Distribution Agreement

In presenting this thesis as a partial fulfillment of the requirements for a degree from Emory University, I hereby grant to Emory University and its agents the non-exclusive license to archive, make accessible, and display my thesis in whole or in part in all forms of media, now or hereafter know, including display on the World Wide Web. I understand that I may select some access restrictions as part of the online submission of this thesis. I retain all ownership rights to the copyright of the thesis. I also retain the right to use in future works (such as articles or books) all or part of this thesis.

Kevin Man Hin Luk April 17, 2013

Role of Topoisomerase I in *Drosophila melanogaster's* mechanism of Dosage Compensation

By

Kevin Man Hin Luk

Dr. John C. Lucchesi, Ph.D.
Advisor

Department of Biology

Dr. John C. Lucchesi, Ph.D.
Advisor

Dr. Paula Vertino, Ph.D.
Committee Member

Dr. Victor Corces, Ph.D.
Committee Member

2013

Role of Topoisomerase I in *Drosophila melanogaster's* mechanism of Dosage Compensation

By

Kevin Man Hin Luk

Dr. John C. Lucchesi, Ph.D.

Advisor

An abstract of

a thesis submitted to the Faculty of Emory College of Arts and Sciences

of Emory University in partial fulfillment

of the requirements of the degree of

Bachelor of Science with Honors

Department of Biology

2013

Abstract

Role of Topoisomerase I in *Drosophila melanogaster's* mechanism of Dosage Compensation

By Kevin Man Hin Luk

Dosage compensation is the equalization of X-linked gene products between male and females. In male *Drosophila melanogaster*, the single X-chromosome is hyper-transcribed twofold and is mediated by the Male Specific Lethal (MSL) complex. This twofold increase in expression of the male X chromosome is achieved by an enhanced rate of transcription elongation. Topoisomerases are a family of enzymes that maintain nucleic acid topology during cellular processes, such as transcription. Based on topoisomerases' role in transcription and experimental evidence revealing topoisomerase II's role in *Drosophila* dosage compensation, an investigation of topoisomerase I's role in dosage compensation was warranted. Using RNA interference, a plasmid model system that reproduces dosage compensation, and qRT-PCR of endogenous genes, this study reveals the ability of the MSL complex to mediate dosage compensation in male *Drosophila* even in the absence of topoisomerase I. This observation further elucidates the relationship of topoisomerase enzymes and MSL complex mediated dosage compensation.

Role of Topoisomerase I in *Drosophila melanogaster's* mechanism of Dosage Compensation

By

Kevin Man Hin Luk

Dr. John C. Lucchesi, Ph.D.

Advisor

A thesis submitted to the Faculty of Emory College of Arts and Sciences

of Emory University in partial fulfillment

of the requirements of the degree of

Bachelor of Science with Honors

Department of Biology

2013

Acknowledgements

I would like to thank Dr. John C. Lucchesi for providing me with my first opportunity to experience academic research in the spring of 2011. I am thankful to Dr. Lucchesi for helping design my experiment and supporting me throughout my time in the lab. My experience has undoubtedly made me realize that I want to continue to conduct original research throughout my medical training years.

A special thank you goes to Simona Cugusi, a post-doctoral fellow in the Lucchesi lab. Since I first entered the lab in the fall of 2011, Simona has patiently and diligently taught me all of the laboratory techniques I know to this day. She taught me both the technical and theoretical aspects of the experiments I performed, and she helped further design and guide the experiments for my project. In addition, Simona co-conducted many of the procedures regarding the topoisomerase I antigen design and synthesis. Thanks to Simona for mentoring me in the lab.

Thank you to everyone in the Lucchesi lab, including Satish, Huiping, and Heiying, for resource advice and discussion. The members of the lab all kindly provided insight into a career in research as well as research protocols. I thank you all for your guidance.

Thank you to Dr. Victor Corces and Dr. Paula Vertino for serving as my honors committee and providing helpful ideas for my project. I understand committing to an honors committee requires a lot of time and dedication, and I thank you for providing these as I accomplished this thesis.

Table of Contents

List of Figures and Tables

General Background

Differences in X-Linked Gene Products

From size, eye color, sex and all other traits, the genome housed within cells directs an organism's form and functionality. In particular, sex determination is often established by dimorphic sex chromosomes. In the heavily characterized "XY system", organisms with the XX genotype are females, while males inherit the XY genotype. The Y chromosome was initially homologous to the X-chromosome, but the Y chromosome's change in size and activity over time has lead to an imbalance in the amount of products produced from the sex chromosomes between males and females (Charlesworth, 1996). In both sexes, the X-chromosome encodes for non-sex-specific products required by the organism for proper development and maintenance. Thus, because inequalities in the number of X-chromosomes exist, equalization strategies of X-linked gene products exist between different groups. For example, in *C.elegans*, both of the X-chromosomes in hermaphrodites are down regulated twofold; effectively "reducing" the amount of X-linked gene products to equal that expressed by the single X-chromosome in males (Gelbart and Kuroda, 2009). In contrast, evolution of the Y-chromosome has lead to the co-evolution of a different way to achieve dosage compensation in male *Drosophila melanogaster*. This mechanism balances the amount of X-linked gene products between male and female flies by up regulating expression of the single X-chromosome in males by

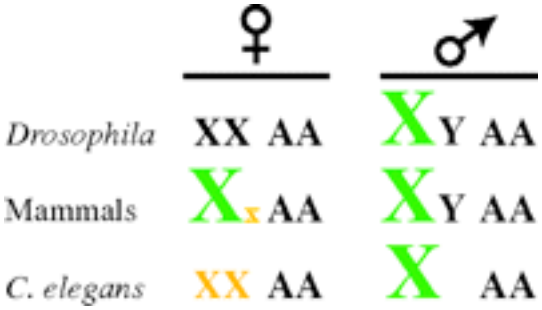

Figure 1. X chromosome dosage compensation mechanisms of D. melanogaster, mammals, and C.elegans (Gelbart and Kuroda, 2009).

approximately twofold (Gelbart and Kuroda, 2009). This allows the amount of X-linked gene products between male and female *Drosophila melanogaster* to be equalized.

The MSL Complex and Dosage Compensation

In male *Drosophila*, the male-specific lethal (MSL) complex mediates dosage compensation. Females express the master sex regulator (SXL) protein that inhibits translation of MSL2 mRNA and thus, prevents the formation of the MSL complex (Hallacli and Akhtar, 2009). The MSL complex consists of five protein subunits and two noncoding RNAs. These include: MSL1, MSL2, MSL 3, males absent on the first (MOF), maleless (MLE) and two roX RNAs (Yokoyama et al., 2007). MSL1 forms an assembly platform by interacting directly with all other components, except MLE, which is associated to the complex in an RNA dependent manner. A leucine zipper like motif in MSL1's N-terminus interacts with MSL2 (Li et al., 2005) while its carboxyl terminus interacts with MSL3 and MOF (Scott et al., 2000). Copps et al. (1998) reported that a mutation of MSL2's RING finger domain and its cysteine rich motif disrupts the interaction between MSL1 and MSL2; suggesting a role for these structures in binding between MSL1 and MSL2. Furthermore, MSL3 and MOF bind

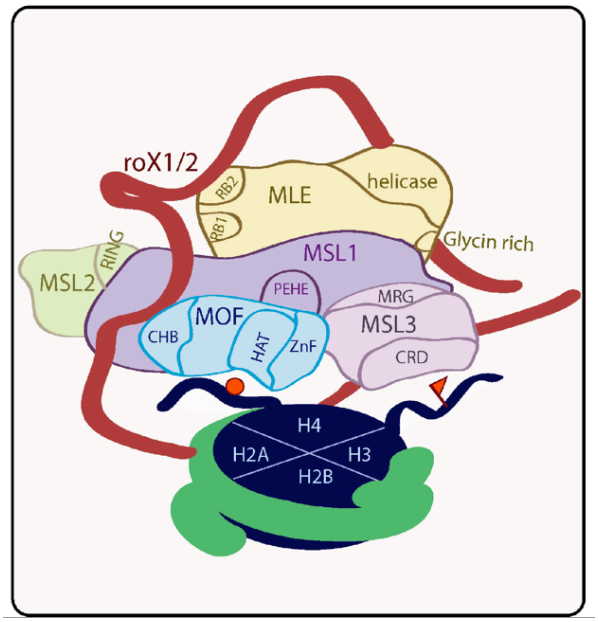

Figure 2. MSL Complex Schematic (Hallacli and Akhtar, 2009).

MSL1 via a zinc finger motif and MRG domain, respectively. MOF's chromobarrel (CHB) domain interacts with RNA, and MSL3's chromo related domain (CRD) binds DNA and nucleosomes. These interactions help anchor the MSL complex to chromatin and maintain its structure (Hallacli and Akhtar, 2009). MOF, MSL2, and MLE exhibit enzymatic activity. MOF is an acetyl-transferase and acetylates H4 at Lysine 16; Recently, it has also been shown that MSL2 acts as an E3 ubiquitin ligase (Wu et al., 2011). MLE is an ATP dependent helicase that unwinds DNA/RNA or RNA/RNA substrates (Gelbart and Kuroda, 2009).

