Thioredoxins and gene regulation in the *Drosophila* germline

av

Malin J. Svensson

Akademisk avhandling

Som med vederbörligt tillstånd av rektorsämbetet vid Umeå Universitet för avläggande av filosofie doktorsexamen i molekylär biologi framlägges till offentlig granskning i sal B, byggnad 1D, 9 tr, fredagen den 9:e mars, kl. 10.00.

Avhandlingen kommer att försvaras på engelska.

Examinator: Prof. Elisabeth Sauer-Eriksson

Fakultetsopponent: Prof. emeritus David Roberts, Genetics Unit, Biochemistry Department, Magdalen College, Oxford, Storbritannien
Abstract
The process of spermatogenesis is in many organisms one of the most dramatic cellular transformations - a normal round diploid cell is ultimately transformed into a needle shaped haploid cell with tightly packaged cell machinery. In *Drosophila melanogaster* this process involves several characteristic stages, one of these being the primary spermatocyte stage, which is the stage prior to meiosis. This stage is characterized by a loose chromatin structure in the nucleus and exceptionally high rates of transcription and translation to produce essentially all the mRNAs and proteins that are needed later during spermatid formation. Two proteins that are expressed in high levels in primary spermatocytes are ThioredoxinT (TrxT) and Painting of fourth (POF).

Thioredoxins are small thiol proteins that reduce disulfides in other proteins, a mechanism that is utilized in many different contexts. In this thesis I show that the *TrxT* gene encodes a testis-specific thioredoxin that specifically associates to Y-chromosome loops in primary spermatocytes. *TrxT* is located right next to *deadhead (dhd)*, a gene that encodes a female-specific thioredoxin that specifically locates to nuclei in the ovaries. A third thioredoxin in *Drosophila* is the ubiquitously expressed Thioredoxin-2 (Trx-2). I have found that flies lacking Trx-2 are viable but have shorter life spans than wild type flies, while over-expression of Trx-2 mediates an increased resistance to oxidative stress. Interestingly, a lack of all three thioredoxins does not result in lethality, contrary to what could be expected. All three thioredoxins are conserved among *Drosophilidae* and the striking genomic organization of *TrxT* and *dhd* is generally conserved.

The gene name *Painting of fourth* originates from the finding that POF stains the 4th chromosome of *Drosophila* in a banded pattern on salivary gland chromosomes. I show in this thesis that POF binding to the equivalent of the 4th chromosome is conserved in genus *Drosophila* and that POF co-localizes with both a protein and a histone modification associated with dosage compensation in species where POF also binds the male X. POF is expressed ubiquitously in both males and females, but at very high levels in male testes. I show that POF is present in nuclei of primary spermatocytes, but also in nuclei of maturing spermatids and that a lack of POF in the male germline causes a global down-regulation of chromosome 4 genes. These results combined suggest a function of POF in the first known case of chromosome-wide gene regulation of an autosome.

Key words: *Drosophila*, thioredoxin, sex-specific, *Pof*, chromosome 4, gene regulation

Language: English      ISBN: 978-91-7264-255-3      Number of pages: 166
Thioredoxins and gene regulation in the *Drosophila* germline

Malin J. Svensson

Umeå Centre for Molecular Pathogenesis
Umeå University
Umeå 2007
To Björn and to my family

“The most exciting phrase to hear in science, the one that heralds new discoveries, is not ‘Eureka!’ (I found it!) but ‘That’s funny ...’ “

Isaac Asimov
CONTENTS

LIST OF PAPERS ........................................................................................................... 6
ABBREVIATIONS ........................................................................................................... 7
PROLOGUE ...................................................................................................................... 8
INTRODUCTION ............................................................................................................ 9

THIOREDOXINS ........................................................................................................... 9
   Biochemistry ............................................................................................................. 10
   Physiological roles .................................................................................................. 11
      Thioredoxins in different cellular compartments .................................................. 11
      Thioredoxins and disease ..................................................................................... 12
      Thioredoxins in the mammalian germline .............................................................. 13
   Thioredoxins in Drosophila .................................................................................... 14
      deadhead ............................................................................................................... 14
      Thioredoxin-2 ...................................................................................................... 15
      ThioredoxinT ......................................................................................................... 16

GENE REGULATION .................................................................................................... 17
   Cellular memory .................................................................................................... 17
      The PcG and trxG complexes ............................................................................. 18
   Chromosome-wide gene regulation ..................................................................... 18
      Dosage compensation in mammals ................................................................ 20
      Dosage compensation in C. elegans ................................................................. 20
      Dosage compensation in Drosophila ................................................................. 21
      Painting of fourth in Drosophila ...................................................................... 23
   Gene regulation in the male germline of Drosophila .............................................. 24
      Spermatogenesis ................................................................................................. 24
      The Y chromosome ............................................................................................. 25
      Testis specific gene regulation ......................................................................... 26
AIMS ................................................................................................................................. 28
MATERIALS AND METHODS ............................................................................................ 29
METHODS USED IN THIS THESIS .................................................................................. 29
In vitro generation of zeste mutants .................................................................................. 29
Protein aggregation experiments ...................................................................................... 29
RESULTS AND DISCUSSION .......................................................................................... 30
PAPERS I-V ....................................................................................................................... 30
TrxT encodes a male-specific thioredoxin ........................................................................ 30
TrxT locates to Y-chromosome loops in primary spermatocytes .................................... 30
Trx-2 mutants have shorter life spans ............................................................................... 31
Thioredoxin redundancy or specialized functions? .......................................................... 32
TrxT, dhd and Trx-2 are conserved in Drosophilids ........................................................ 33
TrxT has an extremely variable C-terminal domain ......................................................... 33
Regulatory motifs in the region between TrxT and dhd .................................................. 34
Pof in genus Drosophila suggests autosomal-specific gene regulation ............................ 35
Deletion mutants in the Pof promoter region are lethal ................................................... 36
Pof in spermatogenesis ................................................................................................... 37
Pof upregulates chromosome 4 genes in the male testes ................................................. 37
NEW RESULTS ................................................................................................................. 39
Thioredoxins can break Zeste aggregates ...................................................................... 39
CONCLUSIONS .............................................................................................................. 42
SUMMARY IN SWEDISH ................................................................................................ 43
ACKNOWLEDGEMENTS ............................................................................................... 44
REFERENCES .................................................................................................................. 46
PAPER I-V
LIST OF PAPERS

