together, can function as the exclusive dimerization partner.

Table 5: Zygotic *CG16815* and *CG16813* do not have a genetic interaction with *sisA*

	Female viability(#)	**Male viability (reference #)**	
X genotype	*sisA*^{1}/FM7	*sisA*^{1}/Y	FM7/Y
2nd chromosome genotype Δ(*CG16815*,*CG16813*)/CyO	95(80)	100(84)	21(18)
Δ (*CG16815*, *CG16813*)/ Δ (*CG16815*, *CG16813*)	95(42)	100(44)	14(6)

FM7/FM7 females are death as most of the FM7 males.

Table 6: Maternally deposited *CG16815* and *CG16813* does not interact with *sisA*

a) *sisA*^{1}/FM7 ; Δ (*CG16815*, *CG16813*)/ Δ (*CG16815*, *CG16813*) X FM7/Y ; Δ (*CG16815*, *CG16813*)/CyO

	Female viability(#)	**Male viability (reference #)**	
X genotype	*sisA*^{1}/FM7	*sisA*^{1}/Y	FM7/Y
2nd chromosome genotype Δ(*CG16815*,*CG16813*)/CyO	109(38)	100(35)	34(12)
Δ (*CG16815*, *CG16813*)/ Δ (*CG16815*, *CG16813*)	105(42)	100(40)	12(5)

FM7/FM7 females are death as most of the FM7 males.

b) *sisA*^{1}/FM7 ; Δ (*CG16815*, *CG16813*)/ CyO X FM7/Y ; Δ (*CG16815*, *CG16813*)/ Δ (*CG16815*, *CG16813*)

	Female viability(#)	**Male viability (reference #)**	
X genotype	*sisA*^{1}/FM7	*sisA*^{1}/Y	FM7/Y
2nd chromosome genotype Δ(*CG16815*,*CG16813*)/CyO	117(33)	100(28)	36(10)
Δ (*CG16815*, *CG16813*)/ Δ (*CG16815*, *CG16813*)	97(30)	100(31)	0(0)

FM7/FM7 females are death as most of the FM7 males.

Unfortunately there were no indications of an interaction between any of the candidate genes obtained from the two hybrid experiment and *sisA* in sex determination or in its vital yolk function. It remains formally possible that redundancy between all three candidate loci is impairing my ability to distinguish the genetic interaction

Alternative Approaches to identify SisA's partner

Co-transformation and co-immune precipitation of SisA and its partner

In *Drosophila melanogaster*, 27 annotated genes encode b-ZIP proteins. The experiment that I will describe next was designed as a quick screen to detect among all known drosophila b-ZIP proteins the proteins that physically interact with SisA. The plan was to co-express tagged forms of SisA and tagged forms of each one of the known b-ZIP proteins in *Drosophila* S2 cells, and determine whether the proteins interacted using co-immune precipitation (co-IP). Using this approach, I was successful in confirming the 2-hyrid interaction between SisA and CG16815, but technical difficulties

with protein expression and protein detection slowed the analysis, and ultimately forced a choice between carrying out the genetic analysis of the 2-hybrid interactors and continuing the screen in cell culture. Because the cell culture approach was logically similar to the completed two-hybrid screen, and because it was subject to artifacts due to its reliance on identifying interactions among overexpressed proteins, I terminated the cell culture project to focus on the genetic analyses. For completeness, I include a brief summary of my experiments and suggestions for anyone wishing to pursue this experimental direction.

Cell culture assay method

All the proteins were cloned in pMT/V5 His B vector from Invitrogen. This vector possesses a metallothionein promoter that can be induced with copper sulfate or cadmium chloride added to the media. It also provided a multicloning site that allowed me to clone each partner protein in frame with a C-terminal V5 tag fallowed by a 6XHis tag. To express *sisA* it was cloned in frame with an N and C-terminal 3X FLAG tag. To assure that *sisA* expression was induced as much as its partner proteins, *sisA* constructs were introduced into the same inducible pMT/V5 His B vector that was used for the partner proteins. Both plasmids (the one containing the partner protein and the one containing *sisA*) were transformed to S2 cells. Expression of *sisA* and its partner was induced with copper sulfate added 24 hours after the cells were seeded. Cells were next harvested 24, 48 and 72 hours post induction. To assess the induction, total amount of protein was measured using the Bradford assay. Detection of SisA and its partner protein

was accomplished using either FLAG, or Anti-V5 antibody respectively on a Western blot. To perform the co-IP, I used Anti-V5 agarose affinity gel to pull down the partner protein followed by a Western blot using FLAG antibody to detect SisA or Jra in case of the control (see below). As a positive control for the co-IP, I used two proteins that were known to interact: Jra and Kay, the fly homologues of human FOS and JUN. The interaction of these two proteins has been well studied in flies as well as in humans [68-70].

Although there are 27 previously described b-ZIP proteins in *Drosophila melanogaster* only a few could qualify as SisA partner. I narrow down the proteins to be analyzed based on the following criteria: 1) The SisA partner protein must be a maternally deposited b-ZIP protein, or express early in development between cycle 8 and before cycle 12 (the first burst of *SxlPe* expression is seen at cycle 12); 2) The SisA partner might also have a non-canonical b-ZIP domain. Since Atf4, CG16815 and CG16813 fulfill the characteristics described above, and at the time the genetic experiments described in the previous section had not yet been concluded, I decided to start the screening with these proteins. I also chose CG14014 because it shares the same time and pattern of expression with *sisA* (fly Base). In addition proteins that share SisA's non-canonical b-ZIP domain Jra and CG15479 were selected [58].

In the first experiment I co-transformed and co-expressed *CG16815*, *CG15479* *CG14014* and *atf4* b-ZIP domain (*atf4* is a large gene so I decided to study the interaction with its b-ZIP domain first) with *sisA* in S2 cells as described above. As controls, the complete *jra* sequence and *kay* b-ZIP domain were also co-transformed and

co-expressed. Expression of *CG15479*, *CG16815*, *sisA*, *jra* and *kay* were found to be low when cells were harvested at the standard 24 hours after the induction. CG14014 and atf4 b-ZIP domain did not show detectable levels of expression. Even though the amount of protein expressed was low, I used it to study protein interaction. The result of the positive control experiment indicated interaction between Jra and Kay as expected. I could also detect a positive interaction between SiasA and CG16815. Although the interactions were clear, the bands in the blot were washed-out due to the low initial expression of the proteins. In addition, I was using a low sensitivity colorimetric reaction to detect the horseradish peroxidase enzyme that was bound to the 3XFlag. As a consequence of these complications I could not photograph or scan the interaction (the signal to noise ratio was too low). To overcome this complication, I attempted to increase the amount of protein produced. This could be accomplished by increased the protein expression or increasing the amount of transformed/harvested cells. Because some of the proteins had not been expressed in the first experiment, I decided to work the induction conditions to obtain higher expression levels as well as expression of all the transformed proteins. I transformed all the chosen proteins individually and increased the induction time from both, 24 to 48 hours, and from 24 to 72 hours. I evaluated induction level in each sample collected. I also changed the detection method from the less sensitive colorimetric to a more sensitive chemiluminescence reaction. As a result I detected expression of CG16815 at both, 24 and 48 hours after induction; however, the 48 hours after induction sample showed a higher amount of protein. CG14014 was only expressed 72 hours post induction, and the expression level was low.

CG15479 was only detected in the sample taken 48 hours post induction. CG16813 was expressed best 48 hours after induction. Atf4 did not show any detectable expression in any of the samples. *SisA* was equally expressed 24 and 48 hours post induction. Finally, the control protein Jra and Kay were expressed 24 and 48 hours after induction but the expression was higher in the 48 hours post induction samples.

Future experiments: Purification and identification of Sis A partner from embryonic extract

At the time I began my experiments, I considered attempting to identify the SisA partner by affinity purifying it from embryonic extracts and identifying the proteins by mass spectrometry. Ultimately, however, I did not actively pursue this strategy because the genetic analysis of the two-hybrid candidates proved so timeconsuming. Because protein purification is the strategy I recommend for any future attempt to find the SisA partner, I will outline my initial steps to develop a purification strategy as a guide to others.

While purification from embryonic extracts is the most straightforward approach to finding the SisA partner, a successful strategy will have to overcome several likely difficulties. First, the SisA/partner complex is likely to be present in small amounts, based on the relatively low expression of *sisA* mRNA, and on the limited developmental period when the gene is expressed. Second, for the mass spec approach to be useful, the number of contaminating proteins in the final preparation should be kept to a minimum. It is important to raise the level of SisA and to employ an effective affinity purification

strategy. My initial plan was to maximize the relative amount of SisA/partner protein in embryo extracts by boosting the levels of SisA by expressing a tagged version of it under the control of the hsp70 promoter. Both approaches aided in the purification of the Knirps repressor from early embryo extracts as described in Struffi et al. [71]. To enable the SisA/partner complex to be purified sufficiently, I designed a triple tag. The modified SisA was to carry a 3X FLAG tag followed by a 6X His tag, and then by a Strep II tag. My hope was that the first two tags will be sufficient for purification, but I wanted also to incorporate the Strep II tag in the event further purification was needed. Due to their high ligand-binding specificity, their ability to concentrate proteins and their capacity to be eluted under native conditions, the double FLAG /His tag system has been demonstrated to be useful in purifying proteins from *Drosophila* embryos and subsequently determining protein identity through mass spectrometry. Previous studies have demonstrated that it is equally effective to affinity purify using the FLAG tag followed by the His tag, so I planned to prepare both versions for injection into flies in the event that one proves more tractable than the other. In brief, the plan is to affinity purify the complex using anti-FLAG M2 Mab agarose, and elute using the 3X FLAG peptide followed by the capture of 6X His tagged protein on the Talon® resin fallowed by Imidazole elution [71,72]. Purified protein complexes will be analyzed by SDS gel electrophoresis and the proteins identified using the Mass spectrometer located at the Biochemistry building at Texas A&M University. If the protein proves insufficiently pure for mass spectrometry analysis, I will employ the additional affinity purification step, utilizing Strep Tactin-Sepharose fallowed by Desthiobiotin elution. Even though

the Strep II appears to be as good as a FLAG in aiding purification of proteins [72], the affinity constant of Strep II tag/ Strep-Tactin is two orders of magnitude lower than FLAG/ anti-FLAG M2 Mab [73], making the Strep II purification more suitable to be used as the last extraction step when the proteins are already present in a small volume.

Chapter Summary

The aim of this chapter was to identify SisA partner protein. The first strategy was to pull out SisA partner protein candidates from a two-hybrid assay. Three genes seemed to fit all the necessary characteristic to be SisA partner, *CG16815*,*CG16813* and *atf4.* Dominant genetic interaction of these genes was first assessed and none were found. Because each gene product is maternally supplied the maternal contribution of these genes. *atf4* maternal contribution was deleted using dominant female-sterile technique. After performing experiments with two different *atf4* alleles I reached the conclusion that *atf4* is not required for sex determination, and it is not the exclusive SisA dimerization partner.

To eliminate all sources of *CG16815*, *CG16813,* I knockout both genes at the same time (they are near each other in the chromosome) using the ends-out technique. Surprisingly, the knockout animals were viable and fertile and showed no decrease in viability of either sex. In addition, decreasing *sisA* gene copy in these flies also did not show any affect in viability demonstrating that neither the individual genes nor both genes together were the sole SisA partner.

Unluckily, none of the genes tested interacted with *sisA*, or showed any involvement in sex determination. This does not eliminate the possibility of redundancy among these genes. To pursue this idea, the maternal contribution of all three candidate genes should be eliminated. This can be done using standard genetic techniques using the strains I have created.