RoX1 and roX2 are male specific noncoding RNAs, localized in the nucleus, and expressed in all tissues. Although the roX RNAs are functionally redundant, they are significantly different in size and sequence. Their large size lends themselves to several possible interaction sites with the MSL complex, but the direct contact points between roX RNAS and MLE, MOF, and MSL3 have not been established (Gelbart and Kuroda, 2009). Both noncoding RNAs colocalize with MSL at various points along the X-chromosome. Furthermore, the roX loci are primary nucleation sites of the complex, thus they play a key role in targeting the MSL complex to the appropriate sites on the X chromosome (Franke and Baker, 1999).

The means by which these components combine into the MSL complex and facilitate the twofold hyper-transcription of the single male X-chromosome in *Drosophila* is currently the subject of investigation. Larschan et al. (2011) recently supported that dosage compensation in flies is achieved by enhanced transcription elongation. Moreover, Conrad et al. (2012) reported an enrichment of polymerase at the promoter region of known dosage compensated regions of the male X-chromosome. The

enrichment of polymerase at the promoter is probably a reflection of increased re-initiation of transcription elongation. The MSL complex's function is likely to be modifying chromatin organization to facilitate transcription elongation through its acetyltransferase, helicase, or ubiquitin ligase. Acetylation levels of H4K16 tend to increase towards the 3' end of X-linked transcriptional domains (Smith et al., 2001), and acetylation of histones has been shown to facilitate chromatin relaxation and enhanced transcription (Gelbart and Kuroda, 2009). It is unclear if or how the MSL complex's helicase or ubiquitin ligase activities facilitate an enhanced rate of gene transcription.

Nucleic Acid Topology and Gene Transcription

Gene transcription, the process of generating ribonucleic acid (RNA) from a deoxyribonucleic acid (DNA) template, is increased twofold on the single X-chromosome in male *Drosophila*. A number of forms of RNA exist, including: rRNA, tRNA, and mRNA. The enzyme, RNA polymerase, synthesizes mRNA as it moves along and reads a single strand of DNA. Messenger RNA (mRNA) is a molecule of RNA used to produce most of the proteins in cells during normal functioning. However, the extended double helical structure of DNA creates unique barriers for gene expression in living organisms. In order for RNA polymerase to read DNA and transcription to occur efficiently, the two strands of double stranded DNA must be separated (Lodish et al., 2008). This DNA strand separation during transcription induces other changes to DNA topology. The progressing RNA polymerase produces localized positive and negative supercoiling in front and behind the transcription bubble, respectively. This structural

stress can affect the progression of transcription and lead to drastic changes in gene expression (Liu and Wang, 1987; French et al., 2011).

Topoisomerase I and II

Topoisomerases are a family of enzymes used in cells to maintain DNA topology – including transcription induced positive and negative supercoiling. Topoisomerases have been implicated in processes such as DNA repair, transcription, replication, and chromosome compaction. In all processes, topoisomerases operate by introducing breaks, twisting the nucleic acid, and ligating the DNA strands back together. This general procedure allows topoisomerases to add or remove DNA supercoils and disentangle DNA strands. Normally, one turn of the DNA helix occurs every 10.5 base pairs on its axis relative to the polymerase during transcription. Positive torsion in front of the polymerase and negative torsion behind the RNA polymerase compacts this distance (French et al., 2011).

Two types of topoisomerases have been discovered. Type I topoisomerases cleave one strand of DNA while type II topoisomerases cleave two strands (French et al., 2011). These two types can be further subdivided. Topoisomerases labeled with even roman numerals belong to the type II family, and odd roman numeral topoisomerases belong to type I. Finally, subtypes A, B, and C are distinguished based on unique amino acid sequences and resultant globular structure. Thus, there is a wide diversity in topoisomerase function and mechanism (Vos et al., 2011).

Type IA topoisomerases use a “strand passage” mechanism. One of the two strands of DNA is cleaved and physically opened to allow the second, uncleaved, DNA

strand to pass through the gap. The cleaved strand is reattached after this pass through. Type IA topoisomerases are found in all cellular domains of life. It has been reported that the primary action of Topoisomerase IA is to relax negatively supercoiled DNA. Type IB and IC topoisomerases differ in structure and mechanism from type IA. Again, one of the two strands of DNA is cleaved. However, type IB and IC topoisomerases allow the cleaved end to rotate with respect to the other around the intact phosphodiester bond. Appropriately, this mechanism is labeled the "swivelase" mechanism. Type IB and IC topoisomerases are distinguished by their active sites (Vos et al., 2011).

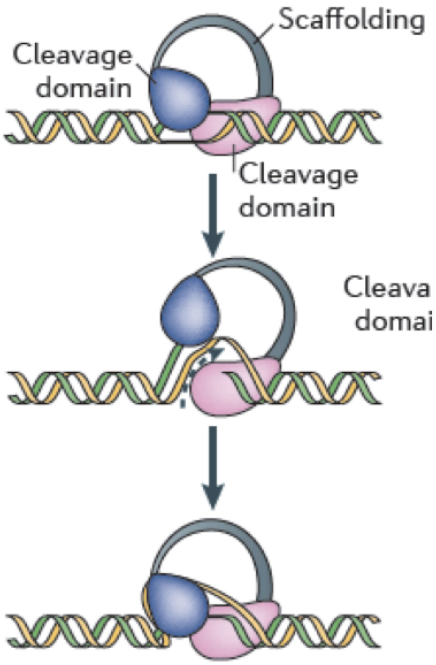

Figure 3. Topoisomerase I Mechanism of Torsional Strain Relief (Vos et al., 2011)

Type IIA and IIB are both known to relieve positive and negative supercoiling. Both also act through a strand passage mechanism. Specifically, type IIA and IIB topoisomerases cleave both strands in DNA and passes an intact duplex through the transient break. The cleavage domains and the ATP-powered strand cleavage are present in both types. However, type IIB is mostly found in archaea, plants, and bacteria, while type IIA is mostly found in eukaryotes (Vos et al., 2011). Because topoisomerases play a prominent role in DNA topology maintenance during transcription elongation, it can be theorized that these enzymes play a role in dosage compensation in *Drosophila melanogaster* cells. However, the mechanisms by which topoisomerases act during gene transcription, and thus dosage compensation, are poorly understood.

A paper currently under review from the Lucchesi lab reported that topoisomerase II is involved in the mechanism of dosage compensation – an observation based on the following findings. Dosage compensated genes exhibit a reduced level of negative supercoiling and topoisomerase II is enriched on X-linked dosage compensated genes in S2 cells. This enhanced recruitment to the male X-chromosome in Drosophila is facilitated by MLE in a RNA dependent manner. Finally, using a plasmid model system of dosage compensation, it was determined that dosage compensation is inhibited in light of topoisomerase II RNA interference. These findings suggest that the MSL complex alters supercoiling during transcription to facilitate dosage compensation.

In the absence of topoisomerase I, plasmid templates exhibited high negative torsion when being transcribed. Extreme negative tension has been implicated in transient blocks of transcription elongation (French et al., 2011). These findings, taken together with recent work done on topoisomerase II mentioned above, suggested a possible role for topoisomerase I in dosage compensation.

Purpose

The working hypothesis in the Lucchesi lab has been that dosage compensation in male *Drosophila melanogaster* cells is achieved by an enhanced rate of transcription elongation – a fact again recently supported by Larschan et al (2011). In light of the crucial role in topology maintenance during transcription by topoisomerase and the lab's recent findings regarding topoisomerase II role in dosage compensation, whether topoisomerase I plays a role in this mechanism requires investigation. The purpose of my

project is to investigate topoisomerase I's relationship, if any, with the MSL complex and its possible role in dosage compensation in male *Drosophila melanogaster* cells.

Experimental Design

Drosophila Schneider line 2 (S2) cells were used for all experimental procedures, because they are an effective male embryonic stem cell line that fully expresses the MSL complex.

The first objective was to perform a chromatin immunoprecipitation followed by DNA sequence (ChIP-seq) analysis in S2 cells in order to visualize the localization of topoisomerase I across the endogenous chromatin. This would help determine the level of association of topoisomerase I with known dosage compensated genes. In order to perform the ChIP, an antigen was produced to obtain an antibody that could efficiently bind a portion of topoisomerase I in its native conformation. Another objective was to perform a preliminary interaction study between components of the MSL complex and topoisomerase I through the use of immunoprecipitation and western blotting.

A third objective was a topoisomerase I knockdown performed in S2 cells by RNA interference (RNAi) treatment to determine its effect on dosage compensation in our lab's plasmid model system as well as endogenous *Drosophila* genes, analyzed through luciferase assay and qRT-PCR, respectively.

Materials and Methods

S2 Cells & Culturing

The S2 cells used for this experiment were grown in HyQ SFX-insect medium (HyClone) with a penicillin-streptomycin antibiotic mixture at 25°C without CO2.

RNA interference (RNAi)

1-2 hours after transferring the cells to the culture dish, they were treated with 10 ug/mL of double stranded RNA (dsRNA). Cells were treated with either topo I dsRNA or GFP dsRNA as a control. The primers used to make the double strand RNA are: GFP forward: 5' ACGTAAACGGCCACAAGTTC 3' reverse: 5' TGCTCAGGTAGT GGTTGTCG 3', topo Ia forward: 5' GCCCTTTACTTCATCGACAA 3' reverse: 5' GCCCTTTACTTCATCGACAA 3'. All dsRNA was synthesized using MEGAscriptT7 kit by Ambion.