This thesis is based on the following papers, which in the text will be referred to by their Roman numerals (I-V):


Papers I, II and IV are reproduced with permission from the publishers.
ABBREVIATIONS

achi/vis  achintya/vismay
aly  always early
AP-1  Activator protein 1
can  cannonball
comr  cookie monster
DCC  Dosage compensation complex
dpy  dumpy
dhd  deadhead
ECFP  Enhanced Cyan Fluorescent Protein
EYFP  Enhanced Yellow Fluorescent Protein
GR  Glutathione reductase
H4K16  Histone 4 at Lysine 16
HP1  Heterochromatin protein 1
mia  meiosis I arrest
mle  maleless
mof  males absent on first
MSL  Male specific lethal
NF-κB  Nuclear factor-κB
nht  no hitter
PcG  Polycomb Group
Pof  Painting of fourth
Ref-1  redox factor 1
ROS  reactive oxygen species
roX  RNA on the X chromosome
rye  ryan express
sa  spermatocyte arrest
sdc  sex determination and dosage compensation
Sptrx  sperm thioredoxin (mammalian)
TAF  TBP-associated factor
TBP  TATA-binding protein
TGR  thioredoxin/glutathione reductase
topi  matotopetli
TPx  Thioredoxin peroxidase
Trx-2  Thioredoxin-2
TrxG  Trithorax Group
TrxR  Thioredoxin reductase
TrxT  ThioredoxinT
Txl  Thioredoxin-like
Xa/Xi  Activated/inactivated X chromosome
Xic  X inactivation center
Xist  Xi-specific transcript
z  zeste
ZIP  Zeste Interacting Protein
It’s hard to tell how it all really started, but if Thomas Hunt Morgan hadn’t found that white eyed fly one day in 1910, this particular thesis would probably never have been written. But he did, and in only a few years the fledgling *Drosophila* research field took a large step forward. But why did these tiny fruit flies become so popular for doing research on? Well, for one, they are very convenient to cultivate. You basically only need some flies, some vials and some rotting fruit and you’re good to go. Now, rotting fruit is perhaps not the nicest thing to be around, so nowadays we use a mix of potato-mash, yeast and agar, but the principle of it is the same. Two other reasons why *Drosophila* became so popular have to do with Morgan and his fellow researchers. First, they happily shared all their fly strains with anyone that wanted them, thereby setting the standard for an entire research community. Second, a large amount of the early publications on *Drosophila* were findings and methods that set the foundation for a large spectrum of genetic tools, that even to this day are being used. Although the first Drosophilists were mostly interested in genetics and heredity, the scientific area of *Drosophila* development soon drew the interest of an increasing amount of scholars. These two fields combined eventually resulted in the discovery of a type of genes that are essential for laying down the body plan in the fly embryo. These genes were called Hox-genes and were later also discovered in mammals and other animals. Mutations in many of the Hox-genes produced peculiar phenotypes, like legs instead of antennae and four wings instead of two. To make sure that the Hox-genes are not on or off at the wrong time two other groups of genes, the Polycomb Group and the Trithorax Group, are there to maintain the active or silenced state. These in turn need regulators and interaction partners of their own, one of them being the gene zeste. Now zeste happens to be, in its own way, related to the very beginning of this story, since certain mutations in it does strange things to the same gene that caused the white phenotype found by Morgan almost one hundred years ago. But what has all of this got to do with how this thesis came to be? Well… In the late 90’s a talented young scientist had just earned his PhD and was in need of a new and interesting project to focus his research on. As it happened the Professor that had been the opponent at his dissertation was interested in a collaboration, and so our young scientist went for two month-long visits to the Professor’s lab. When he came back he brought with him a handful of newly identified *Drosophila* genes from a yeast two-hybrid screen for proteins that interacted with Zeste. The genes were referred to as ZIPs (Zeste Interacting Proteins) followed by a number, since hardly anything was known about them yet. The young scientist conducted experiments on several of the genes, but soon two of them, ZIP16 and ZIP42, stood out as particularly promising. One day a student looking for an exam-project turned up... and the rest, as they say, is history.
INTRODUCTION

Since the work this thesis is based on was divided into two separate parts, this introduction will also be divided into two parts. But as you will see in later sections of this thesis, the two parts are perhaps not as separate as they may seem at first glance.

THIOREDOXINS

The first thioredoxin (Trx) was discovered in 1964, in the bacteria *Escherichia coli*, and along with it thioredoxin reductase (TrxR), a flavoenzyme that maintains thioredoxin in its active, *i.e.* reduced, state (Laurent *et al.* 1964; Moore *et al.* 1964). Only a few years after, thioredoxin was discovered in other bacteria, yeast and rat (Orr & Vitols 1966; Moore 1967; Gonzalez Porqué *et al.* 1970). Since then thioredoxins have been reported in practically every organism studied, including plants, and in the last ten years the field has essentially exploded with publications of thioredoxins in new organisms and new biological contexts. But to understand what thioredoxins are and what they do, one first needs to know a little bit about the basic biochemistry of thioredoxins.