CHAPTER III

SXLPM REGULATION AND *SXLPE -PM* TRANSITION *

While the transient female-specific activation of *SxlPe* in response to X chromosome dose has been the target of much experimental scrutiny, little is known about the control of *SxlPm*. The standard view is that *SxlPm* is a "housekeeping" promoter active in both sexes from around the time of gastrulation through adulthood. Analysis of *Sxl* RNA by Northern blot or RNase protection assays [74-77] suggested a time lag of 1 to 2 hours between the cessation of *SxlPe* activity in early nuclear cycle 14 and the onset of *SxlPm* expression [78] supporting the idea that the two promoters are expressed independently. On the other hand, Barbash and Cline [79] detected *SxlPm*-derived transcripts during cycle 14, and Keyes [80] noted that *SxlPm* appeared to be expressed earlier in XX than in XY embryos raising the possibility of a direct regulatory connection between *SxlPm* and *SxlPe*.

To define when *SxlPm* is active, we developed an *in situ* hybridization assay using an intron-derived probe (Fig. 12) that enabled me to identify nascent *SxlPm*-derived transcripts as focused dots of staining in embryonic nuclei.

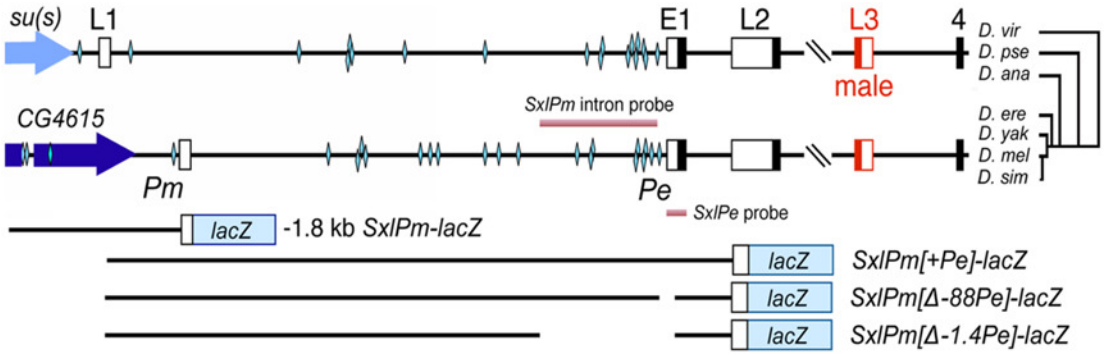


Fig. 12: Map of *Sxl* locus and *SxlPm-lacZ* fusions. (Top) Structure of *Sxl* exons L1, E1, L2, L3, and 4 in *D. pseudoobcura* and *D. melanogaster*. Exons 5-10 are omitted. *SxlPm* and *Pe* promoters are marked. Female splice patterns are E1 to 4 and L1 to L2 to 4. Males splice L1-L2-L3-4. Diamonds represent known or predicted Sc/Da binding sites. Filled portions of *Sxl* exons represent coding regions. A chromosomal rearrangement exchanged the ancestral upstream *su(s)* gene for *CG4615* after divergence of the *D. ananassae* and the *D. erecta*, *D. yakuba*, *D. melanogaster*, and *D. simulans* lineages. *D. virilis* diverged from the other species about 40 million years ago. (Bottom) *SxlPm-lacZ* transgenes. Genomic fragments extended 1.8 kb or 0.5 kb upstream of *Sxl* exon L1. Internal deletions from -88 to +85 or -1,452 to +85 bp relative to exon E1, removed the *SxlPe* promoter and regulatory sequences.

Because *Sxl* is located on the X chromosome, we could differentiate between male (XY) and female (XX) embryos based on the number of dots visible in the nuclei. Progression through cycle 14 was monitored using two different parameters: the ratio between the length and width of the surface nuclei and the extent of cell membrane invagination during the cellularization process [81,82].

Sxl maintenance promoter is activated earlier in females than in males

Inspection of embryos after hybridization with *SxlPm*-specific probes revealed that *SxlPm* is expressed in both sexes from early in the cellularization cycle until the completion of embryonic development (Fig. 13 and data not shown). Analysis of early

embryos, however, showed that the initial expression of *SxlPm* was sexually dimorphic. Transcripts from *SxlPm* first appeared in females during nuclear cycle 13 (Fig. 13, Table 7). Initially, only about 15% of female nuclei expressed *SxlPm*, and many nuclei expressed it from only one of the two X chromosomes, suggesting that activation of *SxlPm* is a stochastic process occurring independently on each X chromosome. During the first minutes of cycle 14, the proportion of expressing nuclei, as well as the number of nuclei expressing both alleles increased, until by 10-15 min every female nucleus expressed both copies of *SxlPm*. In male embryos, expression from *SxlPm* was delayed by about 10 min relative to females (Fig. 13, Table 7). Nascent transcripts from *SxlPm* were first seen in XY embryos as very faint dots in scattered nuclei in early cycle 14. As cycle 14 progressed, the proportion of XY nuclei expressing *SxlPm*, and the staining intensity of the nuclear dots increased, until by 20-30 min every male nucleus transcribed from *SxlPm*. Once *Sxl* was fully active, both sexes maintained expression from *SxlPm* throughout embryogenesis (data not shown).

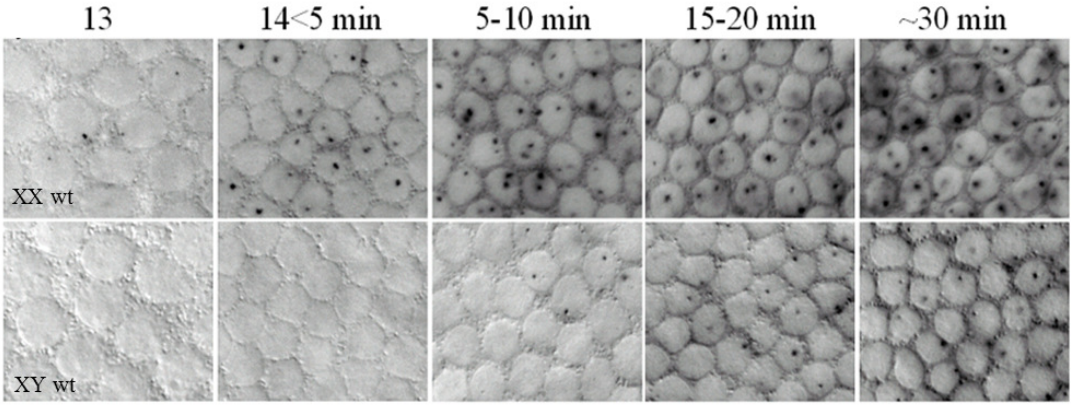


Fig. 13: *SxlPm* activation is sexually dimorphic. Time course of nascent transcripts from *SxlPm* for wild-type, XX and XY embryos. *In situ* hybridization with a probe specific for *SxlPm-derived* pre-mRNA. Surface views of syncytial nuclei.

Table 7: Percentage of expressing nuclei. Percentage of expressing nuclei was determined by counting all nuclei in photographs of embryos. Mean # cycle 13 nuclei counted/embryo (+/- s.e.) = 224 (+/- 26); mean # cycle 14 nuclei counted/embryo = 432 (+/- 38). For samples listed as 0% expressing, number in parentheses indicates photographs examined for evidence of expression. A minimum of 10 additional fixed embryos were examined directly for expressing nuclei for all non-expressing stages. For those listed as 100% expressing, at least one photograph and several other embryos were examined thoroughly for rare non-expressing nuclei.

		Range % nuclei expressing (# embryos counted)				
genotype		cycle 13	cycle 14 < 5 min	cycle 14 5-10 min	cycle 14 10-15 min	cycle 14 15-20 min
Wt	XX	12-16 (4)	48-61 (5)	85-91 (5)	100	100
	XY	0 (10)	6-9 (5)	33-39 (3)	71-85 (3)	100
sc^{sisB3}	XX	0 (2)	12-16 (3)	51-63 (2)	80-90 (2)	100
	XY	0 (3)	0 (2)	7-9 (2)	53 (1)	77-79 (2)
da^{1}	XX	0 (8)	7-13 (3)	50-57 (3)	83 (2)	100
	XY	0 (3)	0 (5)	8-9 (2)	54-60 (2)	76-88 (2)
Dp sc^{+}, *sisA*$^{+}$	XX	71 (1)	89-90 (2)	100	100	100
	XY	27 (1)	74 (1)	81 (1)	100	100
sisA^{1}	XX	11-13 (2)	45-55 (3)	83-88 (2)	100	100
	XY	0 (2)	9-11 (3)	21-35 (2)	82-85 (2)	100
gro^{E48}	XX	11-16 (2)	46-57 (2)	81-85 (2)	100	100
	XY	0 (3)	10 (2)	30-32 (2)	79-85 (2)	100

Our findings demonstrate that *SxlPm* is active earlier than previously estimated from Northern blot and RNase protection analyses [76,77]. They also show that the expression periods of *SxlPe* and *SxlPm* overlap in females during the first 10-20 minutes of cycle 14 (see [41,62,83]). To determine if the sexually dimorphic pattern of *SxlPm* activation is conserved in other *Drosophila* species, we examined *D.virilis*, *D. yakuba*, and *D. simulans* (Fig. 12) using *in situ* hybridization. We found that all three species expressed *SxlPm* similarly to *D. melanogaster* (Fig. 14), suggesting that the female-first pattern of maintenance promoter activation is, like the female-specific activation of *SxlPe* [62], an ancient response to the number of X chromosomes.

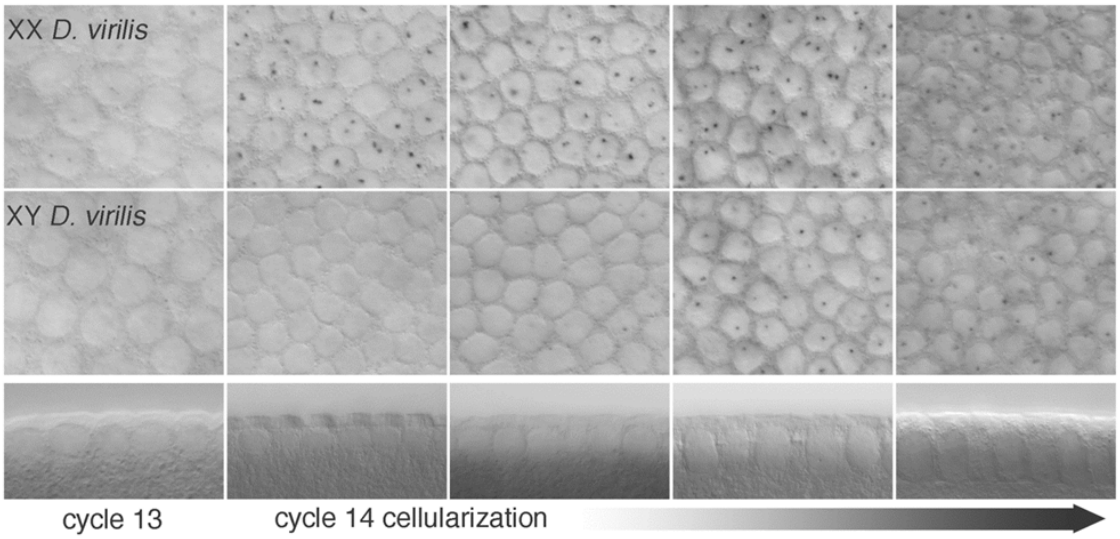


Fig. 14: Time course of nascent *SxlPm* transcripts in *D. virilis* determined by *in situ* hybridizations with a probe specific for *D. virilis SxlPm*-derived pre mRNA.
(Top) Surface views of syncytial nuclei in wild-type XX and XY embryos during cycles 13 and 14. (Bottom) Views of elongating nuclei of the XY embryos pictured above to monitor progression through the cellularization cycle. Nascent transcripts from sister chromatids can be seen in some nuclei.