Plasmids

Plasmids used were: ptTA, copia-Renilla luciferase I, pBluescript (pBS) by Stratagene, and a plasmid containing a firefly luciferase gene; either roX2-FF (roX2) or Nesprin-FF (Nesprin). Both the roX2 and Nesprin containing plasmids also contain the tetracycline resistance operator (tetO) inserted upstream of the firefly luciferase gene of the pGL3-Basic plasmid (Promega). The roX2 plasmid contains a 1,087-base pair fragment of the Drosophila roX2 gene (nucleotides 158-1244 of GenBank sequence

U85981) inserted downstream of the firefly luciferase gene, in the BamHI unique pGL3-Basic site. The Nesprin plasmid is the FF plasmid with 1,140 base pairs from the Nesprin human intron (nucleotides 99300 to 100440 of the emb AL359235 sequence) inserted in the BamHI unique pGL3-Basic site to replace the roX2 gene sequence. The ptTA plasmid expresses the transcriptional activator tTA encoding the tetracycline repressor protein (TetR in a Tet-Off system), which induces transcription of the firefly luciferase genes to very high levels. Transcription of the tTA genes is driven by the *D. melanogaster* constitutive alpha-tubulin 1 promoter (cloned as an XhoI-EcoRI fragment) replacing the cytomegalovirus promoter in the plasmid pUHD15.1

Plasmid Transfection

Three days after treatment with dsRNA, the cells were transferred to 5 mL flasks and transfected. Transfection was carried out following the QIAGEN Effectine protocol with 1.0 ng ptTA, 5.4 ng R plasmid, 0.6 µg pBS, and 30 ng roX2 or Nesprin plasmid. The next day the cells were diluted to a final concentration of 0.6 X 10^6 cells/mL.

Luciferase Assay

Four days after transfection, the cells were collected for the luciferase assay and protein isolation. Luciferase activity was determined by using the dual luciferase reporter assay system (Promega). The firefly luciferase activity was normalized to Renilla luciferase activity from the Renilla plasmid for each sample.

qRT-PCR

RNA was isolated from the S2 cells using the Qiagen Rneasy Plus mini-kit with on-column DNA digestion, following the manufacturer's protocols. Real-time, reverse transcription-PCR was performed using the (iScript RT-PCR kit). Genes tested include: Rp49, GPDH, Spt4, Spt6, S6kii, Act5C, CG2025, and Topoisomerase I. The primers used are shown in table 1. The Pfaffl method was used to measure the effect of topoisomerase I knockdown (Pfaffl, 2001). Primer efficiencies were determined by doing standard curves.

Table 1. Target Genes and Primer Sequences Analyzed through qRT-PCR

Rp49 Fw	5' TGCTAAGCTGTCGCACAAA 3'
Rp49 Rw	5' GTTCGATCCGTAACCGATGT 3'
Gpdh Fw	5' CACCAGTTCATTCCCAACTT 3'
Gpdh Rw	5' CTTGCCTTCAGGTGACGC 3'
Top1 Fw	5' ACGAGGAATCGATCGTAGACAT 3'
Top1 Rw	5' CGTCGTCGTGATCATTGTAGTT 3'
Spt4 Fw	5' AGTGGCAAAGATTGTCCC 3'
Spt4 Rw	5' ATCGTTGACTTCTGTCCC 3'
Act5C fw	5' GTCGTCTAATCCAGAGACAC 3'
Act5C Rw	5' CCAGAGCAGCAACTTCTTCG 3'
Spt6 Fw	5' AGAATCTGGGCGTCAAAGTCG 3'
Spt6 Rw	5' CTGCTCGGCAATCTGCTCA 3'
S6kii Fw	5' ATTGCATCTGCGGTAGCATA 3'
S6kii Rw	5' GCGAAACCCAAATCGCAG 3'
CG2025 Fw	5' AATGCCAAGAACGATGCC 3'
CG2025 Rw	5' TCCACGATGAGATGCAGTTT 3'

Immunoprecipitation

Wild type S2 cells or cells expressing either a FLAG tagged MLE (FLAG-MLE) protein or FLAG tagged MSL3 (FLAG-MSL3) and HA tagged MSL2 (HA-MSL2) were used in the immunoprecipitation experiments. From a stock population, $8x10^6$ S2 cells were transferred into flasks containing 5 mL of HyQ SFX-insect medium. Hygromycin antibiotic was added to the stable cell lines expressing FLAG-MSL3 and HA-MSL2. After 2 -3 hours, the cells were treated with copper sulfate (metallothionein promoter), if needed, to induce the expression of the tagged proteins and incubated for 3 days. The cells were then collected for protein extraction and immunoprecipitation. Cells were collected and washed with 1x PBS. The cells were incubated in lysis buffer for two hours on ice. The lysis buffer contained 150mM NaCl, 50mM Tris HCL(7.5pH), 1% Triton, and 1% protease inhibitor. A designated number of samples were treated with Rnase at 100μg/mL for one hour at room temperature. For the anti-Topo I immunoprecipitations 2 μl of serum or generic IGG were incubated with the extracts over night then protein G/agarose beads (Millipore) were used. For anti-FLAG immunoprecipitations the α-FLAG M2 affinity gel beads (Sigma) were used. Beads were pre-equilibrated in lysis buffer and then incubated with the extracts for at least two hours at 4°C. The samples were centrifuged, and the beads were washed in twenty-minute intervals, four times, with 1% PBS-Tween solution. After the washes and final centrifuge, 30 μL of loading buffer containing β-mercaptoethanol was added to each sample of beads to elute the immunoprecipitated proteins. These samples were stored at -20°C for western blot analysis.

Western Blot

Samples were prepared for SDS-Polyacrylamide Gel Electrophoresis by adding loading buffer containing β-mercaptoethanol to all samples. Samples were loaded into a 7.5% polyacrylamide precast gel (BIO-RAD) and run in a 1X Tris-Glycine buffer solution. The proteins were transferred to a PVDF membrane in a 10% methanol TG buffer solution. The membrane was blocked in PBS buffer with 0.1% tween-20 and 5% powdered milk. The primary antibodies used were: anti-TopoI 1:3000, anti-HA 1:2000, anti-lamin 1:3000, and anti-flag 1:1000. The antisera from the mice after antigen injection were used at a concentration of 1:500. The secondary antibodies used were anti-rabbit HRP 1:10000 and anti-mouse HRP 1:10000. Filter washes and antibody dilutions were performed in PBS-Tween. Enhanced chemiluminescence (ECL-Plus; GE healthcare) was used to develop western blots.

Coomassie Staining

The membrane is submerged in premade coomassie brilliant blue R-250 staining solution for five minutes. The membrane is then submerged in destaining. The destaining solution consisted of 50% methanol, 40% water, and 10% acetic acid (pH=7.5) The membrane is rinsed and air-dried for analysis.

Topoisomerase I Antigen Synthesis

The selected topoisomerase I fragment was amplified by qRT-PCR from S2 cells RNA, after purification the fragment was used for a second PCR amplification with primers containing either a BamHI or a NotI restriction site. The primers used are listed

in the table 2. A pET-30a-c(+) plasmid and the PCR fragment were digested with BamHI and NotI and then ligated. BL21 *E. Coli* cells were transformed with the obtained plasmid and raised in standard LB medium at 37°C. Induction was performed with 2mM IPTG for 3 hours at 37°C. The bacteria were lysed in 400mM Nacl, 100mM KCl, 10mM imidazole, 50 mM Potassium phosphate pH 7.8, 0.5% Triton-X-100, 10% glycerol and protease inhibitor cocktail (Roche) and the insoluble pellet was soaked for 24 hours in lysis buffer plus 10% sarcosyl.

Table 2. C-Terminal Topo I Cloning Primer Sequences

top1 C term Fw	5' TTCGGATTCTGCATGATTGA 3'
top1 C term Rw	5' ATATGAACGGCCCACAGAAA 3'
top1 C term BamHI Fw	5' AATTCTAGGATCCTTCGGATTCTGCATGATTGA 3'
top1 C term NotI Rw	5' AATTCTAGCGGCCGCAGATATGAACGGCCCACAG AAA 3'

Antigenicity Profile

The online website, http://tools.immuneepitope.org/main/, was used to analyze the antigenicity of the topoisomerase I amino acid sequence.

Data & Results

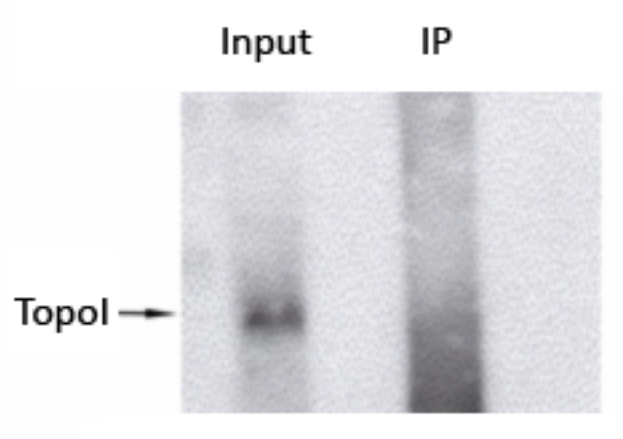

Figure 4. Western Blot of Topo I Immunoprecipitation.