Figure 1. Schematic view of the thioredoxin protein structure. Alpha helices are represented by blocks and beta sheets are represented by arrows. N indicates the N-terminal and C indicates the C-terminal. The two cysteines of the active site are shown as large dots.
INTRODUCTION

Biochemistry

Thioredoxins are a family of small thiol proteins characterized by a similar structure (referred to as the thioredoxin fold) and a conserved active site, trp-cys-gly-pro-cys (WCPC) (Holmgren 1989). In some cases, one or two of the non-cysteine amino acids are exchanged for some other amino acid, but WCGPC is the “classical” and most common active site motif. The typical thioredoxin is a protein with a size of approximately 12 kDa (Gromer et al. 2004). The basic thioredoxin fold, a consensus structure model shared with other similar proteins, like for instance glutaredoxin and DsbA, consists of three alpha helixes and four beta sheets. Thioredoxins have an additional alpha helix, beta sheet and large connecting loop in the N-terminal part of the protein, with the active site located in the N-terminal end of the second alpha helix, close to the second beta sheet (Figure 1) (Martin 1995). The basic biochemical function of thioredoxins is to reduce or oxidize disulfides in other proteins using the two cysteines of its conserved active site. This essentially means that it creates or breaks inter- or intra-molecular disulfide bonds. The most common mode of action for thioredoxins is reduction and in this process the protein itself becomes oxidized and thereby becomes inactivated. To restore thioredoxin to its active state, Thioredoxin reductase (TrxR) reduces the thioredoxin disulfide through a reaction that converts NADPH to NADP⁺ (Figure 2) (Holmgren 1989).

![Figure 2. The thioredoxin system. Thioredoxin (Trx) in its active form (Trx-(SH)₂) reduces disulfides in its target proteins, and thereby becomes oxidized (Trx-S₂) and hence inactivated. To reanimate the oxidized thioredoxin, Thioredoxin reductase (TrxR) reduces the thioredoxin disulfide in a reaction that turns NADPH into NADP⁺.](image)

Thioredoxin reductases can be roughly divided into two classes: those containing selenocysteine, like mammalian TrxRs, but also TrxR-1 from C. elegans and TrxR-2 from Chlamydomonas, and those containing ordinary cysteine, for example TrxR from Drosophila, Anopheles, C. elegans (TrxR-2) and Plasmodium. TrxRs come in two sizes - large (~55 kDa) and small (~35 kDa). A typical example of a small thioredoxin reductase is TrxR from E. coli, while for example
mammals, *Drosophila* and *Plasmodium* all have large TrxRs (Gromer et al. 2004). The conservation among TrxRs is quite high. *Drosophila* TrxR-1, for example, shares 56 % sequence identity with the human TrxR-1 and 49 % identity with *C. elegans* TrxR-1 on amino acid level (Kanzok et al. 2001).

Thioredoxins have been shown to be involved in many different types of reactions. For instance, thioredoxin was first reported as an electron donor for ribonucleotide reductase, a protein that is necessary for the production of deoxyribonucleotides used in the synthesis of DNA (Orr & Vitols 1966). Later on thioredoxin was shown to function as a structural component of DNA polymerase in the T7 bacteriophage (Huber et al. 1987). It has also been shown to be important for maintaining intracellular proteins in a reduced state (Fernando et al. 1992; Stewart et al. 1998) and to function as a hydrogen donor for 3'-phosphoadenyl-sulphate (PAPS) reductase in *E. coli* (Lillig et al. 1999). In the literature in recent years, thioredoxins have also been increasingly implicated in the cellular system protecting against reactive oxygen species (ROS) and oxidative stress (Nonogaki et al. 1991; Ritz et al. 2000; Jee et al. 2005). A common target for thioredoxins is the peroxiredoxins, proteins with the major function of reducing hydrogen peroxide (*H₂O₂*) (Hofmann et al. 2002). Peroxiredoxins rely on various thiol proteins and thiol substances for their reducing functions. Peroxiredoxins that rely exclusively on thioredoxins are called thioredoxin peroxidases (TPx:s). These are however only a few of the functions that have been described for thioredoxins, countless more thioredoxins and biological roles have been reported.

**Physiological roles**

*Thioredoxins in different cellular compartments*

Thioredoxins are often described as essentially cytoplasmic proteins that are responsible for the maintenance of the redox-status of other proteins, but already in 1989 thioredoxins specifically located in mitochondria of both animals and plants were reported (Bodenstein-Lang et al. 1989) and ten years later a mitochondrial thioredoxin system in yeast was also published (Pedrajas et al. 1999). In plants there are more than 20 different thioredoxins (Gelhaye et al. 2005). Among these are thioredoxins specific for both the mitochondria and the chloroplasts. The mitochondrial thioredoxin system in *Arabidopsis*, consists of AtTRX-o1, a novel thioredoxin with regards to sequence and intron structure, and a thioredoxin reductase, AtNTRB, with high similarity to the cytosolic thioredoxin reductase AtNTRA (Laloi et al. 2001; Reichheld et al. 2005). The chloroplast-specific thioredoxin system is made up of thioredoxins, ferredoxin/thioredoxin reductase
and the iron-sulphur protein ferredoxin. This system is very important for the light-regulation of carbon metabolism (Schürmann 2003).

In mammals, the mitochondrial thioredoxin system is comprised by one thioredoxin (Trx-2) and one thioredoxin reductase (TrxR2). The mitochondrial thioredoxin has an important function as hydrogen donor for peroxiredoxin (Miranda-Vizuete et al. 2000). Mitochondrial peroxiredoxin is a thioredoxin-dependent protein involved in regulation of cell proliferation and in protection against the oxidative stress caused by the hydrogen peroxide that is produced by the mitochondrial metabolism (Immenschuh & Baumgart-Vogt 2005). But the mitochondrion is not the only cellular compartment in mammals, besides the cytoplasm, where thioredoxins can be found. Mammalian Trx-1 translocates to the nucleus when cells are exposed to oxidizing radiation in the form of UV-radiation but also when exposed to nitric oxide (Hirota et al. 1999; Arai et al. 2006). In the nucleus thioredoxin enhances the DNA binding capacity of the transcription factor NF-κB (Hirota et al. 1999) and activates the AP-1 transcription factor through interactions with redox factor-1 (Ref-1) (Hirota et al. 1997; Wei et al. 2000). NF-κB is a well known regulator of the immune response, but is also involved in regulating the response to different forms of stress, like ROS and UV-radiation (Kratsovnik et al. 2005). Thioredoxin has also been found to have functions in the extracellular compartment. A truncated form of human Trx-1 (in those days referred to as eosinophil cytotoxicity-enhancing factor, but later referred to as Trx80) functions as a cytokine, regulating the proinflammatory functions of human eosinophils (Silberstein et al. 1993). It also works as a chemokine for monocytes, T lymphocytes and polymorphonuclear leukocytes (Bertini et al. 1999).