The XSEs *scute*, *runt* and maternal *daughterless* regulate *SxlPm*

SxlPe is activated during nuclear cycle 12 and expressed through the first 10-20 min of cycle 14 in response to the two X dose of XSEs [41,79,83]. To determine if the same X-linked signal elements (XSEs) that control the on-or-off response of *SxlPe* also regulate *SxlPm*, we analyzed mutations in several XSEs and cofactors to determine if they affected transcription from the maintenance promoter.

The XSE *scute* encodes a dose-sensitive bHLH transcription factor that dimerizes with maternally-supplied daughterless protein to directly activate *SxlPe* [45]. We found that loss of zygotic *scute* (*sc*) or maternal *daughterless* (*da*) also affected *SxlPm*. In sc^{sisB3} and the maternal da^{1} mutant, progeny expression of *SxlPm* was delayed in both sexes by about 5 to 10 minutes compared to wild type (Figs. 15A, B). In XX embryos, no expression was observed during cycle 13 and only a fraction of nuclei showed expression by 5 min in cycle 14. Thereafter, however, the proportion of expressing nuclei increased, until by about 20 min in cycle 14, all female zygotic nuclei expressed *SxlPm* in a manner indistinguishable from wild type. Expression in sc^{sisB3} and da^{1} males was similarly delayed. About one half of XY nuclei expressed *SxlPm* at 15 min into cycle 14 and all had stably activated the maintenance promoter by 30 min into the cellularization process.

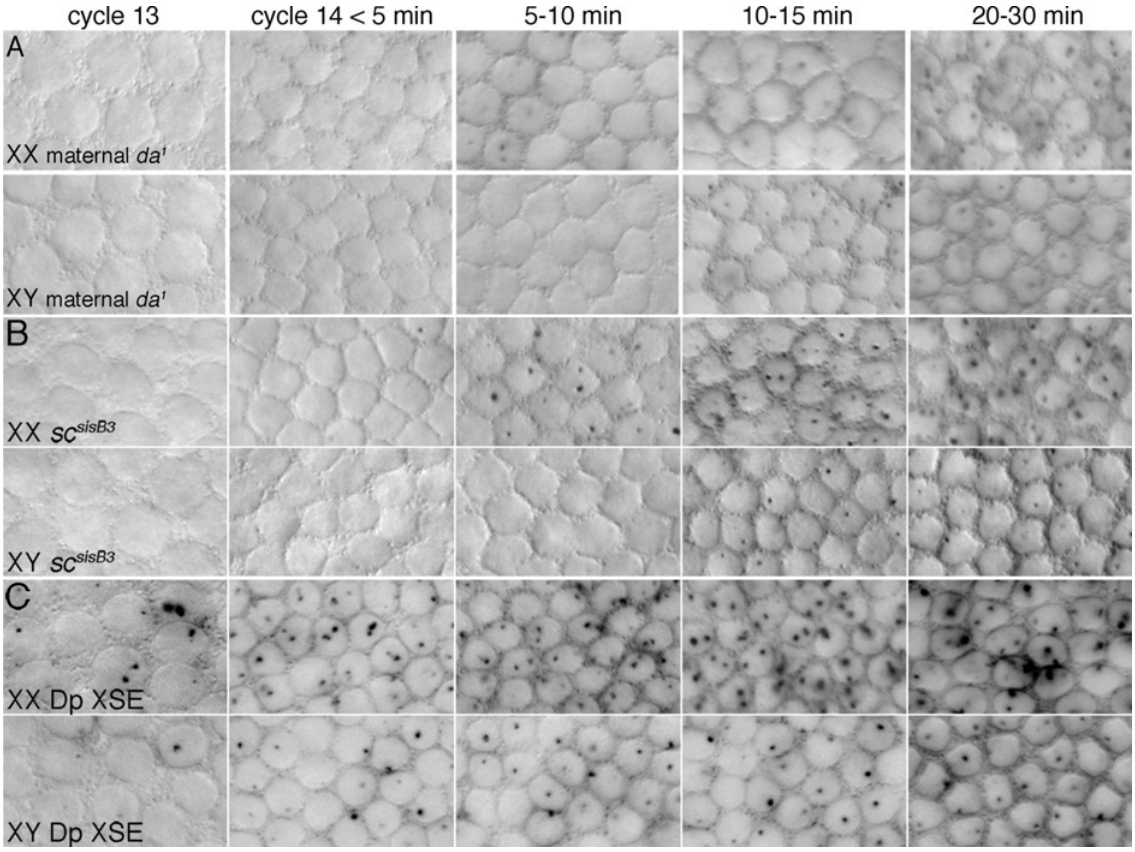


Fig. 15: Time course of nascent *SxlPm* transcripts in maternal or zygotic sex signal mutants. A) Progeny of homozygous *da1* mothers. B) sc^{sisB3} mutant XX and XY embryos. C) XX and XY embryos with two extra copies of the XSEs sc^{+} and $sisA^{+}$. Embryos in B) were progeny of: y sc^{sisB3}/*y* sc^{sisB3} Sxl^{M4} *sn* females and *y* sc^{sisB3}/Y males. Embryos in C) were progeny of: $^{wild\ type}$ females and *y w cm* Sxl^{f1} *ct6*/Y; 2X P(mini-w^{+},$sisA^{+}$)and 2X P(mini-w^{+},sc^{+})/CyO males.

We observed similar results with the XSE *runt* which is required to activate *SxlPe* in the central region of the embryo [51]. In homozygous $\Delta runt^{3}$ females, the number of nuclei expressing *SxlPm*, and the staining intensity of the individual dots, was noticeably decreased between 5 and 30 min in cycle 14 (Fig. 16). This caused a diminution of overall embryo staining intensity in central regions relative to the poles during early cycle 14. Similar effects were seen in $\Delta runt^{3}$ males but the lower contrast

resulting from the lowered expression level of their single *Sxl* allele made it more difficult to document photographically (Fig. 16).

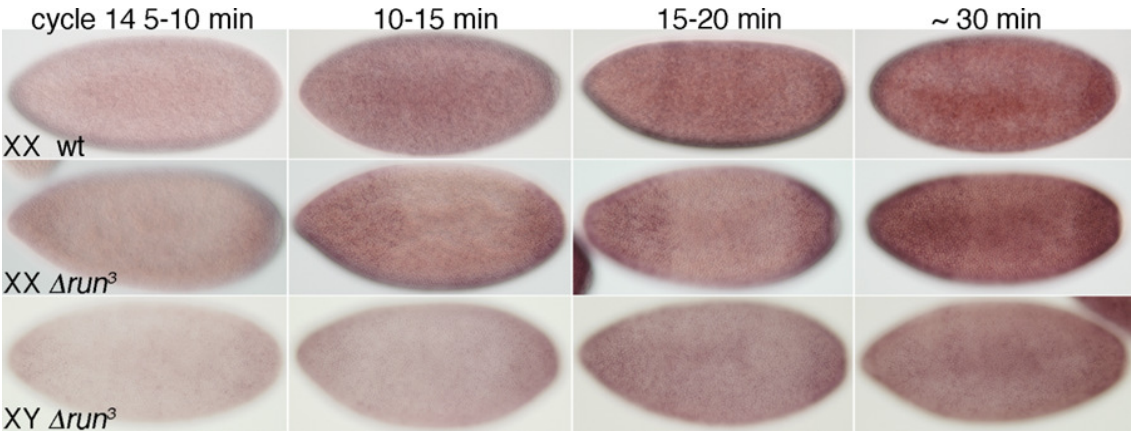


Fig. 16: The XSE *runt* controls expression from *SxlPm*, as determined by *in situ* hybridizations with a *SxlPm*-specific intron probe. (Top) Wild-type females at the indicated times in nuclear cycle 14. (Middle) Homozygous run^3 mutant XX progeny. (Bottom) Hemizygous *run3* XY embryos. Mutant XX embryos were progeny of *w f* run^3/FM7c females and *w f* run^3/$Y y^+$, *mal_* males. One-half of the XX embryos exhibited the abnormal staining pattern. The XY embryos were progeny of *w f* run^3/FM7c females and *FM7c*/Y males. Wild-type females were from *w1118* parents.

Our findings demonstrate that Sc/Da and Runt regulate the onset of *SxlPm* expression in both sexes, but also show that none of these three proteins are required for maintenance promoter activity during the remainder of embryonic development. This implies that at least two temporally distinct mechanisms control *SxlPm* activity: one regulating the sexually dimorphic onset of transcription in response to X chromosome dose, and the second conferring constitutive expression throughout the remainder of the life cycle.

To test the notion that XSE dose specifies the timing of *SxlPm* activation, we asked if an increase in XSE copy number can cause earlier activation of the promoter [40]. To increase XSE dose, we crossed males heterozygous for a second chromosome carrying two transgenic copies each of sc^+ and $sisA^+$ to normal females [33,84]. This created a population consisting of XX embryos with 4 copies of sc^+ and $sisA^+$, XY embryos with 3 copies of sc^+ and $sisA^+$, and normal XX and XY embryos. We found that *SxlPm* was activated earlier in both sexes when XSE dose was increased (Fig. 15 C , Table 7). In females, the extra XSE genes caused about 70% of nuclei to express *SxlPm* during cycle 13 and all XX nuclei expressed *SxlPm* from the earliest stages of cycle 14. In males, the additional XSE copies caused nearly 30% of XY nuclei to express *SxlPm* ectopically during cycle 13 and nearly all expressed *SxlPm* by 10 min into cycle 14.

Our results demonstrate that the sexually dimorphic activation of *SxlPm* is controlled by some of the same determinants that signal female-specific expression of *SxlPe*. They do not, however, allow us to distinguish whether the XSE proteins directly regulate *SxlPm* or whether their effects on *SxlPm* reflect indirect effects, due perhaps to local chromatin changes associated with the activation of the adjacent *SxlPe* (Fig. 12). As a first step to answering the question of whether activation of *SxlPm* is linked in cis to that of *SxlPe*, we asked if we could find conditions where we could uncouple the activities of the two promoters.

SxlPe* activity is not needed for proper regulation of *SxlPm

Duplication of XSE activators leads to strong ectopic expression of *SxlPe* in male embryos. Reciprocally, elimination of the maternal corepressor, Groucho, also causes strong ectopic *Sxl* expression in males by decreasing the threshold XSE concentrations needed to activate *SxlPe* [52]. We reasoned that if the initial female-specific response of *SxlPm* is coupled to the activation of the *SxlPe*, then loss of *groucho* should result in premature expression from *SxlPm*. On the other hand, if Sc/Da and Runt activate the two promoters directly, then the sex-specific response of *SxlPm* may well be independent of *groucho*.

We found that embryos derived from mothers lacking *groucho* germline function expressed *SxlPm* in a wild type pattern (Fig. 17 B). Females first activated *SxlPm* in cycle 13 and expressed the promoter fully by 10-15 min into cycle 14. Males initiated expression early in cycle 14 and fully expressed *SxlPm* some 20-30 min later. Our observations are consistent with direct regulation of the two promoters by the XSEs scute and runt. This conclusion is further supported by our finding that the XSE *sisA* does not regulate *SxlPm*.