A polyclonal Topoisomerase I antibody, kindly provided by Dr. Tao-shih Hsieh, was used to perform an immunoprecipitation of S2 cells extracts and the following western blot analysis (Fig 4). The antibody recognizes a specific band of the right size in the input while a similar band is not present in the immunoprecipitated fraction. Thus, this antibody interacts specifically with the denatured protein but does not efficiently recognize topoisomerase I in its native conformation and therefore cannot be used for ChIP-seq analysis. Therefore, a new, more efficient polyclonal topoisomerase I antibody was needed in order to perform a ChIP-seq analysis.

Topoisomerase Antibody Induction

In order to obtain a new topoisomerase I antibody we decided to express the protein in bacteria. Topoisomerase I is approximately 130 kDa, thus it is difficult to purify. We therefore planned to target the N-terminal and the C-terminal portions of the

protein producing two antigens of about 50 kDa. We skipped the first 80 amino acids because they are hydrophobic thus not good candidates for antibody targeting, we then designed a first antigen spanning amino acids 84-494 and a second antigen spanning amino acids 563-966 as represented in Figure 5. However, the analysis of the protein antigenicity profile with the Kotaskar and Tougaonkar method revealed that the overall N-terminal portion antigenicity was rather weak while the C-terminal peptide was predicted to be highly immunogenic (Fig 6). Based on the antigenicity profile, we decided to initially focus on the C-terminal portion and to eventually synthesize short peptides targeting the N-terminal portion in a second moment.

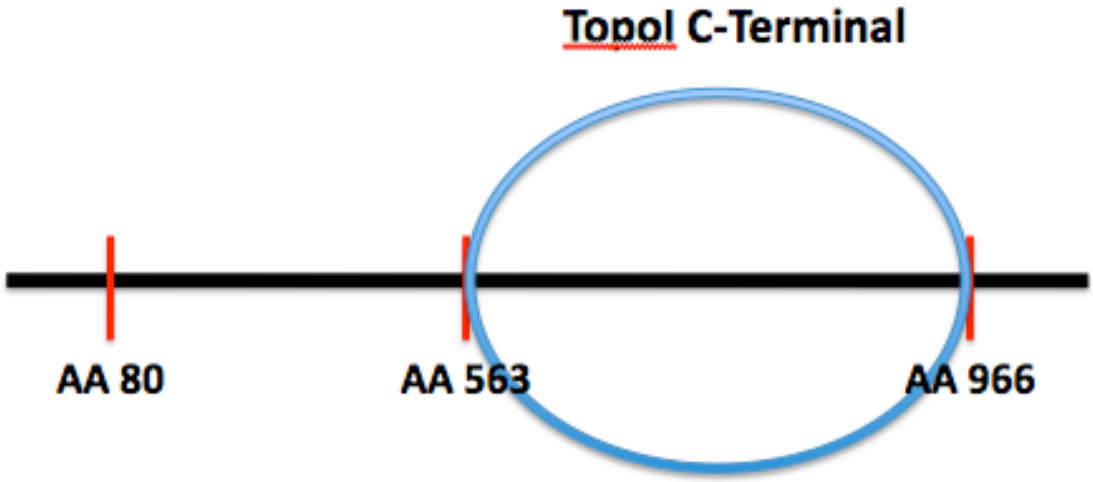

Figure 5. Topoisomerase I Sequence Targets for Antibody Synthesis. Target portion includes amino acid 563 to 966.

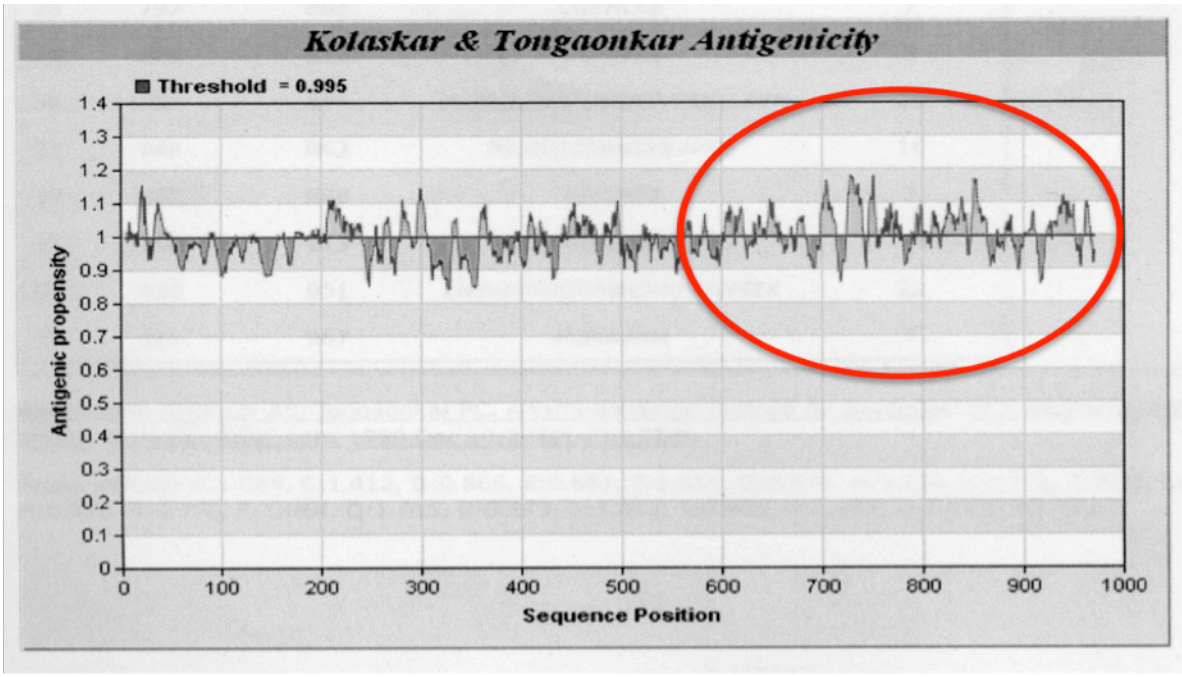

Figure 6. Antigenicity Profile of Topoisomerase I. Antigenic propensity values greater than 1 represent highly immunogenic potential.

For the bacterial expression we used a pET-30a-c(+) expression vector (Fig 7) which contains a strong-viral T7 promoter, a lac operon sequence for Isopropyl β-D-1-thiogalactopyranoside (IPTG) induction and two histidine tags (His-Tag) for protein purification. A fragment coding for the selected region of topoisomerase I was amplified by qRT-PCR using S2 cells RNA and cloned in the expression vector in frame with the N-terminal His-tag.

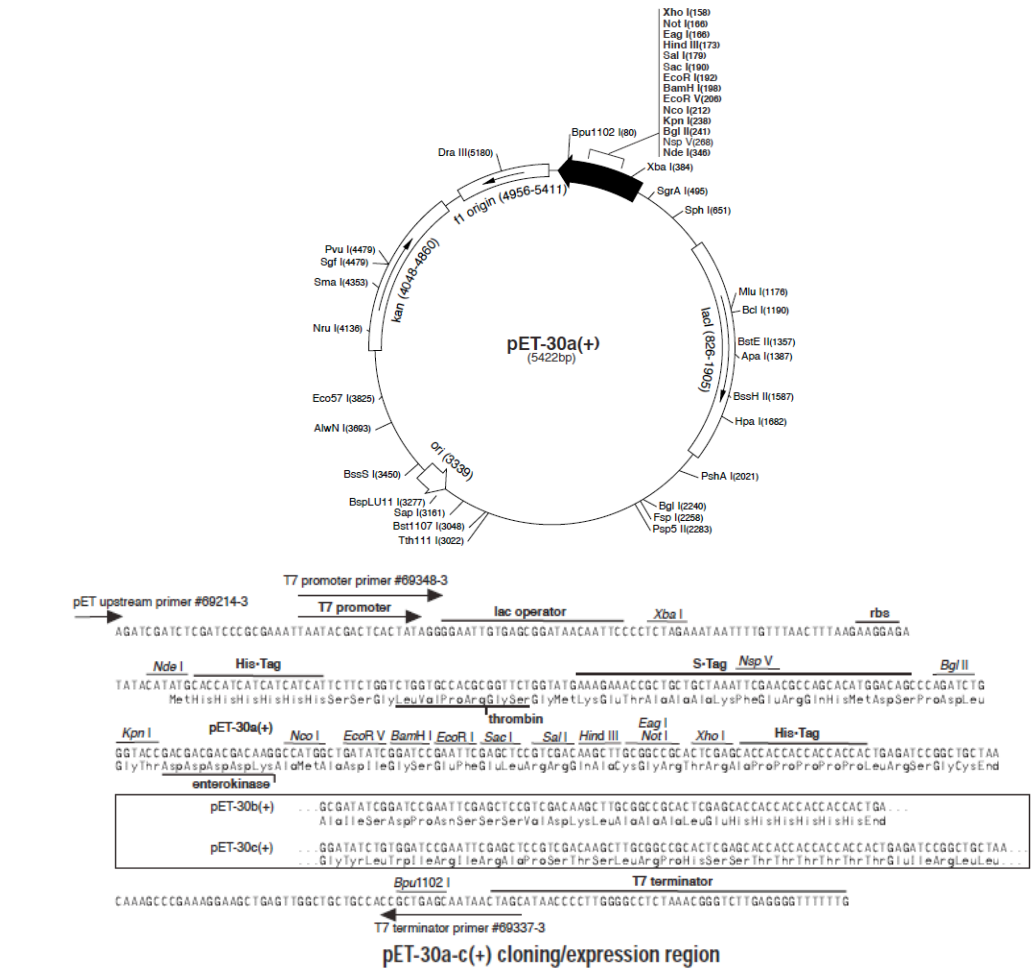

Figure 7. PET-30a-c(+) Vector Map and Cloning/Expression Region

Protein Isolation and Antibody Induction

Protein extracts from bacteria transformed with the expression vector and induced with IPTG were separated in a soluble fraction (SN) and an insoluble fraction (pellet) and loaded in a protein gel to evaluate the expression (Fig 8), a sample of bacteria transformed but not induced with IPTG was loaded as a control. The coomassie staining of the gel shows that a protein that runs at the predicted size of our target protein is present in the induced samples while one is absent in the non-induced sample. Therefore, the IPTG induction leads to an adequate expression of our protein, however it is mainly

concentrated in the pellet which means that it is highly expressed and requires solubilization prior the isolation step.