Thioredoxins and disease

The thioredoxin that undeniably has been most investigated with regards to human diseases is human Trx-1. It is a mainly cytosolic protein that under certain conditions can be translocated into the nucleus or be secreted for extracellular functions. Changes in Trx-1-levels have been associated with a number of different diseases. In HIV infected individuals thioredoxin levels are elevated in blood serum (Nakamura et al. 1996, 2001). This may be due to the increase in oxidative stress that has repeatedly been reported for HIV-patients (Pace & Leaf 1995). Although full length Trx-1 suppresses production of the virus, Trx80, the truncated form of Trx-1, counteracts this effect and enhances HIV-production (Newman et al. 1994). Increased thioredoxin levels have also been found in rheumatoid arthritis and related diseases (Maurice et al. 1999). Thioredoxin seems to be a double edged sword in many studied human diseases, not only in HIV, as mentioned above, but also in cancer and other diseases. In early stages it can
INTRODUCTION

counteract the disease through protection against induced oxidative stress, while in later stages it promotes the disease by increasing cell proliferation (Iwata et al. 1997). Many carcinogens exert their deleterious mechanisms through free radicals and ROS, a step where the antioxidative properties of thioredoxin and the other components of the thioredoxin system has the potential to protect cells through elimination of harmful radicals (Gromer et al. 2004). Thioredoxin thereby plays a preventive role in the mechanisms that can lead to cancer, but when the growth of a tumor once has started the positive effects of Trx-1 on cell proliferation and its protective effects against the immune system promotes further tumor growth and thereby worsens the condition (Andersson et al. 1996; Grogan et al. 2000).

Interest in thioredoxins and the thioredoxin system as targets for a range of therapeutic measures in a variety of diseases has increased dramatically in recent years. Although very promising due to the extensive number of conditions where the thioredoxin system has been implicated, the diverse and sometimes ambiguous functions of thioredoxins in vivo complicates matters and demands further research and extended knowledge before thioredoxins can serve as large scale therapeutic targets.

Thioredoxins in the mammalian germline

The mammalian testis-specific thioredoxin system is composed of a thioredoxin reductase (TGR) and four thioredoxins or thioredoxin-like proteins, Sptrx-1, Sptrx-2, Sptrx-3 and Txl-2. The first three are testis/spermatid specific while the last, Txl-2 is associated with microtubules in flagella and cilia (Miranda-Vizuete et al. 2004). TGR differs from most other thioredoxin reductases in that it is not exclusively reducing disulfides in thioredoxins, but also reduces glutathione (Sun et al. 2001). This is accomplished by two separate domains in the protein, a small N-terminal glutaredoxin-like domain and a C-terminal thioredoxin reductase domain. TGR is highly expressed in the testis, especially at the site of mitochondrial sheath formation in early spermatids, but it is also expressed elsewhere, for example in the heart, the lung and the liver (Su et al. 2005). Sptrx-1 was the first spermatid specific mammalian thioredoxin to be identified (Miranda-Vizuete et al. 2001). It is a 486 amino acid protein, with two distinct domains. The N-terminal domain is constituted of a 15-residue motif repeated 23 times. This is a novel domain predicted to have a coiled-coil structure, while the second domain can be described as a typical thioredoxin (Jimenez et al. 2002). Sptrx-1 protein is expressed in elongating spermatids, specifically located to longitudinal columns of the fibrous sheath of the sperm tail (Yu et al. 2002). The second sperm specific thioredoxin to be discovered, Sptrx-2, is a 588 amino acid protein (Sadek et al. 2001). It consists of two parts, just like Sptrx-1, but in this case, the thioredoxin is located in the N-terminal part of the protein and then followed by a repeat of three NDP kinase domains. Sptrx-2 has been suggested to be a structural part of the
fibrous sheath, since enzymatic activity of neither the thioredoxin nor NDP kinase domains could be found in vitro and the protein locates to the principal piece of the flagellum (Miranda-Vizuete et al. 2003). Txl-2, the third thioredoxin with a strong localization to testis and spermatids, is very similar to Sptrx-2 but differs in two regards. Like Sptrx-2 it has an N-terminal thioredoxin, however, the thioredoxin domain is only followed by one NDP kinase domain (Sadek et al. 2003). Another thing that differs from the other spermatid specific thioredoxins is that, although Txl-2 is expressed at high levels in spermatids and located to the microtubule-based spermatid manchette and axoneme, it is also expressed in cilia of the lung and at low levels in various other tissues (Sadek et al. 2003). The fourth, and as of yet, last mammalian spermatid specific thioredoxin, Sprtx-3, stands out from the others in that it lacks any extra domains besides the thioredoxin domain (Miranda-Vizuete et al. 2004). The Sprtx-3 gene is located only 50 kB downstream of Trx-1 in the human genome and shares large sequence similarity with Trx-1, from which it is believed to have originated through a duplication event (Jiménez et al. 2004). Sprtx-3 is located in the Golgi apparatus in elongating spermatids and pachytene spermatocytes, but also briefly located in the developing achrosome. It is important to keep in mind that 3 of the four mammalian testis specific thioredoxins would probably not be classified as thioredoxins due to their large extra domains, but rather thioredoxin-like proteins, if they were to be conventionally named.

**Thioredoxins in Drosophila**

In most organisms there are two separate systems for thioredoxins and glutaredoxins and they function side by side, performing similar tasks with the same basic function of donating electrons. However, in *Drosophila melanogaster*, as well as in *Anopheles gambiae*, there is no gene coding for glutathione reductase (GR) (Kanzok et al. 2001; Bauer et al. 2003a). It is therefore quite likely that the thioredoxin system (and thioredoxin reductase especially) substitutes for GR in reducing the oxidized form of glutathione in both fruit flies and mosquitoes. *Drosophila* thioredoxin reductase (TrxR) has been extensively characterized in vitro (Bauer et al. 2002, 2003b), however very few studies have attempted any studies of DmTrxR in vivo (Orr et al. 2003).