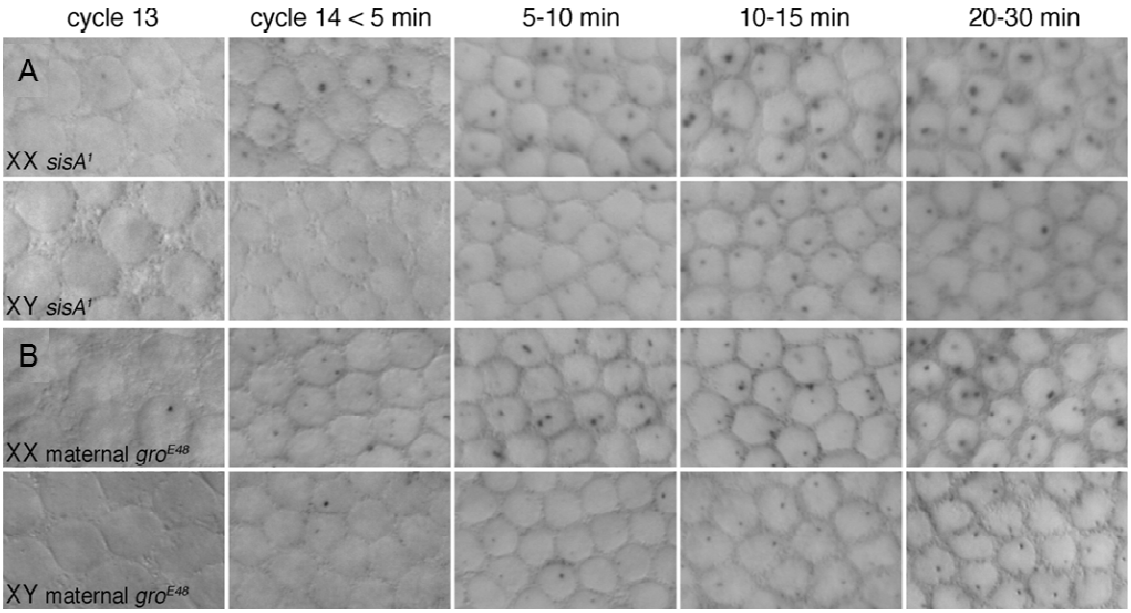


Fig. 17: Time course of nascent *SxlPm* transcripts in maternal or zygotic sex signal mutants. A) $sisA^1$ mutant embryos from crosses of *y pn* $cmSxl^{M4}$ *v sisA1*/*y pn* $cmSxl^{f1}$ *ct6 v sisA1* females and *y pn* $cmSxl^{f1}$ *ct6 v sisA1*/Y males. B) Progeny of mothers carrying gro^{E48} germ-line clones.

The X-linked *sisA* gene encodes a bZIP transcription factor needed for the female-specific activation of *SxlPe* in all somatic cells of the embryo [33,85]. When we analyzed the strong loss-of-function allele $sisA^1$, we found that neither homozygous $sisA^1$ females, nor hemizygous $sisA^1$ males, exhibited any delay in *SxlPm* activation or diminution of staining of *SxlPm*-derived nascent transcripts when compared to wild type (Fig. 17 A). Taken together, our findings with *sisA* and *groucho* show that *SxlPe* activity can be blocked without affecting expression from *SxlPm*, and also that *SxlPe* activity can be induced ectopically without activating *SxlPm*. This strongly suggests that some, but not all, XSEs regulate the two promoters directly, but leaves open the question of

whether the promoters share common enhancers or utilize independent cis regulatory elements.

SxlPm and *SxlPe* share a common regulatory element that responds to X chromosome dose

Our finding that *SxlPm* is activated earlier in XX than in XY embryos in response to Scute, Da, and Runt suggests that *SxlPm*, like *SxlPe*, directly responds to the number of X chromosomes present in the embryo. To determine whether the XSEs and other proteins regulate the two promoters through independent regulatory elements or whether they instead share a common enhancer, we analyzed the structure of *SxlPm* by creating a series of transgenes that fused different portions of the *Sxl* gene to a *lacZ* reporter.

We first assessed the function of the region upstream of the *SxlPm* transcription start site by fusing sequences from -1.8 kb to + 34 bp within exon L1 to *lacZ* (Fig. 12). We found that none of the four reporter lines tested expressed detectable *lacZ* mRNA in embryos, indicating that key regulatory elements needed for *SxlPm* activity are likely located downstream of exon L1 (data not shown). Considering that the sequences upstream of *Sxl* are not conserved in all *Drosophila* species, having been exchanged by a chromosomal rearrangement some 10-15 million years ago, and that the 3' ends of the upstream genes are located within about 200-500 bp of the *SxlPm* start site, we next tested a shorter upstream (-0.8 kb) and large downstream (+ 6.0 kb) segment encompassing the *SxlPe* regulatory elements for *SxlPm* function (Fig. 12). We found

these transgenes expressed *lacZ* mRNA in a manner consistent with them having both normal *SxlPe* and *SxlPm* promoter activity.

To analyze *SxlPm* independent of *SxlPe* activity, we created a modified version of the full-length transgene in which we removed a 171 bp segment that included the *SxlPe* basal promoter and part of the E1 exon (Fig. 12). This construct, *SxlPm*[Δ-88Pe]-lacZ, was expressed in a manner indistinguishable from the endogenous *SxlPm* promoter (Fig 18 A, B, 19 A,B). Weak *lacZ* expression was detected in cycle 13 nuclei in XX embryos and by 10-15 min into cycle 14 every nucleus appeared to express both copies of *SxlPm*[Δ-88Pe]-*lacZ*. Male embryos first expressed *SxlPm*[Δ-88Pe]-*lacZ* in cycle 14 with full activation occurring about 20 min later. Notably, XX embryos expressed *SxlPm*[Δ-88Pe]-*lacZ* mRNA at higher levels than did XY embryos even when the transgenes were present in two copies in both sexes (Fig. 18 A). This difference was maintained through cycle 14 and then gradually disappeared during gastrulation and germ band extension (data not shown). These results establish that all the sequences necessary for normal expression of *SxlPm* lie between -0.8 and +6.0 kb and confirm that a functional *SxlPe* is not required for the early onset of *SxlPm* activity in females.

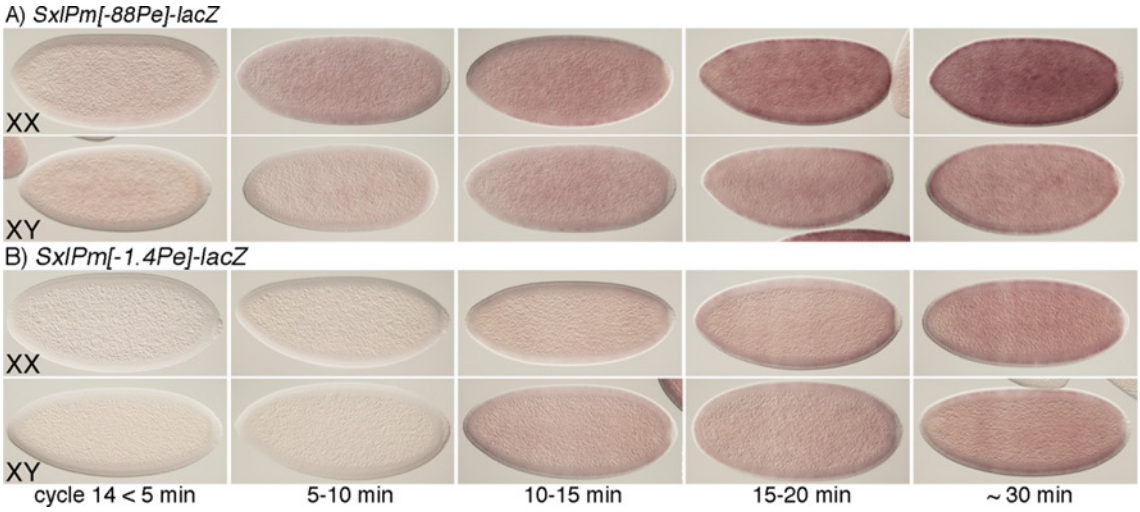


Fig. 18: A 1.4-kb deletion of *SxlPe* regulatory DNA equalizes *SxlPm* activity in the sexes. *In situ* hybridization was used to detect steady-state *lacZ* mRNA levels.
A) Embryos carried two copies of an autosomal *SxlPm [Δ88Pe]-lacZ* transgene deleted for the *SxlPe* basal promoter region. B) Embryos with two copies of an autosomal *SxlPm[Δ 1.4Pe]-lacZ* transgene deleted for *SxlPe* and its regulatory sequences to -1.4 kb. Sex was determined by fluorescent detection of endogenous *SxlPe*-derived transcripts. Times after the onset of cycle 14 are indicated. Four independent lines of each transgene were examined, with indistinguishable results.

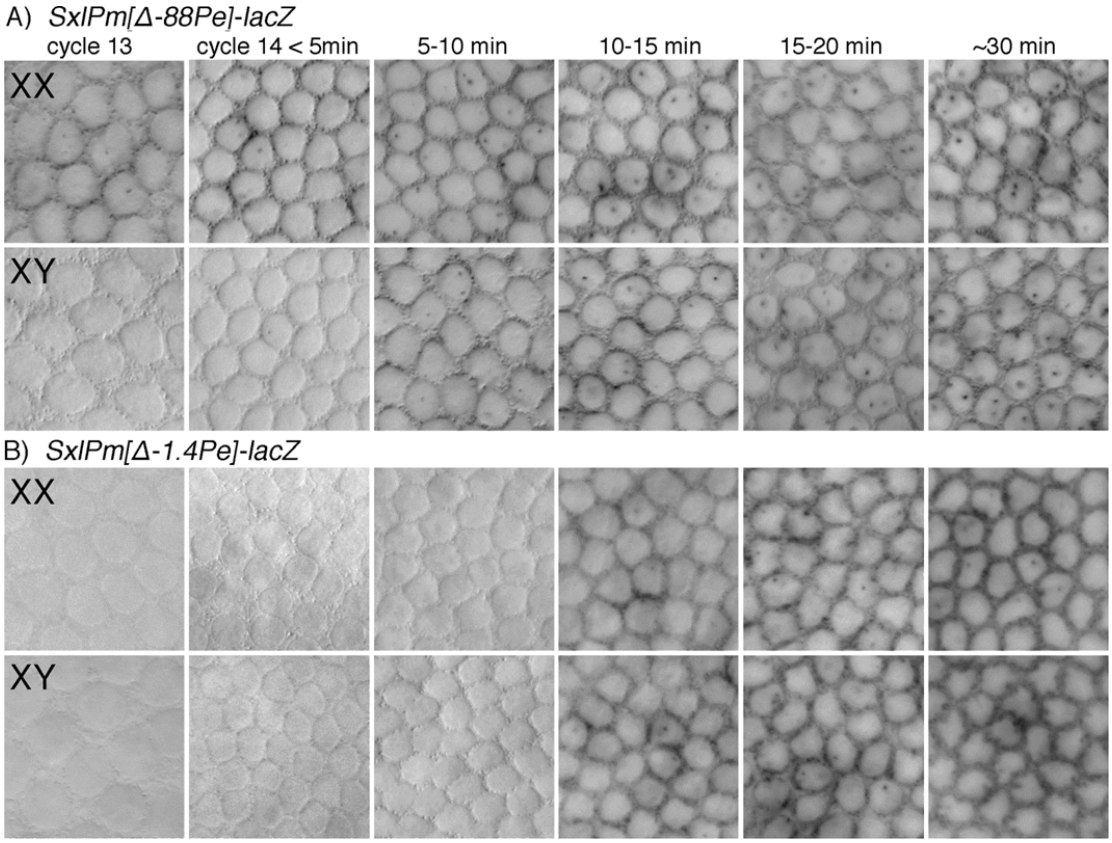


Fig. 19: Time course of nascent transcripts from *SxlPm-lacZ* fusions during cycles 13 and 14 detected by *in situ* hybridization with a *lacZ* probe. Surface views of syncytial nuclei. A) Embryos homozygous for an autosomal *SxlPm [Δ -88Pe]-lacZ* line. B) Embryos homozygous for an autosomal *SxlPm [Δ -1.4Pe]-lacZ* line. Sex was determined by fluorescent detection of endogenous *SxlPe*-derived transcripts. Four independent lines of each transgene were examined, with indistinguishable results. Nascent transcripts from *lacZ* transgenes were more difficult to detect than those from the endogenous *Sxl* locus, resulting in patchier lower-intensity staining.