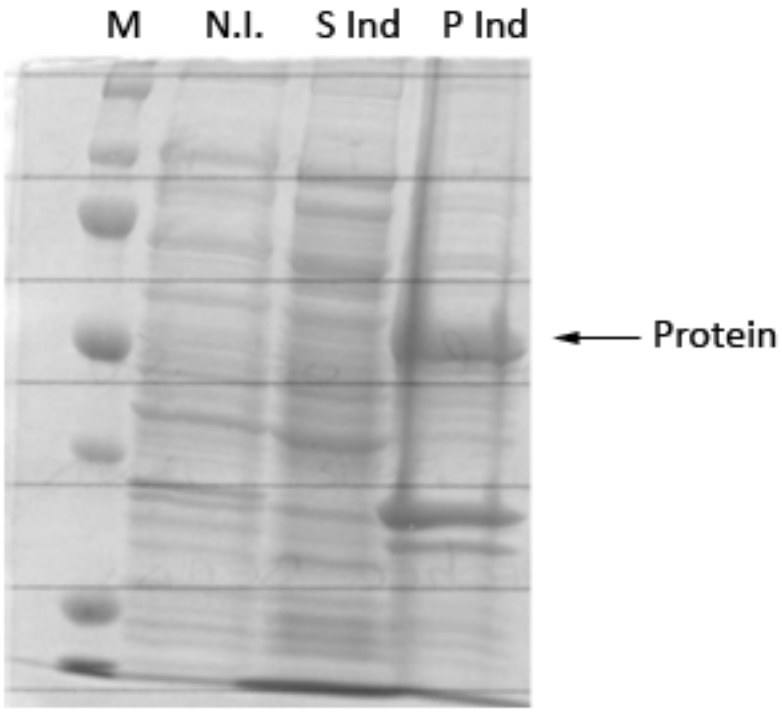

Figure 8. Coomassie Stain of Protein Induction Gel. Loaded as marker (M), non-induced sample (N.I.), induced supernatant (S Ind), and induced pellet fraction (P Ind).

The pellet fraction, containing the protein of interest, was soaked in a buffer with a high concentration of sarcosyl detergent in order to disrupt inclusion body membranes. The supernatant from a subsequent centrifugation was incubated with Ni2+ beads, which have a high affinity for histidine. As shown in Figure 9., the protein was successfully solubilized and efficiently purified.

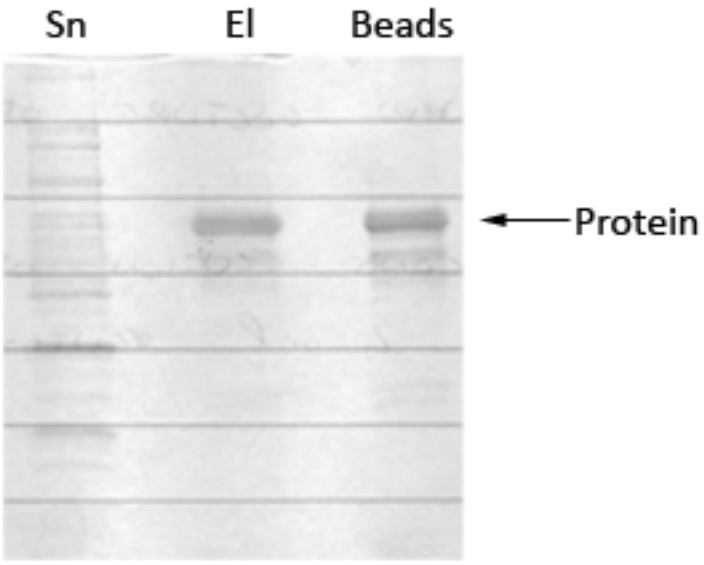

Figure 9. Coomassie Stain of Protein Solubilization Gel. Loaded as supernatant unbound to Ni2+ beads (Sn), elution fraction from Ni2+ beads (El), and Ni2+ bead fraction (Beads).

As shown in figure 10., the isolated protein was compared to known concentrations of an albumin standard in order to quantify our protein for injection.

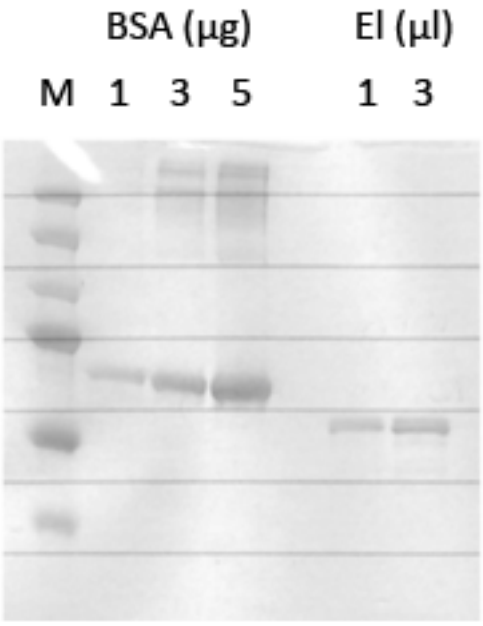

Figure 10. Coomassie Stain of Protein Quantification Gel. Loaded as marker (M), Bovine serum albumin (BSA), and elution containing the synthesized antigen (El).

The purified antigens were then used for mice immunization. The efficiency of the antisera was tested by western blot. As shown in figure 11., all of the five injected mice yielded antibodies that efficiently recognized a protein that runs at the size expected of topoisomerase I. Furthermore, a similar band is absent in the preimmune serum, which is the serum of the mice before being injected with our antigen. We also tested the specificity of the detected band by western blot analysis of samples knocked down for Topo I (Fig 12). A band of the correct size is detected in the cells treated with GFP dsRNA but not in the topoisomerase I dsRNA treated sample. Therefore, the antisera obtained specifically recognizes topoisomerase I. As shown in figure 13., the induced antibody of mice one, three, and four were used for an immunoprecipitation of topoisomerase I in S2 cells. The western blot analysis clearly indicates that these antibodies efficiently bind topoisomerase I in its the native conformation, allowing their usage for detection of topoisomerase I across the endogenous chromatin in a ChIP-seq experiment.

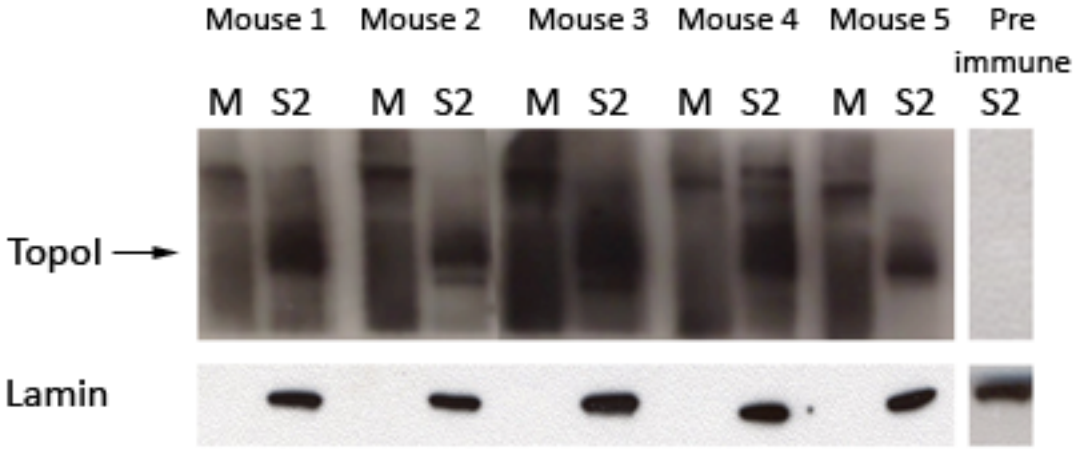

Figure 11. Western Blot Detected with Mice Antisera. Loaded as repeated marker and S2 cell extract. Separately blotted with antisera from mice one through five. Lamin Control.