**deadhead**

The first thioredoxin gene to be discovered in *Drosophila melanogaster* was the female specific thioredoxin *deadhead* (*dhd*), also referred to as *DmTrx-1*. *dhd* was found in a screen for recessive female-sterile mutations in genes that mapped within a certain region (4E1-2 - 4F11-12) on the X chromosome (Salz et al. 1994).
Female homozygous *dhd* mutants were shown to lay eggs that seemed morphologically normal, but that in a majority of cases failed to initiate development. In 10% of the cases, where embryos of later stages could be found, the embryos had asynchronous nuclear divisions. At developmental stage ten these embryos had areas of the cortex that lacked nuclei and some embryos also had nuclei that varied in size, suggesting polyploidy. These phenotypes together suggested an arrest in meiosis, but it was later shown by another research group to be an arrest immediately after meiosis, but before the S-M cycles (Page & Orr-Weaver 1996; Elfring et al. 1997). In 5% of the cases the embryos developed far enough to secrete cuticle, but had severe head defects. Male fertility was not affected in these mutants. Using Northern blots and in situ hybridizations Salz and colleagues show that *dhd* RNA is expressed mainly in the female ovaries after oogenesis stage nine, specifically expressed in the nurse cells, but not in the somatic follicle cells (Salz et al. 1994). They also show that *dhd* RNA is maternally contributed to the embryo, where it is evenly distributed and prevails until the four hour stage. However, another research group have later shown that maternally contributed *dhd* RNA is enriched in the pole cells in stage five embryos, and that the embryo itself expresses *dhd* RNA in the pole cells and the embryonic gonad at stage ten and sixteen (Mukai et al. 2006).

The DHD protein consists of 107 amino acids and is 32% and 36% identical to yeast and human thioredoxins, respectively. The two cysteines of its conserved active site are necessary for the function of the protein in vivo, which suggests that the protein has a redox-function in the fly body. However, DHD does not play a role in DNA synthesis (Pellicena-Palle et al. 1997). DHD has been biochemically characterized in vitro and was found to be a 12.4 kD protein with a classical thioredoxin fold (Kanzok et al. 2001). Kanzok and colleagues have also shown that DHD is able to efficiently reduce glutathione disulfide and that it is effectively reduced by Thioredoxin reductase (TrxR), but that it does not function as a substrate for Thioredoxin peroxidase (TPx-1) (Kanzok et al. 2000, 2001; Bauer et al. 2002). The details of DHD’s physiological functions in vivo are still largely unknown.

**Thioredoxin-2**

The next *Drosophila* thioredoxin gene to be cloned and published was *Thioredoxin-2* (*Trx-2* or *DmTrx-2*) (Bauer et al. 2002). The Trx-2 protein has been biochemically characterized; it was found to be able to reduce glutathione as efficiently as DHD. It was also found to be oxidized by TrxR and to be able to function as a substrate for TPx-1 (Bauer et al. 2002). Bauer and colleagues also showed that Trx-2 is expressed in larvae, adults of both sexes and the embryonically derived Schneider cell line. The difference in potential as substrate for TPx-1
INTRODUCTION

between Trx-2 and DHD, coupled with a fairly low amino acid sequence identity (only 38 %) between the two proteins and the difference in expression pattern suggests a difference in functionality (Bauer et al. 2002). Trx-2 has also been crystallized in both reduced an oxidized form (Figure 3) (Wahl et al. 2005). Trx-2 was shown to be structurally very similar to thioredoxins from many other organisms (considering the large evolutionary time span since the divergence of ancestral species), including *E. coli* and humans, which is consistent with the similarities in amino acid sequence (34 % and 47 % identity, respectively). Despite the large similarities with human thioredoxin, oxidized Trx-2 did not form dimers under experimental conditions, and was thereby deemed not likely to form dimers *in vivo*, in contrast to the human thioredoxin (Wahl et al. 2005).

![Figure 3. Ribbon plot of oxidized *Drosophila melanogaster* Trx-2. α1-α4 indicates the four alpha-helices and β1-β5 indicates the five beta-sheets. Alpha-helices and beta-sheets are numbered from the N to the C terminus. C32 and C35 indicate the redox-active disulfide. Adapted from Wahl et al. (2005).](image)

*ThioredoxinT*

The third thioredoxin gene that was discovered in *Drosophila* is *ThioredoxinT* (*TrxT*). It was found with a yeast two-hybrid screen for proteins that interact with the PcG-group protein Zeste (Chen 1992), and was in the initial stages of this work referred to as ZIP42 (short for Zeste-interacting protein 42).
INTRODUCTION

GENE REGULATION

Gene regulation can be thought of as the multitude of different ways in which the genes of an organism can be controlled. There are many different systems and mechanisms controlling the ‘wheres’ and ‘whens’ of the expression of an organism’s genes. The most important basal mechanism that genes in general encounter is transcription, where the genetic information is copied into RNA, the medium that is utilized as a template in the mechanism that converts gene information into a functional protein. Many of the mechanisms that regulate genes are in some way related to transcription, either by directly affecting the transcription machinery by helping it or interfering with it, or more indirectly affecting it by attracting it or repelling it. A good example of a category of proteins that directly affect transcription is the transcription factors. These are proteins that regulate the initiation of transcription and the binding of RNA polymerase (the major component in the transcription machinery) to DNA (Turner 2001). Transcription factors can either be enhancers or repressors, that is, they either assist or block RNA polymerase binding. Many transcription factors bind both DNA (often in the promoter or enhancer region of a gene) and other transcription factors. Transcription and transcription factors are regulated both by the chromatin structure of the gene region that they bind, by epigenetic markers present on the DNA and by the presence of other DNA binding proteins. Two of the important indirect mechanisms that contribute to gene regulation are the cellular memory that maintains correct expression of the genes governing development and the chromosome-wide gene regulation often associated with dosage compensation.