Normal sex-specific regulation of *SxlPe* requires sequences extending to -1.4 kb upstream of exon E1 [42]. Within these sequences, two regions spanning from +20 to -400 bp and from -800 to -1,400 bp have been identified as crucial for full *SxlPe* activity. To determine if these *SxlPe* regulatory sequences also regulate *SxlPm*, we created a modified *SxlPm*-lacZ fusion carrying an internal deletion spanning the region from -

1452 to +85 relative to the start site of exon E1 (Fig. 12). We found that the onset of expression from the *SxlPm*[Δ-1.4Pe]-lacZ transgenes was delayed relative to the *SxlPm*[Δ-88Pe] lines, and furthermore, that there was no longer any difference in the timing or level of expression between the sexes (Figs. 18 and 19). Expression from *SxlPm*[Δ-1.4Pe]-*lacZ* lines was first seen in a few nuclei in both XX and XY embryos about 10-15 min into cycle 14 (Fig. 19). The number of expressing nuclei increased thereafter, reaching a maximum 15 to 20 min later. Mature transcripts accumulated over time, but XX and XY embryos expressed equal *lacZ* mRNA levels at all times examined. We conclude that the 1.4 kb region that controls the female-specific expression of *SxlPe* is also required for the sex differential expression of *SxlPm*. We also note that the effects of the [Δ-1.4Pe] mutation on *SxlPm* appeared stronger than those of *sc*, *da*, and *runt* mutations on the endogenous locus, as the mutations in the trans-acting regulators did not abolish male/female differences in *SxlPm* expression. This suggests that factors in addition to Sc/Da and Runt interact with the 1.4 kb region to control the initial activation of *SxlPm*.

CHAPTER IV

CONCLUSIONS AND METHODS*

The *Drosophila* sex determination pathway elegantly illustrates the use of pre-mRNA splicing control in development, but the establishment of sex-specific splicing depends, ultimately, on the coordinated activities of two promoters for the master regulatory gene *Sxl*. This dissertation explores *SxlPe* activation by *sisA* and unravels *SxlPm* expression and regulation, resulting in an innovative view of the transition between promoters.

SisA partner protein(s)

In this dissertation I studied the interaction between one of the most quantitatively important XSEs, *sisA*, and the promoter that regulates early *Sxl* expression, *SxlPe*. Although much effort was applied, I was unable to uncover the mechanism by which this member of the b-ZIP family of transcription factors activates *SxlPe*. Because SisA is thought to activate *SxlPe* by binding the promoter as heterodimer, three plausible binding partners from two hybrids experiment were studied: *atf4*, *CG16815* and *CG16813*. Although the genes that code for these proteins seem to

---

*Some material in this chapter is reprinted with permission "A shared enhancer controls a temporal switch between promoters during Drosophila primary sex determination" by Alejandra N. González, Hong Lu, and James W. Erickson., 2008. PNAS, 105, 18436–18441, Copyright 2008 by The National Academy of Sciences of the USA.

follow all the necessary requirements to be SisA partners (see chapter II), none of them were to be the sole SisA partner. It remains formally possible that the system is redundant and a phenotype can only be observed in absence of all three genes. A more likely possibility is that SisA partner is a different b-ZIP protein. In chapter II I proposed to identify SisA partner by extracting the SisA-Partner hetherodimer from an embryonic extract, and identify the proteins through mass spectrometry. Because SisA interacts with proteins of many different molecular natures and functionalities in the two hybrid experiment (for more explanation see chapter II) it is possible that an attempt to affinity purify SisA and its partner from embryo extracts could also result in many false positive interactions. This is an important concern as a lesson from my work is how time consuming it can be to rigorously prove, or disprove, the functions of candidate proteins. One solution to narrowing the focus to authentic interactors would be to consider the molecular nature and predicted 3D structure of SisA itself and look for proteins that could complement its structure. Many DNA binding proteins have intrinsically disordered domains [86]. These domains are regions, or perhaps few amino acids, that are predicted to lack stable intramolecular interaction, and usually form loops that can provide flexibility to the protein structure. These domains can aid in DNA binding or protein-protein interactions by energetically stabilizing the protein-protein, protein-DNA dimers [86,87]. To analyze disorder domains within a protein there are multiple different programs that base their predictions in different structural details (reviewed in ([87]). Intrinsically disordered domains had never been explored within SisA. To determine if SisA had such a domains I used two different programs, DisEMBL (http://dis.embl.de/)

[88] and ANCHOR (http://iupred.enzim.hu/) [89]. The first program search for disordered regions within the coils (sections of the protein that do not follow any clear secondary structure). In addition it examines disorder regions among the hot loops (subgroup of coils that have high mobility) and ultimately it compares the sequence of your protein with previously defined disordered sequences [88]. After running the DisEMBL program on SisA, the program found 2 regions disorder by loops encompassing amino acid from 1-41 and from 73-110. Two regions from 97-151 and from 165-175, were found using the hot loops and two regions 25-43 and 94-121 were found to have homology to previously define disorder region (look at graph and raw data in Apendix A). ANCOR uses IUPred to detect intrinsically disorder regions and then search for protein domains that have the property of stabilizing the protein conformation when bounded to a globular domain, within the disorder region. In other words this programs finds sites on the protein were the amino acid side chain adopt a conformation similar to a cavity in the protein surface. These cavities, although unstable on their own, can be stabilized by interacting with a globular domain in the surface of a binding partner resulting in the stabilization of the dimer structure [89]. After running the ANCHOR program with SisA sequence 5 binding regions (cavities in SisA surface) were found, 6-21, 47-60, 65-78, 89-95, 128-153 (Appendix A for graph and raw data). Though the methods these two programs use to uncover the disorder region are very different and the results from each prediction do not have to match disordered region between 94 and 110 matches the three different predictions. In the future, we could analyze the amino acids contained in each one of the five predicted binding regions and

find possible charged or uncharged regions or regions that could suggest the nature of a binding partner. After that, the structure of partner proteins could be analyzed and searched for domains that carry amino acid side chains complementary to those found in SisA [89]

***SxlPm* regulation and *SxlPe-Pm* transition**

In this dissertation I showed that the switch from the initial assessment of X chromosome dose at *SxlPe* to the stable autoregulatory control of *Sxl* pre-mRNA splicing, exploits an unexpected level of transcriptional control of the *Sxl* maintenance promoter. I demonstrated that, contrary to the prevailing view, *SxlPm* responds to X chromosome dose and that it does so by sharing common X-signal elements and a common enhancer with *SxlPe*. The switch between *Sxl* promoters thus serves as a tractable model to explore the logical circuitry and molecular mechanisms that control the fidelity of developmental switches and that coordinate the uses of multiple promoters for a single gene [90].

Why is *SxlPm* regulated?

A priori, a female embryo must do two things to establish and then remember its sex. It must produce a pulse of SXL protein by transiently activating *SxlPe* in response to the XX signal, and it must activate *SxlPm,* so that its transcripts can be spliced to produce yet more SXL protein. A male embryo need only keep *SxlPe* off so that no SXL protein is present when *SxlPm* is active. The system would seem to impose no

requirement for sexually dimorphic expression from *SxlPm*, nor even for a temporal overlap in transcription from the two promoters [78], yet both features are conserved across the breadth of *Drosophila* species. We believe the resolution to this paradox lies in the recognition that the transition to stable autoregulatory *Sxl* splicing requires that substantial amounts of SXL protein be present [74,78] rather than being driven by trace quantities of SXL protein (see [91,92]. Given this, we consider that overlapping expression from the two promoters ensures that XX cells rapidly engage autoregulatory *Sxl* splicing, while the delayed activation of *SxlPm* in XY cells buffers against improper *Sxl* activation due to random variations in regulatory protein concentrations. In effect, we believe that robustness is conferred on the system by rapid reinforcement of correct decisions. In XX embryos, strong induction of *SxlPe,* coupled with early activation of *SxlPm,* ensures that high levels of SXL protein and its pre-mRNA substrate are present during the transition to splicing control. In XY embryos, chance fluctuations in XSE or inhibitor concentrations that caused low-level activation of *SxlPe,* would not persist to activate *SxlPm,* thus preventing rare mistakes from being amplified into the fully on state. We seeing that a logically similar, two-target, control process operates in *C. elegans* primary sex determination. There, four XSE proteins exert primary control of the master regulator *xol-1* at the level of transcription, while a fifth XSE acts post-transcriptionally to ensure the fidelity of X chromosome counting [93,94]. Inclusion of multiple regulatory steps may prove a general mechanism for conferring robustness on dose-sensitive regulatory switches.

How is *SxlPm* regulated?

SxlPm appears to be equally active in both sexes after the onset of gastrulation. Prior to that, *SxlPm* is expressed in a graded fashion, becoming active earlier and being expressed more strongly in XX than in XY embryos. Sequences governing the early sexually dimorphic expression of *SxlPm* are included in the same 1.4 kb DNA segment that controls the on-or-off regulation of *SxlPe* [95]. Importantly, the 1.4kb region must work as an enhancer for *SxlPm*, rather than exerting an indirect effect *in cis*, via activation of *SxlPe*, because deletion of the *SxlPe* core promoter had no effect on *SxlPm* activity. This effect, combined with the involvement of the XSEs *scute* and *runt* in *SxlPm* regulation, suggests that *SxlPm*, like *SxlPe*, responds directly to the number of X chromosomes present in the embryo. However, the fact that neither loss of the strong XSE *sisA*, nor loss of the potent maternal co-repressor *gro* affects *SxlPm,* argues that the mechanism of X-counting at *SxlPm* differs from that at *SxlPe,* despite their sharing common *cis*- and *trans*-acting components. We suspect that additional transcription factors contribute to both early *SxlPm* activation and the female/male differences in timing.

The existence of a regulatory region shared between *SxlPe* and *SxlPm* raises the question of how enhancer activity is directed to the correct promoter at the appropriate time. The 1.4 kb region regulates *SxlPe* from cycle 12 through early 14, yet the enhancer does not lead to significant expression from *SxlPm* until cycle 14. Expression from the two promoters overlaps briefly before *SxlPe* is silenced and *SxlPm* fully controls *Sxl* transcription. We imagine two general mechanisms that might explain how

the enhancer could choose between the two promoters [96,97]. First, an insulator situated between the enhancer and the upstream promoter might block the 1.4 kb region from interacting with *SxlPm* until the insulating protein was removed from the DNA or its activity overcome by additional positive signals. Second, promoter choice could be dictated by differences in the transcription machinery at the two promoters [96], or by a temporally-restricted transcription factor that recruits the enhancer to one of the two *Sxl* promoters. The developmentally regulated competition between the promoters of the chicken ε-globin and β-globin genes for their common enhancer provides precedent for the latter mechanism [98]. The rapid changeover from *SxlPe* to *SxlPm* coincides with the *Drosophila* maternal to zygotic transition, when expression of the zygotic genome begins in earnest and numerous early mRNAs and proteins are eliminated from the embryo [99]. It would not be surprising if the rapid changes at *Sxl* were directly connected to more general regulatory events occurring during this dynamic period of development.