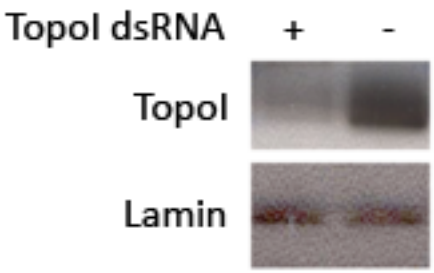

Figure 12. Western Blot of Topoisomerase I Knockdown with Mice Antisera

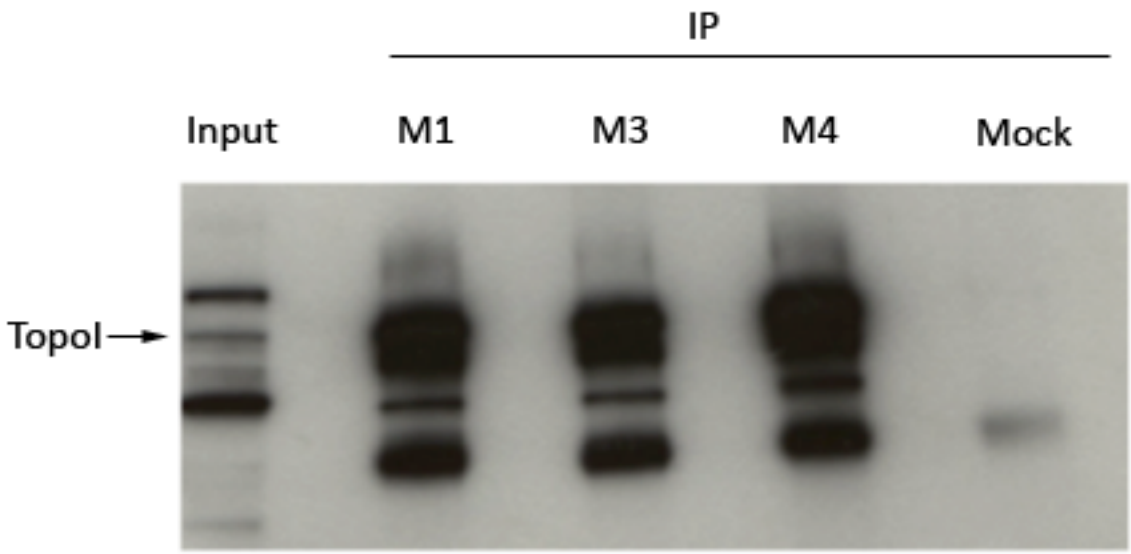

Figure 13. Western Blot of Immunoprecipitation with Mouse Antisera

Topoisomerase I Interaction with MSL Components

To investigate the possibility of an interaction between topoisomerase I and the MSL complex, MLE was immunoprecipitated from cells expressing a FLAG-tagged MLE form and topoisomerase I antibody was used for western blot analysis (Fig 14). The topoisomerase I signals are relatively constant across the wild type input and the RNase treated input. The RNase treated FLAG-MLE immunoprecipitation sample had a weak topoisomerase I signal relative to the immunoprecipitation not treated with RNase. The feint signal in the RNase treated immunoprecipitation could either be a result of a small

amount of topoisomerase I present or a background signal of MLE. This is due to the fact that MLE and topoisomerase I run the same distance during gel electrophoresis. A coomassie stain was performed and revealed a strong band located exactly where the signal is present on the western blot in the RNase treated immunoprecipitation sample. Because the shape and height are similar, it strongly suggests that the feint signal seen in the western blot is due to MLE background. This also suggests that topoisomerase I interacts with MLE in a RNA dependent fashion.

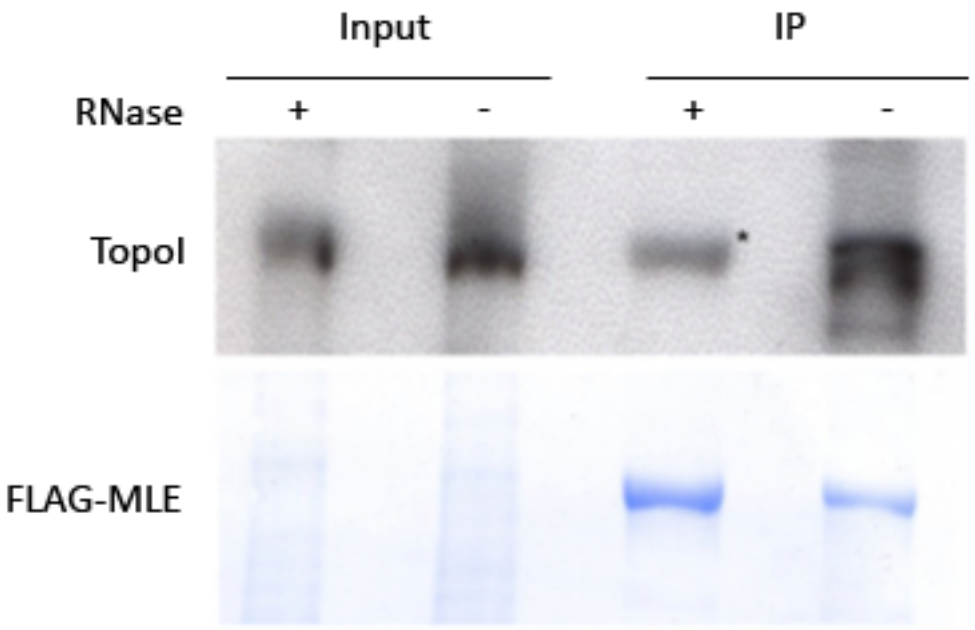

Figure 14. Western Blot of MLE IP & Coomassie Staining of MLE

In order to understand if the Topo I interaction with MLE is in the context of the MSL complex, we performed an immunoprecipitation with another component of the complex, MSL3, and tested for the binding of Topo I. As shown in Figure 15., topoisomerase I interacts with MSL3 and its signal is stronger in the wild type than in the RNase treated sample in the shorter and longer exposures. Thus, topoisomerase I associate with MSL3 in a RNA dependent manner. This is further supported by the fact that, according to the FLAG MSL3 signal, more RNase treated sample was loaded, but we see a drastic difference in signal relative to the non-RNase treated sample. Therefore, topoisomerase I does associate with parts of the MSL complex. Specifically, topoisomerase I interacts with MLE and MSL3 in a RNA dependent manner.

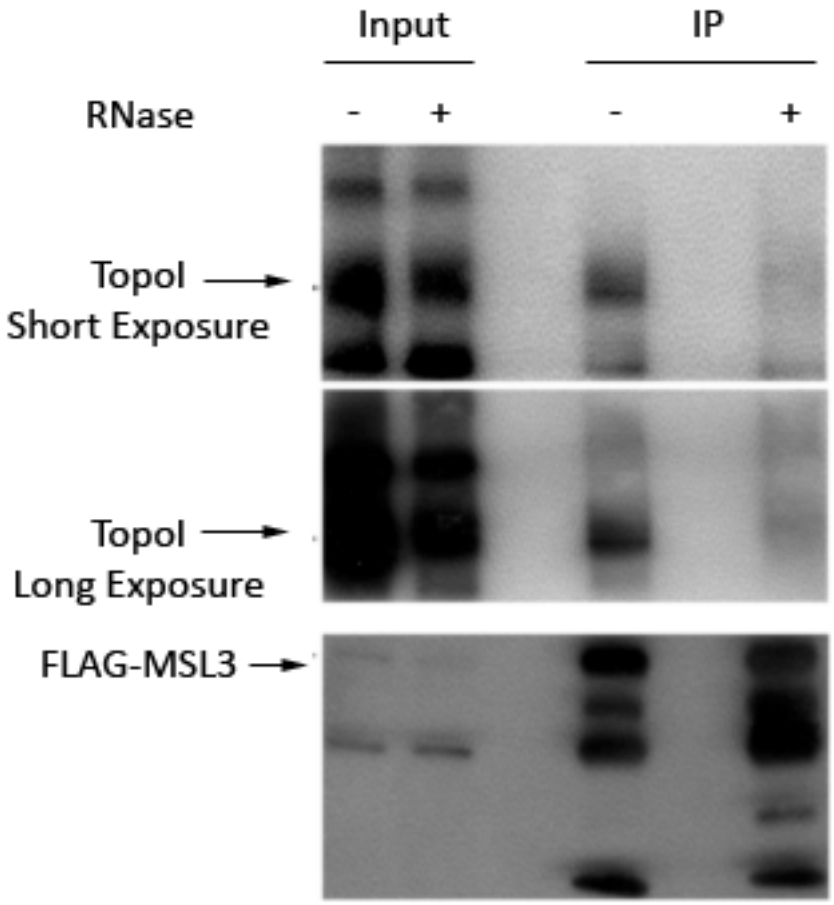

Figure 15. Western Blot of FLAG-MSL3 with Topoisomerase I Detection

Topoisomerase I Knockdown and Dosage Compensation

To explore the potential role of Topoisomerase I in dosage compensation we tested the effect of topoisomerase I knock down on a plasmid system that mimics dosage compensation. Confirmation of RNAi knockdown was done through western blot as shown in Figure 16 and through qRT-PCR, as shown in Figure 17.

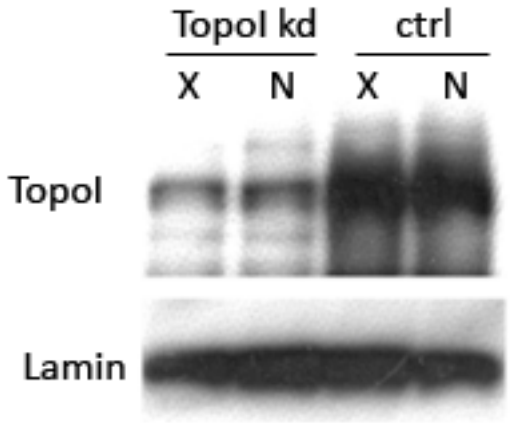

Figure 16. Western Blot of Topoisomerase I Knockdown Confirmation.