Cellular memory

In basically all higher organisms (from yeast to plants and animals) there is a category of genes called the homeobox genes (Levine et al. 1984; Kaufmann 1993; Williams 1998). These genes all have a conserved domain, the homeobox, which in its protein form can bind to DNA. The Homeobox genes are generally transcription factors. In nematodes and higher animals, a certain category of the homeobox genes are called homeotic genes or Hox genes (after the Hox gene cluster in which they are situated) (Garber et al. 1983; Scott et al. 1983; Way & Chalfie 1988). The homeotic genes are very important to any organism that has them since they determine the layout of the organism’s body plan and they have been described as the master regulators of development (Turner 2001). Hox genes are highly conserved through evolution and are generally organized in conserved clusters in the genome. The first homeotic genes, before they were even called Hox genes, were found in Drosophila (Lewis 1978). There are two clusters of
INTRODUCTION

homeotic genes in Drosophila: the Antennapedia complex (ANT-C) and the Bithorax complex (BX-C) (Lewis 1978; Lewis et al. 1980a-b). The Antennapedia complex is responsible for the segment identity of the head and the thorax and consists of labial (lab), proboscipedia (pb), Deformed (Dfd), Sex combs reduced (Scr) and Antennapedia (Antp). The Bithorax complex is responsible for the segments in the abdomen and consists of Ultrabithorax (Ubx), abdominal-A (abd-A) and Abdominal-B (Abd-B). If all homeotic genes would be active all the time, every bodysegment in an organism would look exactly the same. To make sure that only the right homeotic genes are on at the right time there is a transient regulation performed by gap genes and segmentation genes, that activates or silences the homeotic genes. To maintain the activated or silenced states of the homeotic genes there is a controlling system commonly referred to as the cellular memory. The cellular memory generally consists of two systems, one that mediates maintenance of the active state and another that mediates maintenance of the repressed state.

The PcG and trxG complexes
The cellular memory system in Drosophila that maintains the correct expression of the homeotic genes is made up of two basic units; the Polycomb Group (PcG) complex, which is responsible for maintaining the inactive state of genes that have initially been turned off, and the trithorax Group (trxG) complex, which is responsible for maintaining the active state of genes that have initially been turned on. The complexes bind specific regulatory elements in the vicinity of the homeotic genes, where they exert their activating or silencing functions. These regulatory elements are commonly called PcG or TrxG response elements (PREs/TREs) (Simon et al. 1993). Both of these protein complexes contain a number of different proteins. Some examples of genes coding for PcG components are Polycomb (Pc), Posterior sex combs (Psc), polyhomeotic (ph), Enhancer of zeste (E(z)), extra sexcombs (esc), and Suppressor of zeste 12 (Su(z)12), and examples of trxG component genes are trithorax (trx), brahma (brm), absent, small or homeotic discs 1 and 2 (ash-1, ash-2) and zeste (z), and there are many more (reviewed by Ringrose & Paro 2004, and Schwarts & Pirrotta 2007). It is not known exactly how the PcG and trxG complexes maintain the silenced or activated states of the homeotic genes, despite decades of research, but histone methylation has been repeatedly shown to be involved in both cases. Both mono, di and tri-methylation of different lysines of the histone tails have been shown for the PcG complexes (Schwarts & Pirrotta 2007). Histone de-acetylation has also been correlated with PcG silencing, while acetylation has been connected to trxG. In recent years, more and more reports suggest that silencing is the default state and at least two of the TrxG proteins have been proposed to be repressors of silencing rather than true activators (Poux et al. 2002; Klymenko & Müller 2004).
INTRODUCTION

Chromosome wide gene regulation

In most organisms the loss of one of the two copies of any chromosome leads to lethality (with some exceptions). This is because the total transcript dose of all expressed genes on that chromosome goes down to half that of the normal level and the combined effect of all these drastic alterations in transcript level is too much for the cell to handle. In many organisms however, males and females have different numbers of sex-determining chromosomes, i.e. one sex is homogametic (e.g. XX for mammalian females and ZZ for male birds) while the other is heterogametic (e.g. XY for mammalian males and ZW for female birds). In most cases the Y or W chromosome has lost the majority of genes it perhaps once had in common with its partner chromosome and mainly contains sex-determining genes and/or genes that are exclusively involved in fertility. The heterogametic sex is thereby effectively haploid for all the genes on the X or Z chromosome. This would present a problem for the cell if it were not for the phenomenon of dosage compensation.

Dosage compensation is the way in which an organism compensates for the inequality in gene dosage on the sex-chromosome common to both sexes. This has been solved in many different ways through evolution, the best known being those of mammals, nematodes and fruit flies. The modes of dosage compensation in these three organisms differ quite a lot. However, recent research in the field suggests that they are all based on an initial general up-regulation of the X chromosomes, followed by specific mechanisms to equalize the dose between the sexes (Gupta et al. 2006). In mammals, one X chromosome becomes inactivated (X-inactivation), while in *C. elegans* transcription on both hermaphrodite X chromosomes is downregulated to half (hypo-transcription), in both cases compensating for a general up-regulation of the X chromosomes compared to the autosomes (Figure 4). *Drosophila* differs from the other two in that the dosage compensation occurs in males, where transcription of the male X is upregulated two-fold (hyper-transcription) (Figure 4).