Methods

P-element vectors and transformation

Sxl genomic fragments were made using Expand Long Template PCR System (Roche), cloned into pCRII-TOPO (Invitrogen), and ligated into P-element transformation vectors based on pCaSpeR-AUG-ßgal. Germline transformations were performed by Genetic Services Inc. (Cambridge, MA). Transgenes with internal deletions were cloned as upstream and downstream fragments and joined at primer-derived *Pac* I sites. The

construct -1.8kb*SxlPm-lacZ* was made with primers 1.8Pm5' and 1.8Pm3'; *SxlPm[+Pe]-lacZ* with primers 1' and 4'; *SxlPm[Δ-88Pe]-lacZ* with primers 1' and 2'/PacI and 3''PacI and 4'; *SxlPm[Δ-1.4Pe]-lacZ* with primers 1' and 5'/PacI and 3''PacI and 4'. Control transgenes *SxlPe[L2]-lacZ* and *SxlPe[Δ-88,L2]-lacZ* were similar to *SxlPm[+Pe]-lacZ* and *SxlPm[Δ-88Pe]-lacZ* except sequences distal to 1.4 kb upstream of exon E1 were absent and vector was pPelican. Two independent lines of *SxlP[L2]-lacZ* were expressed similarly to previous 1.4 kb *SxlPe-lacZ* lines [41,95], but deletion of the core *SxlPe* promoter left both tested lines of *SxlPe[Δ-88,L2]-lacZ* inactive. Primer sequences: 1.8Pm5'-ctcacgctagagaacaccgatcattc; 1.8Pm3'-gactttccttcttcggcaaC; 1'-CCATCCGATCCGCGAGTCCA; 4'-GCACGCTCACTGTGCTTTCCTCTC; 2'/PacI-CCAttaattaaGGAGGCAAGGTGCGCGT; 3''/PacI-CCAttaattaaCGTAACTTTGTGATTATCCC; 5'/PacI-CCttaattaaATGCGAGCAGCGGAGAAGGG.

***In-situ* hybridization**

Non-fluorescent *in situ* hybridization used digoxygenin or fluoroscene-labeled probes [100]. *D. melanogaster* and *D. simulans SxlPm* intron probes (1.4 kb) were transcribed from templates made using primers: Pm5'-CCCTTCTCCGCTGCTCGCAT and T3Pm-aattaaccctcactaaagggCCAGGTAGAAGATCGAAGGA. Templates for corresponding *D. yakuba* and *D. virilis SxlPm* probes were made with yakPm5'-CACCACCCCATTCCACCCG and T3Pm, or virPm5'-CGAGCCTTTCCGTAACTGTTCG and virT3Pm-

aattaaccctcactaaagggTGCGCTACCTGTTGACAGTG. Probes for *lacZ* and exon E1 (*SxlPe*) have been described [41,83,101]. Fluorescent detection of *SxlPe* transcripts was as detailed: http://superfly.ucsd.edu/~davek/. Nascent transcripts, visible as dots within stained nuclei, were seen with all probes but were more difficult to detect from *lacZ* transgenes. For X-linked genes the number of nuclear dots indicates chromosomal sex. Times within cycle 14 were estimated by nuclear shape and length, and by the extent of membrane furrow invagination [82,81]. Specific developmental time estimates were based on published literature, but embryos grouped within specified time periods were staged as closely as possible to each other.

Genetic analysis

Alleles $sisA^{1}$, da^{1}, and sc^{sisB3} are near null for sex determination. gro^{E48} and run^{3} are null alleles. Embryos homozygous or hemizygous for sc^{sisB3} and $sisA^{1}$ were generated using the constitutive Sxl^{M4} allele to bypass female-lethal effects [81]. Null allele Sxl^{fl} suppressed the male-lethality of the 2X *P(mini-*w^{+}*,* $sisA^{+}$*)* & 2X *P(mini-*w^{+}*,* sc^{sisB+}*)* chromosome [81]. Nascent transcripts from Sxl^{fl} and Sxl^{M4} are not detectably different from wild type . Germline clones [63] were generated in larvae of flies with the genotype: *P[hsFLP]1,* y^{1} w^{1118}/w^{1118}; *P[neoFRT]82B* ry^{506} gro^{E48}/*P[neoFRT]82B P[ovoD1-18]3R and P[hsFLP]1,* y^{1}. Females with recombinant germlines were crossed to w^{1118}/Y males. Embryos were collected at 25°C. Other mutations and chromosomes are described: http://flybase.bio.indiana.edu. The sc^{sisB3} allele and transgenic XSE duplications were provided by T. Cline (University of California, Berkeley). *FRT82B*

gro^{E48} was from by P. Simpson (University of Cambridge). *D. virilis* was provided by S. Johnson (Texas A&M University). *D. simulans*, *D. yakuba* and *D. simulans* were from D. Barbash (Cornell University). Fly stocks for FLP/FRT recombination were from the Bloomington *Drosophila* stock center.

REFERENCES

1. Bell LR, Maine EM, Schedl P, Cline TW (1988) Sex-lethal, a Drosophila sex determination switch gene, exhibits sex-specific RNA splicing and sequence similarity to RNA binding proteins. Cell 55: 1037-1046.

2. MacDougall C, Harbison D, Bownes M (1995) The developmental consequences of alternate splicing in sex determination and differentiation in Drosophila. Dev Biol 172: 353-376.

3. Cline TW, Meyer BJ (1996) Vive la difference: males vs females in flies vs worms. Annu Rev Genet 30: 637-702.

4. Penalva LO, Sanchez L (2003) RNA binding protein sex-lethal (Sxl) and control of Drosophila sex determination and dosage compensation. Microbiol Mol Biol Rev 67: 343-359.

5. Sosnowski BA, Belote JM, McKeown M (1989) Sex-specific alternative splicing of RNA from the transformer gene results from sequence-dependent splice site blockage. Cell 58: 449-459.

6. Valcarcel J, Singh R, Zamore PD, Green MR (1993) The protein Sex-lethal antagonizes the splicing factor U2AF to regulate alternative splicing of transformer pre-mRNA. Nature 362: 171-175.

7. Boggs RT, Gregor P, Idriss S, Belote JM, McKeown M (1987) Regulation of sexual differentiation in D. melanogaster via alternative splicing of RNA from the transformer gene. Cell 50: 739-747.

8. Salz HK, Erickson JW (2010) Sex determination in Drosophila: The view from the top. Fly (Austin) 4: 60-70.

9. Burtis KC, Baker BS (1989) Drosophila doublesex gene controls somatic sexual differentiation by producing alternatively spliced mRNAs encoding related sex-specific polypeptides. Cell 56: 997-1010.

10. Baker BS (1989) Sex in flies: the splice of life. Nature 340: 521-524.

11. Heinrichs V, Ryner LC, Baker BS (1998) Regulation of sex-specific selection of fruitless 5' splice sites by transformer and transformer-2. Mol Cell Biol 18: 450-458.

12. Song HJ, Taylor BJ, Edgar BA editor (1997) Analysis of lethal fru mutants reveal larval neuronal phenotypes. Cold Spring Harbor (New York): Cold Spring Harbor Press. 198p.

13. Dulac C (2005) Sex and the single splice. Cell 121: 664-666.

14. Goodwin SF, Villella A, Anand A, Ryner LC, Edgar BA editor (1997) Behavioral and molecular dissection of fruitless (fru). Cold Spring Harbor (New york): Cold Spring Harbor Press. 248p.

15. Villella A, Gailey DA, Berwald B, Ohshima S, Barnes PT, et al. (1997) Extended reproductive roles of the fruitless gene in Drosophila melanogaster revealed by behavioral analysis of new fru mutants. Genetics 147: 1107-1130.

16. Demir E, Dickson BJ (2005) fruitless splicing specifies male courtship behavior in Drosophila. Cell 121: 785-794.

17. Ito H, Fujitani K, Usui K, Shimizu-Nishikawa K, Tanaka S, et al. (1996) Sexual orientation in Drosophila is altered by the satori mutation in the sex-determination gene fruitless that encodes a zinc finger protein with a BTB domain. Proc Natl Acad Sci U S A 93: 9687-9692.

18. Gergen JP (1987) Dosage Compensation in Drosophila: Evidence That daughterless and Sex-lethal Control X Chromosome Activity at the Blastoderm Stage of Embryogenesis. Genetics 117: 477-485.

19. Birchler J, Sun L, Fernandez H, Donohue R, Xie W, et al. (2011) Re-evaluation of the function of the male specific lethal complex in Drosophila. J Genet Genomics 38: 327-332.

20. Lucchesi JC, Kelly WG, Panning B (2005) Chromatin remodeling in dosage compensation. Annu Rev Genet 39: 615-651.

21. Muller HJ (1932) Further studies on the nature and causes of gene mutations. Proc 6th Int Congr Genet. pp. 213-255.

22. Gelbart ME, Kuroda MI (2009) Drosophila dosage compensation: a complex voyage to the X chromosome. Development 136: 1399-1410.

23. Kelley RL, Meller VH, Gordadze PR, Roman G, Davis RL, et al. (1999) Epigenetic spreading of the Drosophila dosage compensation complex from roX RNA genes into flanking chromatin. Cell 98: 513-522.

24. Park Y, Kelley RL, Oh H, Kuroda MI, Meller VH (2002) Extent of chromatin spreading determined by roX RNA recruitment of MSL proteins. Science 298: 1620-1623.

25. Deng X, Meller VH (2006) Non-coding RNA in fly dosage compensation. Trends Biochem Sci 31: 526-532.

26. Alekseyenko AA, Peng S, Larschan E, Gorchakov AA, Lee OK, et al. (2008) A sequence motif within chromatin entry sites directs MSL establishment on the Drosophila X chromosome. Cell 134: 599-609.

27. Park SW, Oh H, Lin YR, Park Y (2010) MSL cis-spreading from roX gene up-regulates the neighboring genes. Biochem Biophys Res Commun 399: 227-231.

28. Gilfillan GD, Dahlsveen IK, Becker PB (2004) Lifting a chromosome: dosage compensation in Drosophila melanogaster. FEBS Lett 567: 8-14.

29. Gebauer F, Merendino L, Hentze MW, Valcarcel J (1998) The Drosophila splicing regulator sex-lethal directly inhibits translation of male-specific-lethal 2 mRNA. Rna 4: 142-150.

30. Gebauer F, Grskovic M, Hentze MW (2003) Drosophila sex-lethal inhibits the stable association of the 40S ribosomal subunit with msl-2 mRNA. Mol Cell 11: 1397-1404.

31. Duncan KE, Strein C, Hentze MW (2009) The SXL-UNR corepressor complex uses a PABP-mediated mechanism to inhibit ribosome recruitment to msl-2 mRNA. Mol Cell 36: 571-582.

32. Medenbach J, Seiler M, Hentze MW (2011) Translational control via protein-regulated upstream open reading frames. Cell 145: 902-913.

33. Erickson JW, Cline TW (1993) A bZIP protein, sisterless-a, collaborates with bHLH transcription factors early in Drosophila development to determine sex. Genes Dev 7: 1688-1702.

34. Cline TW (1993) The Drosophila sex determination signal: how do flies count to two? Trends Genet 9: 385-390.

35. Horabin JI, Schedl P (1993) Regulated splicing of the Drosophila sex-lethal male exon involves a blockage mechanism. Mol Cell Biol 13: 1408-1414.

36. Bell LR HJ, Schedl P, Cline TW (1991) Positive autoregulation of sex-lethal by alternative splicing maintains the female determined state in Drosophila. Cell 65: 229-239.

37. Cline TW (1984) Autoregulatory functioning of a Drosophila gene product that establish es and maintains the sexually determined state. Genetics 107: 231-277.