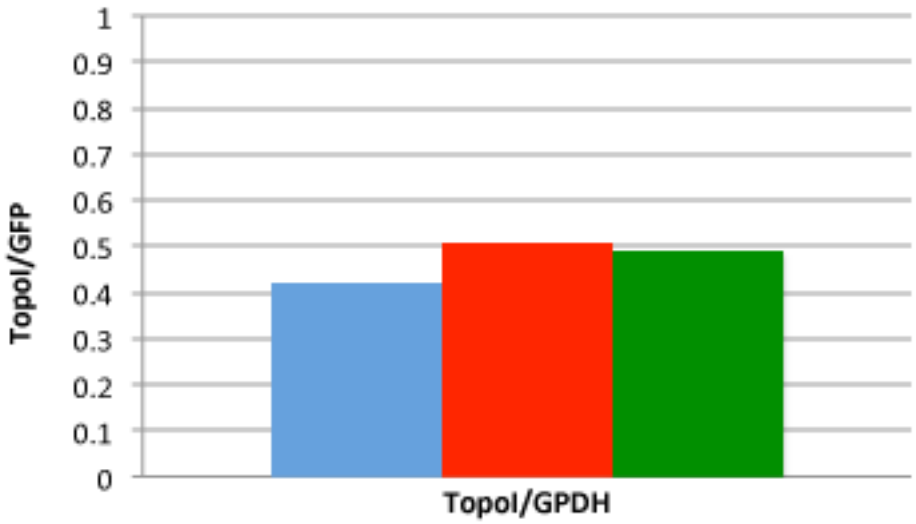

Figure 17. qRT-PCR of Topoisomearse I Knockdown Normalized to GPDH

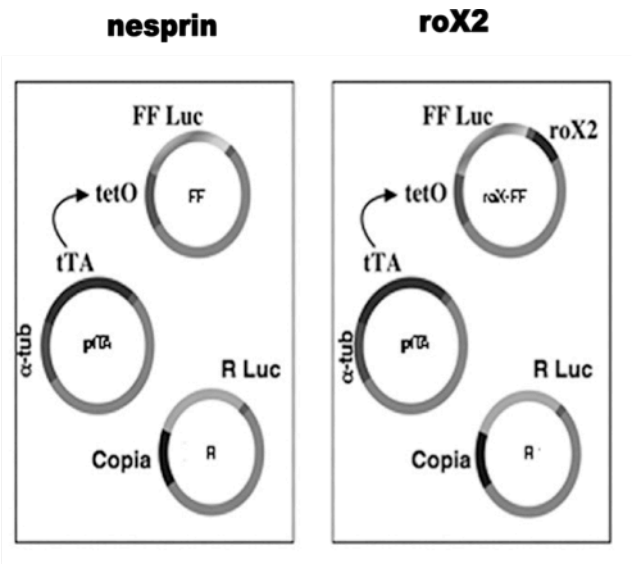

Figure 18. Plasmid model of dosage compensation. Plasmids transfected in Nesprin (control) cells: FF (Nesprin- FF), ptTA, copia-Renilla luciferase. OR Plasmids transfected in roX2 model of dosage compensation: roX2-FF, ptTA, copia-Renilla luciferase (Yokoyama et al., 2007).

Plasmid Model System and Luciferase Assay

In 2007, the Lucchesi lab developed the plasmid model system utilized in this project. This plasmid system facilitates the study of the mechanism underlying dosage compensation. Cells are transfected with one of two sets of plasmids. The first plasmid set contains the Renilla luciferase gene-containing "R plasmid", the tTA gene containing "ptTA plasmid", and the reporter gene firefly luciferase gene-containing "Rox2 plasmid". The R plasmid is used to normalize the firefly signal for the number of cells successfully transfected with one of the two plasmid sets. The ptTA plasmid produces tTA, which induces the inducible promoter on the firefly containing plasmid. The Rox2 plasmid reproduces dosage compensated levels of Firefly luciferase, because a fragment of the rox2 gene is immediately downstream of the reporter gene firefly luciferase. The presence of this fragment recruits the endogenous MSL complex onto the plasmid and allows the complex to up regulate transcription of the plasmid by twofold. The second plasmid set used also contain the R plasmid and ptTA plasmid, but rather than the Rox2 plasmid, this set contains the control Nesprin plasmid. The Nesprin plasmid also contains the firefly luciferase reporter gene, but an intron of the human

Nesprin gene is downstream rather than a rox2 gene fragment. Consequently, the Nesprin plasmid is not dosage compensated, because the rox2 sequence, critical for recruitment of the MSL complex is replaced by a trivial DNA sequence of the same length. The relative luciferase activity in topoisomerase I knockdown versus the control GFP treated S2 cells were measured in order to determine if differences in dosage compensation occurred after topoisomerase I RNAi treatment. This would reveal a role, if any, of topoisomerase I in the dosage compensation mechanism of *Drosophila melanogaster*.

Depicted in table 3. are the relative firefly signal against the Renilla signal of the same sample. This ratio is reported for four experimental groups that include S2 cells treated with topoisomerase I dsRNA transfected with roX2 (TX), topoisomerase I dsRNA transfected with Nesprin (TN), GFP dsRNA transfected with roX2 (GX), or GFP dsRNA transfected with Nesprin (GN). Values for four different experiments containing these same experimental groups were taken and are shown in the FF/Ren column. The ratio of roX2 transfected versus those transfected with Nesprin of the same dsRNA treatment are shown in the roX/Nesp column. As shown, very little to no difference in the ratio of roX2 over Nesprin transfected is measured following treatment with topoisomerase I or with GFP dsRNAs. This relationship was observed in all four experiments. The slightly elevated figures in the roX transfected versus Nesprin FF/Ren value may be due to inconsistencies in plasmid transfection, or an unequal amount of roX2 and Nesprin plasmids were transfected. The final column, titled TopoI/GFP, is provided to highlight the lack of an effect on dosage compensation in this plasmid model system with topoisomerase I dsRNA treatment. However, there is an increase in absolute values of firefly and Renilla for the topoisomerase I dsRNA treated samples relative to the GFP

dsRNA treated samples. As shown in Figure 19., the same increase in transcription is observed in the non-dosage compensated plasmid. Therefore, a global increase in transcription is occurring consistently with topoisomerase I dsRNA treatment.

Table 3. Luciferase Assay Data for Four Experiments. Relative FF/Ren, roX/Nesp, and TopoI/GFP values reported.

	Sample	FF/Ren	roX/Nesp	TopoI/GFP
1	TX	11.35	2.23	.9955
	TN	5.093		
	GX	7.36	2.24	
	GN	3.26		
2	TX	8.52	2.63	1.044
	TN	3.24		
	GX	5.3	2.52	
	GN	2.1		
3	TX	2.1	3.28	1.055
	TN	.64		
	GX	1.09	3.11	
	GN	.35		
4	TX	2.21	2.8	.8889
	TN	.79		
	GX	1.04	3.15	
	GN	.33		

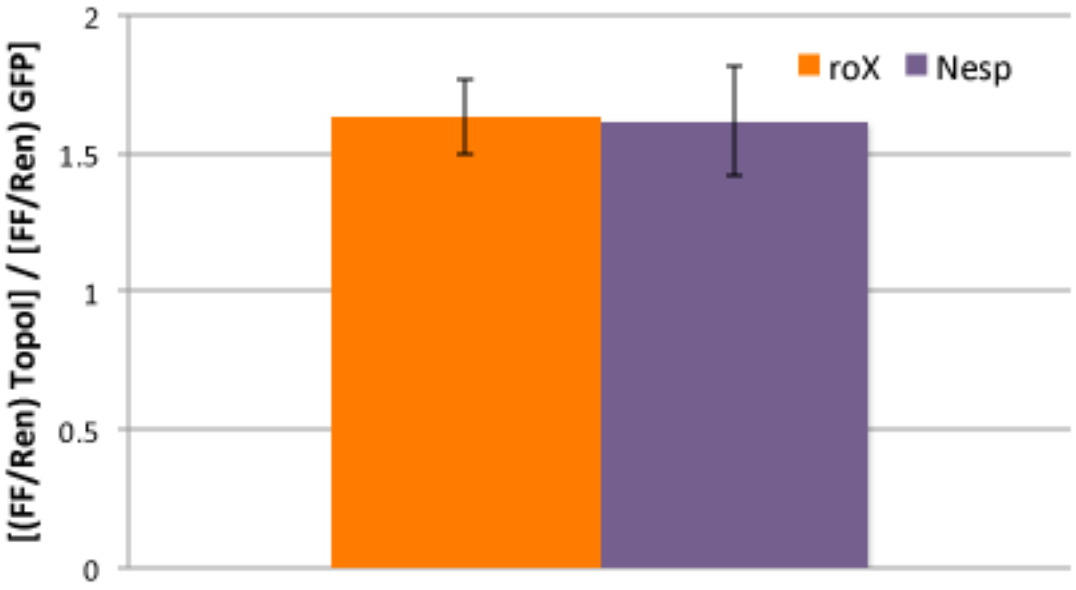

Figure 19. FF/Ren Expression of all TopoI RNA treated versus GFP RNAi treated. Both RNAi treatment samples either transfected with roX2 or Nesprin.