**Figure 4.** Three variants of dosage compensation. A) X-inactivation in mammals. B) Hypotranscription in the nematode *Caenorhabditis elegans*. C) Hypertranscription in the fruit fly *Drosophila melanogaster*. 
Dosage compensation in mammals
In mammalian dosage compensation there are basically three important steps: 

i. up-regulation of the X chromosome in males and females, 

ii. inactivation of one of the female X chromosomes and 

iii. maintenance of silencing of the inactivated X (Xi) (Heard & Disteche 2006). It has been known for a long time that one of the female X chromosomes becomes inactivated, forming the Barr body (Lyon 1961). However, it was not until resent years that it became clear that there is also up-regulation of the X chromosomes in relation to the autosomes, i.e. step one listed above (Adler et al. 1997; Gupta et al. 2006; Nguyen & Disteche 2006). This up-regulation is particularly strong in the brain and absent in haploid germ-cells. It is, as of yet, not known what is controlling this up-regulation, whether it is caused by a protein complex or by evolutionary changes in the regulatory regions of X genes as an indirect response to loss of the corresponding genes on the Y chromosome (Nguyen & Disteche 2006).

Much more is known about what factors are controlling the inactivation and maintenance of silencing of one of the female’s X chromosomes. For X inactivation to occur at all, a cell must have (at least) two copies of a complex X chromosome region called the X inactivation centre (Xic) (Rastan & Robertson 1985). The Xic, which is also involved in the “chromosome counting” that determines whether X inactivation should occur at all (Alexander & Panning 2005), contains the Xist gene, which codes for a non-coding RNA (Brown et al. 1991). The Xist RNA is expressed exclusively from the inactivated chromosome and coats its entire length (Brown et al. 1992; Clemson et al. 1996). Partially overlapping the Xist gene lies Tsix, which codes for an Xist antisense RNA that is involved in early regulation of Xist (Lee et al. 1999). The mechanism for this is not fully known, but transient colocalization of the two Xics, suggesting cross-talk, has recently been shown (Bacher et al. 2006). Other elements located in the Xic, e.g. Xite, have also been connected with Xist regulation (Ogawa & Lee 2003).

The major function of Xist RNA is to draw silencing complexes to the inactivated X, for example PRC1 and PRC2 (Plath et al. 2003). These are involved in histone modifications of various kinds, but other modifications, for instance DNA methylation, are also typical for the Xi. Some examples of histone modifications are the hypoaetylation of histone H4 (Keohane et al. 1996), the enrichment of H3 Lys-27 trimethylation (Plath et al. 2003; Silva et al. 2003) and H2A K119 monoubiquitylation (de Napoles et al., 2004; Fang et al. 2004). Many of these modifications have been, or are suggested to be, connected with irreversible silencing of the Xi (Kohlmaier et al. 2004).
Dosage compensation in *C. elegans*

In the nematode *C. elegans* both the sex of an individual and the dosage compensation is determined by the ratio of X chromosomes to autosomes. All ratios below 2/3 result in male development, while ratios of 3/4 and above result in hermaphrodite development and dosage compensation in the form of down-regulation of both X chromosomes (Cline & Meyer 1996). In *C. elegans* a dosage compensation complex (DCC) has been identified, a complex that in turn consists of two subcomplexes with several members, some of which are also involved in sex determination. One subcomplex, the DPY-complex (named after genes with the dumpy phenotype) consists of DPY-26, DPY-27, DPY-28 and MIX-1, while the other subcomplex, SDC (named after a group of genes found in a screen for sex determination and dosage compensation mutants) is composed of SDC-1, SDC-2, SDC-3, DPY-21 and DPY-30 (Lucchesi *et al.* 2005). The DPY-complex can be described as an adapted form of the condensin complex (the MIX-1 protein is for example a member of both complexes). The condensin complex is a chromosome associated complex necessary for sister chromatid resolution and DNA compaction, both during mitosis and meiosis (Hagstrom & Meyer 2003). Recruitment of the DPY-complex to X chromosomes requires at least three of the components of the SCD-complex: SDC-2, SDC-3 and DPY-30.

The *C. elegans* Dosage Compensation Complex is initially recruited to the X chromosome at a number of recruitment sites and later spread in *cis* along the chromosome (Csankovszki *et al.* 2004). The DCC not only binds to the X chromosome, it also locates to the *her-1* gene, where it acts as a switch for hermaphrodite development by a 20-fold repression of *her-1* (Davis & Meyer 1997). The master switch for both dosage compensation and sex determination is the gene *xol-1*. The functional role of *xol-1* is to repress *sdc-2*, thereby also blocking *her-1* repression, resulting in simultaneous repression of dosage compensation and activation of male differentiation (Miller *et al.* 1988). It was not known until very recently that also the nematode X chromosomes are generally upregulated compared to the autosomes, in both males and females (Gupta *et al.* 2006). Very little has so far been reported in the subject, but the suggested up-regulation would certainly provide an important and logical piece in the puzzle, since it is otherwise hard to explain how an organism can survive with a general X chromosomal transcription level corresponding to that of a single autosome, *i.e.* leading to haploinsufficiency. With a general up-regulation of the male and female X chromosomes, a dosage compensation system that halves the transcription level of the female X:as would lead to an equalized transcription level of all chromosomes, in both sexes.
INTRODUCTION

Dosage compensation in Drosophila
Contrary to mammals and nematodes, the fruit fly has a form of dosage compensation that affects males only. In Drosophila the single male X is upregulated two-fold, equalizing the expression from the male X to that of the two X chromosomes in females (Larsson & Meller 2006). The core dosage compensation complex (DCC) in Drosophila consists of the five proteins MSL-1, MSL-2, MSL-3, MLE and MOF and either one of two non-coding RNAs. The genes encoding the five proteins are named after their characteristic male-lethal phenotypes when mutated: male specific lethal -1, -2 and -3, maleless and males absent on first. mle was first discovered by Fukunaga and coworkers (1975) but was not really connected with dosage compensation of the single male X chromosome until 1980 (Belote & Lucchesi 1980b). The same year, Belote and Lucchesi had published the finding of two additional sex specific lethals: msl-1 and msl-2 (Belote & Lucchesi 1980a), but it took many years before the two remaining components were found; characterization of msl-3 was first published by Gorman and coworkers (1995) and mof was found by Hilfiker and coworkers (1997). The two non-coding RNAs, roX1 and roX2, were published in the same issue of Cell (Meller et al. 1997; Amrein & Axel 1997). They were later shown to be essential components of the dosage compensation machinery (Franke & Baker 1999).