38. Johnson ML, Nagengast AA, Salz HK (2010) PPS, a large multidomain protein, functions with sex-lethal to regulate alternative splicing in Drosophila. PLoS Genet 6: e1000872.

39. Nagengast AA, Stitzinger SM, Tseng CH, Mount SM, Salz HK (2003) Sex-lethal splicing autoregulation in vivo: interactions between SEX-LETHAL, the U1 snRNP and U2AF underlie male exon skipping. Development 130: 463-471.

40. Cline TW (1988) Evidence that sisterless-a and sisterless-b are two of several discrete "numerator elements" of the X/A sex determination signal in Drosophila that switch Sxl between two alternative stable expression states. Genetics 119: 829-862.

41. Avila FW, Erickson JW (2007) Drosophila JAK/STAT pathway reveals distinct initiation and reinforcement steps in early transcription of Sxl. Curr Biol 17: 643-648.

42. Estes PA, Keyes LN, Schedl P (1995) Multiple response elements in the Sex-lethal early promoter ensure its female-specific expression pattern. Mol Cell Biol 15: 904-917.

43. Torres M, Sanchez L (1989) The scute (T4) gene acts as a numerator element of the X:a signal that determines the state of activity of sex-lethal in Drosophila. EMBO J 8: 3079-3086.

44. Torres M, Sanchez L (1991) The sisterless-b function of the Drosophila gene scute is restricted to the stage when the X:A ratio determines the activity of Sex-lethal. Development 113: 715-722.

45. Yang D, Lu H, Hong Y, Jinks TM, Estes PA, et al. (2001) Interpretation of X chromosome dose at Sex-lethal requires non-E-box sites for the basic helix-loop-helix proteins SISB and daughterless. Mol Cell Biol 21: 1581-1592.

46. Hou SX, Zheng Z, Chen X, Perrimon N (2002) The Jak/STAT pathway in model organisms: emerging roles in cell movement. Dev Cell 3: 765-778.

47. Gergen JP, Butler BA (1988) Isolation of the Drosophila segmentation gene runt and analysis of its expression during embryogenesis. Genes Dev 2: 1179-1193.

48. Manoukian AS, Krause HM (1993) Control of segmental asymmetry in Drosophila embryos. Development 118: 785-796.

49. Kania MA, Bonner AS, Duffy JB, Gergen JP (1990) The Drosophila segmentation gene runt encodes a novel nuclear regulatory protein that is also expressed in the developing nervous system. Genes Dev 4: 1701-1713.

50. Aronson BD, Fisher AL, Blechman K, Caudy M, Gergen JP (1997) Groucho-dependent and -independent repression activities of Runt domain proteins. Mol Cell Biol 17: 5581-5587.

51. Kramer SG, Jinks TM, Schedl P, Gergen JP (1999) Direct activation of Sex-lethal transcription by the Drosophila runt protein. Development 126: 191-200.

52. Lu H, Kozhina E, Mahadevaraju S, Yang D, Avila FW, et al. (2008) Maternal Groucho and bHLH repressors amplify the dose-sensitive X chromosome signal in Drosophila sex determination. Dev Biol 323: 248-260.

53. Younger-Shepherd S, Vaessin H, Bier E, Jan LY, Jan YN (1992) deadpan, an essential pan-neural gene encoding an HLH protein, acts as a denominator in Drosophila sex determination. Cell 70: 911-922.

54. Paroush Z, Finley RL, Jr., Kidd T, Wainwright SM, Ingham PW, et al. (1994) Groucho is required for Drosophila neurogenesis, segmentation, and sex determination and interacts directly with hairy-related bHLH proteins. Cell 79: 805-815.

55. Garrell J, Modolell J (1990) The Drosophila extramacrochaetae locus, an antagonist of proneural genes that, like these genes, encodes a helix-loop-helix protein. Cell 61: 39-48.

56. Bridges CB (1921) Triploid Intersexes in Drosophila Melanogaster. Science 54: 252-254.

57. Schutt C, Nothiger R (2000) Structure, function and evolution of sex-determining systems in Dipteran insects. Development 127: 667-677.

58. Fassler J, Landsman D, Acharya A, Moll JR, Bonovich M, et al. (2002) B-ZIP proteins encoded by the Drosophila genome: evaluation of potential dimerization partners. Genome Res 12: 1190-1200.

59. Fields S, Song O (1989) A novel genetic system to detect protein-protein interactions. Nature 340: 245-246.

60. Hewes RS, Schaefer AM, Taghert PH (2000) The cryptocephal gene (ATF4) encodes multiple basic-leucine zipper proteins controlling molting and metamorphosis in Drosophila. Genetics 155: 1711-1723.

61. Siegel JG (1981) Genetic characterization of the region of the Drosophila genome known to include the histone structural gene sequences. Genetics 98: 505-527.

62. Erickson JW, Cline TW (1998) Key aspects of the primary sex determination mechanism are conserved across the genus Drosophila. Development 125: 3259-3268.

63. Chou TB, Perrimon N (1996) The autosomal FLP-DFS technique for generating germline mosaics in Drosophila melanogaster. Genetics 144: 1673-1679.

64. Preston CR, Engels WR (1996) P-element-induced male recombination and gene conversion in Drosophila. Genetics 144: 1611-1622.

65. Ugarkovic D (2009) Centromere. New York: Springer.

66. Rong YS, Golic KG (2000) Gene targeting by homologous recombination in Drosophila. Science 288: 2013-2018.

67. Rong YS, Golic KG (2001) A targeted gene knockout in Drosophila. Genetics 157: 1307-1312.

68. Chen L, Glover JN, Hogan PG, Rao A, Harrison SC (1998) Structure of the DNA-binding domains from NFAT, Fos and Jun bound specifically to DNA. Nature 392: 42-48.

69. Ciapponi L, Bohmann D (2002) An essential function of AP-1 heterodimers in Drosophila development. Mech Dev 115: 35-40.

70. Ciapponi L, Jackson DB, Mlodzik M, Bohmann D (2001) Drosophila Fos mediates ERK and JNK signals via distinct phosphorylation sites. Genes Dev 15: 1540-1553.

71. Struffi P, Arnosti DN (2005) Functional interaction between the Drosophila knirps short range transcriptional repressor and RPD3 histone deacetylase. J Biol Chem 280: 40757-40765.

72. Lichty JJ, Malecki JL, Agnew HD, Michelson-Horowitz DJ, Tan S (2005) Comparison of affinity tags for protein purification. Protein Expr Purif 41: 98-105.

73. Gloeckner CJ, Boldt K, Schumacher A, Roepman R, Ueffing M (2007) A novel tandem affinity purification strategy for the efficient isolation and characterisation of native protein complexes. Proteomics 7: 4228-4234.

74. Bernstein M, Lersch RA, Subrahmanyan L, Cline TW (1995) Transposon insertions causing constitutive Sex-lethal activity in Drosophila melanogaster affect Sxl sex-specific transcript splicing. Genetics 139: 631-648.

75. Keyes LN, Cline TW, Schedl P (1992) The primary sex determination signal of Drosophila acts at the level of transcription. Cell 68: 933-943.

76. Salz HK, Maine EM, Keyes LN, Samuels ME, Cline TW, et al. (1989) The Drosophila female-specific sex-determination gene, Sex-lethal, has stage-, tissue-, and sex-specific RNAs suggesting multiple modes of regulation. Genes Dev 3: 708-719.

77. Samuels ME, Schedl P, Cline TW (1991) The complex set of late transcripts from the Drosophila sex determination gene sex-lethal encodes multiple related polypeptides. Mol Cell Biol 11: 3584-3602.

78. Louis M, Holm L, Sanchez L, Kaufman M (2003) A theoretical model for the regulation of Sex-lethal, a gene that controls sex determination and dosage compensation in Drosophila melanogaster. Genetics 165: 1355-1384.

79. Barbash DA, Cline TW (1995) Genetic and molecular analysis of the autosomal component of the primary sex determination signal of Drosophila melanogaster. Genetics 141: 1451-1471.

80. LN K (1995) Sex-specific activation of the Sex-lethal gene in Drosophila melanogaster Princeton University. 650 p.

81. Foe VA OG, Edgar BA editor (1993) Mitosis and morphogenesis in the Drosophila embryo: Point and Counterpoint: Cold Spring Harbor (New York): Cold Spring Harbor Press. 149-300 p.

82. Grosshans J, Muller HA, Wieschaus E (2003) Control of cleavage cycles in Drosophila embryos by fruhstart. Dev Cell 5: 285-294.

83. Erickson JW, Quintero JJ (2007) Indirect effects of ploidy suggest X chromosome dose, not the X:A ratio, signals sex in Drosophila. PLoS Biol 5: e332.

84. Erickson JW, Cline TW (1991) Molecular nature of the Drosophila sex determination signal and its link to neurogenesis. Science 251: 1071-1074.

85. Walker JJ, Lee KK, Desai RN, Erickson JW (2000) The Drosophila melanogaster sex determination gene sisA is required in yolk nuclei for midgut formation. Genetics 155: 191-202.

86. Fuxreiter M, Simon I, Bondos S (2011) Dynamic protein-DNA recognition: beyond what can be seen. Trends Biochem Sci 36: 415-423.

87. Mittag T, Kay LE, Forman-Kay JD (2010) Protein dynamics and conformational disorder in molecular recognition. J Mol Recognit 23: 105-116.

88. Linding R, Jensen LJ, Diella F, Bork P, Gibson TJ, et al. (2003) Protein disorder prediction: implications for structural proteomics. Structure 11: 1453-1459.

89. Meszaros B, Simon I, Dosztanyi Z (2009) Prediction of protein binding regions in disordered proteins. PLoS Comput Biol 5: e1000376.

90. Davuluri RV, Suzuki Y, Sugano S, Plass C, Huang TH (2008) The functional consequences of alternative promoter use in mammalian genomes. Trends Genet 24: 167-177.

91. Graveley BR (2002) Sex, AGility, and the regulation of alternative splicing. Cell 109: 409-412.

92. Lallena MJ, Chalmers KJ, Llamazares S, Lamond AI, Valcarcel J (2002) Splicing regulation at the second catalytic step by Sex-lethal involves 3' splice site recognition by SPF45. Cell 109: 285-296.

93. Gladden JM, Farboud B, Meyer BJ (2007) Revisiting the X:A signal that specifies Caenorhabditis elegans sexual fate. Genetics 177: 1639-1654.

94. Gladden JM, Meyer BJ (2007) A ONECUT homeodomain protein communicates X chromosome dose to specify Caenorhabditis elegans sexual fate by repressing a sex switch gene. Genetics 177: 1621-1637.

95. Estes PA, Keyes LN, Schedl P (1995) Multiple response elements in the Sex-lethal early promoter ensure its female-specific expression pattern. Mol Cell Biol 15: 904-917.

96. Ohtsuki S, Levine M, Cai HN (1998) Different core promoters possess distinct regulatory activities in the Drosophila embryo. Genes Dev 12: 547-556.

97. Arnosti DN (2002) Design and function of transcriptional switches in Drosophila. Insect Biochem Mol Biol 32: 1257-1273.

98. Foley KP, Engel JD (1992) Individual stage selector element mutations lead to reciprocal changes in beta- vs. epsilon-globin gene transcription: genetic confirmation of promoter competition during globin gene switching. Genes Dev 6: 730-744.