Effect of Topoisomerase Knockdown on Endogenous X-Linked Genes

We proceeded further analyzing the impact of topoisomerase I knock down on endogenous genes using qRT-PCR. To analyze the data we used the Pfaffl method, which consists in the calculation of the ΔCt adjusted for the primer efficiency. Target genes were normalized against the housekeeping genes, GPDH and RP49.The relative expressions for autosomal and X-linked genes were compared to determine if topoisomerase I knockdown had any effect on normally dosage compensated genes compared to non-dosage compensated. As shown in Figure 20., we see that there is no difference in gene expression following topo I RNAi. In fact, for each gene tested the ratio of the expression values in Topo I knock down versus GFP is near 1 for both X-linked and autosomal genes. Therefore, topoisomerase I knockdown does not seem to have an effect on dosage compensated or non-dosage compensated genes. In addition, the

universal increase in transcription observed in the plasmid system is not seen in the tested endogenous genes.

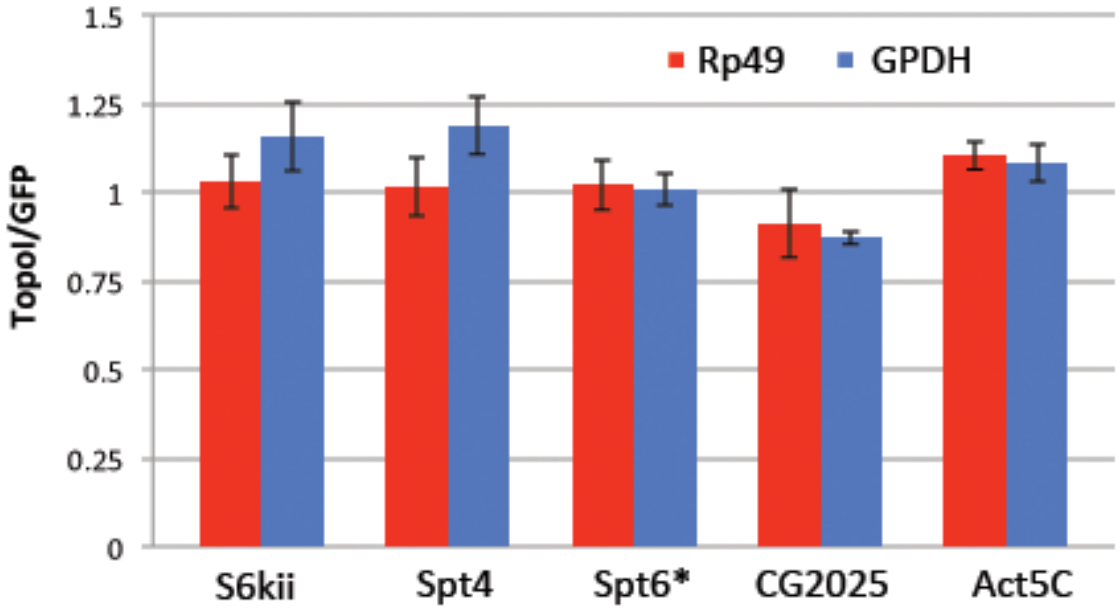

Figure 20. qRT-PCR Data for All Target Genes Normalized to RP49 and GPDH. All genes, except for Spt6, are an average of three independent samples. Spt6 is an average of two independent samples.

Discussion

In light of topoisomerase II's role in the process of dosage compensation in *Drosophila melanogaster*, as well as topoisomerase I's ability to relieve transcription induced negative torsional strain, a role for topoisomerase I in Drosophila dosage compensation warranted investigation. However, based on the results of both the dosage compensation plasmid model and qRT-PCR of endogenous genes, topoisomerase I does not seem to play a role in the mechanism of dosage compensation beyond the role that it likely plays in general transcription. Surprisingly, topoisomerase I's association with components of the MSL complex mirrors that of topoisomerase II's RNA-dependant association with components of the MSL complex. Therefore, the significance of topoisomerase I's association with MLE and MSL3 remains to be investigated.

The global increase in transcription in the plasmid system following topoisomerase I knockdown may be explained by the topological model of transcription proposed by the model of Liu and Wang (1987). If topoisomerase I preferentially restores the negative supercoils that occur behind the elongating polymerase to pre-transcription levels, its absence would leave the DNA of genes more negatively supercoiled, which might facilitate the successive rounds of transcription and elongation. The fact that this increase in overall transcription is not evident in endogenous genes may reflect the difference in organization between a circular plasmid and linear genes whose DNA is presumably anchored in insulator sites.

Future Work

While previous work in the lab had demonstrated the absence of dosage compensation in the plasmid system when topoisomerase II was knocked down, a comparable effect was not seen with endogenous genes. Knockdown of topoisomerase I also did not show an effect on the dosage compensation of endogenous genes. It is possible that on endogenous genes, the absence of one topoisomerase is made up by the presence of the other. Therefore, the effect of a double knockdown of both enzymes on dosage compensation should be investigated. In addition, as a result of this work's antibody synthesis, a chromatin immunoprecipitation followed by DNA sequencing can now be efficiently performed. Finally, using the newly obtained antibody, further co-immunoprecipitations could further elucidate the association of the topoisomerase enzymes with other components of the MSL complex.

References

1. Charlesworth B (1996) The evolution of chromosomal sex determination and dosage compensation. *Curr. Biol*. **6**, pp.149-162.
2. Conrad T, Cavalli FMG, Vaquerizas JM, Luscombe NM, Akhtar A (2012) Drosophila Dosage Compensation Involves Enhanced Pol II Recruitment to Male X-Linked Promoters. *Science* **337**, pp. 742-746.
3. Copps K, Richman R, Lyman LM, Chang KA, Rampersadammons J, Kuroda MI (1998) Complex formation by the *Drosophila* MSL proteins: role of the MSL2 RING finger in protein complex assembly. *EMBO J* **17**, pp.5409-17.
4. Franke A, Baker BS (1999) The rox1 and rox2 RNAs are essential components of the compensasome, which mediates dosage compensation in *Drosophila*. *Mol Cell* **4**, pp.117-22.
5. French SL, Sikes ML, Hontz RD, Osheim YN, Lambert TE, Hage AE, Smith MM, Tollervey D, Smith JS, Beyer AL (2010) Distinguishing the Roles of Topoisomerases I and II in Relief of Transcription-Induced Torsional Stress in Yeast rRNA Genes. *Mol Cell Biol.* **31**, pp.482-494.
6. Gelbart ME, Kuroda MI (2009) *Drosophila* dosage compensation: a complex voyage to the X chromosome. *Development* **136**, pp. 1399-1410.
7. Hallacli E, Akhtar A (2009) X chromosomal regulation in flies: when less is more. *Chromosome Research*. **17**, pp.603-619.
8. Larschan E, Bishop EP, Kharckenko PV, Core L, Lis JT, Park PJ, Kuroda MI (2011) X Chromosome dosage compensation via enhanced transcriptional elongation in *Drosophila*. *Nature* **471**. pp.115-118.
9. Li F, Parry DA, Scott MJ (2005) The amino-terminal region of *Drosophila* MSL1 contains basic, glycine-rich, and leucine zipper-like motifs that promote X chromosome binding, self-association, and MSL2 binding, respectively. *Mol Cell Biol*. **25**, pp.8913-24.
10. Liu LF, Wang JC (1987) Supercoiling of the DNA template during transcription. *Proc. Natl. Acad. Sci.* **84,** pp. 7024-7027.
11. Lodish H, Berk A, Kaiser CA, Krieger M, Scott MP, Bretscher A, Ploegh H, Matsudaira P (2008) *Molecular Cell Biology* (6th ed.). New York: W. H. Freeman.
12. Pfaffl MW (2001) A new mathematical model for relative quantification in real-time RT-PCR. *Nucleic Acid Res.* **29,** pp.2002-2007.
13. Scott MJ, Pan LL, Cleland SB, Knox AL, Heinrich J (2000) MSL1 plays a central role in assembly of the MSL complex, essential for dosage compensation in *Drosophila. EMBO J* **19**, pp.144-55.
14. Smith ER, Allis CD, Lucchesi JC (2001) Linking Global histone acetylation to the transcription enhancement of X-chromosomal genes in *Drosophila* males. *J. Biol. Chem.* **276**, pp.31483-31486.
15. Vos SM, Tretter EM, Schmidt BH, Berger JM (2011) All tangled up: how cells direct, manage and exploit topoisomerase function. *Mol Cell Biol.* **12**, pp.827-841.
16. Wu L, Zee BM, Wang Y, Garcia BA, Dou Y (2011) The RING finger protein MSL2 in the MOF complex is an E3 ubiquitin ligase for H2B K34 and is involved in crosstalk with H3 K4 and K79 methylation. *Mol Cell Biol.* **43**, pp.132-144
17. Yokoyama R, Pannuti A, Ling, H, Smith ER, Lucchesi JC (2007) A plasmid

model system shows that *Drosophila* dosage compensation depends on the global acetylation of histone H4 at lysine 16 and is not affected by depletion of common transcription elongation chromatin marks. *Molecular and Cellular Biology* **27 (22)**, pp.7865-7870.