99. Tadros W, Westwood JT, Lipshitz HD (2007) The mother-to-child transition. Dev Cell 12: 847-849.

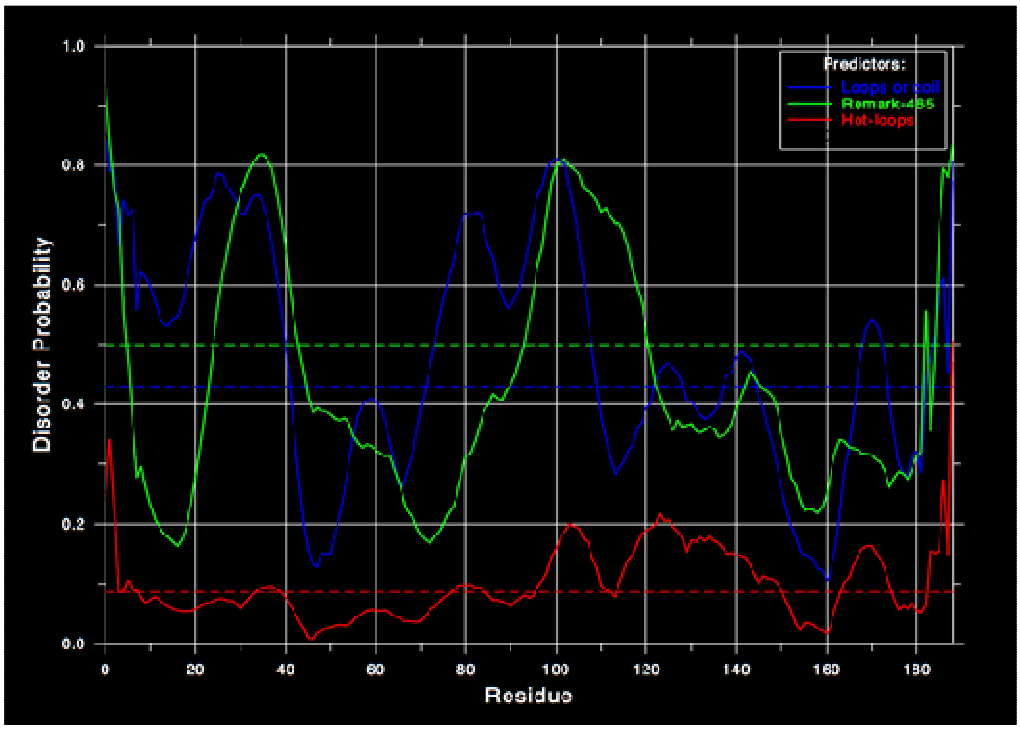
100. Lehmann R, Tautz D (1994) In situ Hybridization to RNA. In: Goldstein LSB, Fyrberg EA, editors. Drosophila melanogaster: Practical Uses in Cell and Molecular Biology. San Diego: Academic Press. pp. 576-597.

101. Erickson JW, Cline TW (1993) A bZIP protein, Sisterless A, collaborates with bHLH transcription factors early in Drosophila development to determine sex. Genes Dev 7: 1688-1702.

102. Cline TW (1988) Evidence that sisterless-a and sisterless-b are two of several discrete "numerator elements" of the X:A sex determination signal in Drosophila that switch Sex-lethal between two alternative stable expression states. Genetics 119: 829-862.

APPENDIX A

SisA intrinsically disorder regions were determined using the DisEMBL program. The data obtained from running this program is shown below.



Disordered by Loops/coils definition

>SisA_LOOPS 1-41, 73-110
MERSHLYLPT LSYAAMGHVY APYRGSSSPA LSTASSTSSK Peqieelvsq qlhhlkmhya deeqryvdqm llENPIVVER RAPPPLKTEL AMDCRGSGSG SGSGSGSDVK daqrqraesc rksrynnkik kaklrfrhkf vsgqlkksav mldtmrdvia qaerqllerg ypaatlermr atfglemeq

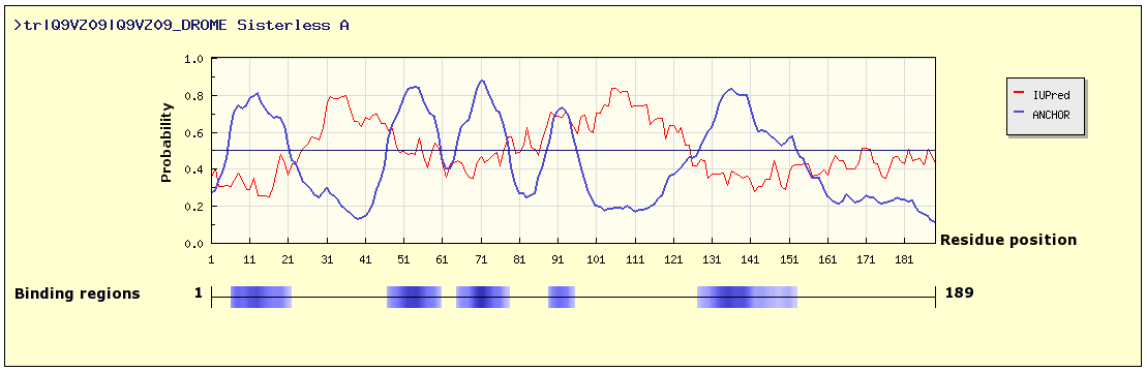
Disordered by Hot-loops definition

>SisA_HOTLOOPS 97-151, 165-175
mershlylpt lsyaamghvy apyrgssspa lstasstssk peqieelvsq qlhhlkmhya deeqryvdqm llenpivver rappplktel amdcrgSGSG SGSGSGSDVK DAQRQRAESC RKSRYNNKIK KAKLRFRHKF VSGQLKKSAV Mldtmrdvia qaerQLLERG YPAATlermr atfglemeq

Disordered by Remark-465 definition

>SisA_REM465 25-43, 94-121
mershlylpt lsyaamghvy apyrGSSSPA LSTASSTSSK PEQieelvsq qlhhlkmhya deeqryvdqm llenpivver rappplktel amdCRGSGSG SGSGSGSDVK DAQRQRAESC Rksrynnkik kaklrfrhkf vsgqlkksav mldtmrdvia qaerqllerg ypaatlermr atfglemeq

Data collected from ANCHOR



Predicted Disordered Binding Regions			
	From	To	Length
1	6	21	16
2	47	60	14
3	65	78	14
4	89	95	7
5	128	153	26

Position Specific Score			
Position	Residue	ANCHOR Probability	Output
1	M	0.2684	0
2	E	0.2824	0
3	R	0.3372	0
4	S	0.3858	0
5	H	0.4590	0
6	L	0.6233	1
7	Y	0.7114	1
8	L	0.7435	1
9	P	0.7269	1
10	T	0.7442	1
11	L	0.7831	1
12	S	0.7917	1
13	Y	0.8088	1
14	A	0.7676	1
15	A	0.7305	1
16	M	0.6978	1
17	G	0.6736	1
18	H	0.6810	1
19	V	0.6765	1
20	Y	0.6375	1
21	A	0.5475	1
22	P	0.4454	0
23	Y	0.4315	0
24	R	0.3717	0

25	G	0.3303	0
26	S	0.3064	0
27	S	0.2832	0
28	S	0.2594	0
29	P	0.2481	0
30	A	0.2761	0
31	L	0.3001	0
32	S	0.2694	0
33	T	0.2513	0
34	A	0.2280	0
35	S	0.1945	0
36	S	0.1781	0
37	T	0.1597	0
38	S	0.1413	0
39	S	0.1308	0
40	K	0.1319	0
41	P	0.1447	0
42	E	0.1706	0
43	Q	0.2087	0
44	I	0.2891	0
45	E	0.3445	0
46	E	0.4256	0
47	L	0.5652	1
48	V	0.6371	1
49	S	0.6801	1
50	Q	0.7200	1
51	Q	0.7790	1
52	L	0.8342	1
53	H	0.8353	1
54	H	0.8450	1
55	L	0.8383	1
56	K	0.7764	1
57	M	0.7282	1
58	H	0.6958	1
59	Y	0.6798	1
60	A	0.5871	1

61	D	0.4630	0
62	E	0.4042	0
63	E	0.4007	0
64	Q	0.4558	0
65	R	0.5428	1
66	Y	0.6311	1
67	V	0.6538	1
68	D	0.6646	1
69	Q	0.7408	1
70	M	0.8356	1
71	L	0.8773	1
72	L	0.8715	1
73	E	0.8124	1
74	N	0.7659	1
75	P	0.7123	1
76	I	0.7008	1
77	V	0.6353	1
78	V	0.5435	1
79	E	0.4065	0
80	R	0.3238	0
81	R	0.2682	0
82	A	0.2670	0
83	P	0.2483	0
84	P	0.2587	0
85	P	0.2677	0
86	L	0.3535	0
87	K	0.4101	0
88	T	0.4912	0
89	E	0.5698	1
90	L	0.6765	1
91	A	0.7121	1
92	M	0.7340	1
93	D	0.7080	1
94	C	0.6878	1
95	R	0.5763	1
96	G	0.4969	0

97	S	0.4235	0
98	G	0.3484	0
99	S	0.2760	0
100	G	0.2266	0
101	S	0.2030	0
102	G	0.1930	0
103	S	0.1752	0
104	G	0.1819	0
105	S	0.1856	0
106	G	0.1871	0
107	S	0.1880	0
108	D	0.1856	0
109	V	0.1986	0
110	K	0.1835	0
111	D	0.1668	0
112	A	0.1797	0
113	Q	0.1800	0
114	R	0.1846	0
115	Q	0.1946	0
116	R	0.2062	0
117	A	0.2418	0
118	E	0.2564	0
119	S	0.3086	0
120	C	0.3648	0
121	R	0.3678	0
122	K	0.3906	0
123	S	0.4076	0
124	R	0.4341	0
125	Y	0.4617	0
126	N	0.4585	0
127	N	0.4715	0
128	K	0.5205	1
129	I	0.5663	1
130	K	0.6014	1
131	K	0.6252	1
132	A	0.6893	1

133	K	0.7484	1
134	L	0.7975	1
135	R	0.8188	1
136	F	0.8338	1
137	R	0.8131	1
138	H	0.7975	1
139	K	0.7993	1
140	F	0.7996	1
141	V	0.7423	1
142	S	0.6661	1
143	G	0.6023	1
144	Q	0.6064	1
145	L	0.6052	1
146	K	0.5779	1
147	K	0.5616	1
148	S	0.5482	1
149	A	0.5236	1
150	V	0.5425	1
151	M	0.5704	1
152	L	0.5731	1
153	D	0.5028	1
154	T	0.4651	0
155	M	0.4497	0
156	R	0.3991	0
157	D	0.3521	0
158	V	0.3527	0
159	I	0.3510	0
160	A	0.2988	0
161	Q	0.2517	0
162	A	0.2338	0
163	E	0.2168	0
164	R	0.2135	0
165	Q	0.2296	0
166	L	0.2604	0
167	L	0.2406	0
168	E	0.2155	0

169	R	0.2216	0
170	G	0.2348	0
171	Y	0.2589	0
172	P	0.2483	0
173	A	0.2449	0
174	A	0.2241	0
175	T	0.2108	0
176	L	0.2176	0
177	E	0.2237	0
178	R	0.2290	0
179	M	0.2459	0
180	R	0.2326	0
181	A	0.2370	0
182	T	0.2250	0
183	F	0.2308	0
184	G	0.1946	0
185	L	0.1653	0
186	E	0.1559	0
187	M	0.1433	0
188	E	0.1219	0
189	Q	0.1134	0

VITA

Name:	Alejandra Noemi González Rojos
Address:	Department of Biology, Texas A&M University, 3258 TAMU, College Station, TX 77843.
Email Address:	agonzalezrojos@gmail.com
Education:	B.A., Biochemistry, Pontificia Universidad Catolica de Chile, 2000
	Ph. D., Biology, Texas A&M University, 2012