All MSL proteins, except MSL-2, are expressed also in females. This means that MSL-2 acts as a master switch for the entire complex. In the presence of the sex-determining protein SXL (which is the component that acts first in the cascade leading to female-specific development) MSL-2 is repressed, but in its absence MSL-2 is expressed, thereby enabling complex formation (Bashaw & Baker 1995, 1997). If MSL-2 is ectopically expressed in females, the dosage compensation system is switched on, resulting in up-regulation of the female X chromosomes, virtually rendering them transcriptionally tetraploid, which leads to lethality (Kelley et al. 1995). The first step in the formation of the DCC is an interaction between MSL-1 and MSL-2, via MSL-1s coiled-coil domain and a RING finger in MSL-2 (Lucchesi et al. 2005), enabling both binding of the other components and binding of the complex to DNA (Lyman et al. 1997; Copps et al. 1998). Next, the histone acetyltransferase MOF binds to MSL-2 and MSL-3 associates with MSL-1 after acetylation by MOF (Gu et al. 1998; Scott et al. 2000; Buscaino et al. 2003; Morales et al. 2004). The fifth protein subunit, MLE, is an ATP dependant DNA/RNA helicase that only seems to associate with the other proteins through one of the two roX RNAs (Richter et al. 1996; Lee et al. 1997). All five DCC proteins are essential for male survival, while either one of roX1 or roX2, but not both, can be removed without effect (Meller & Rattner 2002; Deng et al. 2005).
The dosage compensation complex is targeted to the male X via 25-35 entry sites, the roX genes being two of the first sites where the complex binds (Kelley et al. 1999; Gu et al. 2000; Demakova et al. 2003). From these sites, binding of the complex spreads in cis to sites with lower affinity. Recent publications suggest that the complex binds to actively transcribed genes (Alekseyenko et al. 2006). MSL localization to the X can be seen as a banded pattern along the whole chromosome when visualized with antibody staining against any of the complex-components. Intriguingly, the MSL complex does not locate to the X chromosome in the male germline, even though it has been shown to be dosage compensated (Rastelli & Kuroda 1998; Gupta et al. 2006). Essentially nothing is so far known about either the mechanism or the components that might be involved in the up-regulation of the X chromosome in the male germline.

Painting of fourth in Drosophila
The tiny, largely heterochromatic, 4th chromosome in Drosophila is the only autosome that can be present in only one copy without causing lethality (Hochman 1976). This has often been explained by the fact that the 4th chromosome is so small; it only contains about one hundred annotated genes, compared to several thousand each for the other chromosomes. It was also the most plausible explanation until the gene Painting of fourth (Pof) was discovered by Larson and co-workers (2001). Pof is situated in region 60E on chromosome 2R and encodes a novel protein with an RNA-binding domain (RRM1). Pof RNA levels are the highest in males, with lower levels in pupae and larvae. Pof RNA is also present in females but at much lower levels and there is hardly any maternal contribution of Pof RNA. Antibody staining of POF protein on salivary gland polytene chromosomes reveals a specific localization of POF to the 4th chromosome, hence the name Painting of fourth (Larsson et al. 2001). The staining is equally strong in males and females, despite the difference in Pof RNA levels in male and female flies. Larson and coworkers could, through a number of translocations between the 4th and some of the other chromosomes, show that binding of POF to chromosome 4 was dependent on a proximal region of the chromosome. The binding then spreads in cis over the remainder of the chromosome. The spreading is also dependent on sequences or chromatin structures specific to chromosome 4, since it will not spread into autosomal pieces translocated to the 4th, but can spread to translocated chromosome 4 regions in trans.

Different numbering systems of the chromosomes have been used in different Drosophila species. Therefore it is better to use Muller’s classification of Drosophilid chromosomes, where D. melanogaster’s 4th chromosome corresponds to the F-element (Muller 1940). In the related species Drosophila busckii the F-element is part of the X-chromosome. In males, but not females, of this species
POF staining is seen along the entire X-chromosome, suggesting an involvement of POF in the dosage compensation system of *D. busckii* (Larsson *et al.* 2001). Since spreading of POF in *D. melanogaster* is dependent on chromosome 4 specific sequences or structures, Larsson and colleagues suggested that the *DINE-1* element, which is a short repetitive sequence highly enriched in chromosome four but also in other heterochromatic chromosome regions (Locke *et al.* 1999), is a possible candidate as a target for POF. A *DINE*-probe did however not specifically hybridize to the X chromosome in *D. busckii*, either suggesting that the *DINE* element is not involved in POF targeting or that the element has diverged to a point where a *D. melanogaster DINE*-probe will no longer cross-hybridize with a *D. busckii DINE* fragment. A few years later it was shown that a group of short sequence motifs, which highly contributed to separating the 4th chromosome from the other chromosomes in a multivariate sequence signature analysis, resided within *DINE-1* elements (Stenberg *et al.* 2005). Regions of chromosome 4 with low numbers of these motifs correlate well with regions where POF does not bind on polytene chromosomes. This suggests that *DINE-1* elements, or sequence motifs colocalized with them, are directly or indirectly involved in POF’s binding to chromosome four (Stenberg *et al.* 2005). All of these findings together suggest that POF is part of a regulatory complex for the 4th chromosome, which is quite extraordinary since it is the first autosome specific regulatory complex.

**Gene regulation in the male germline of *Drosophila***

*Spermatogenesis*

The *Drosophila* male reproductive tract consists of several different organs; the testes, the testicular ducts, the seminal vesicles, the accessory glands, the anterior ejaculatory duct, the ejaculatory bulb and the posterior ejaculatory duct, with the major part of spermatogenesis carried out in the testis (Fuller 1993). Located in the tip of the testis is a cluster of cells, referred to as the germinal proliferation center. This cluster consists of three different cell types; the apical cells, the germline stem cells and the cyst progenitor cells. The first step in spermatogenesis is when one of the stem cells goes through a radially oriented mitotic division, resulting in two daughter cells, one stem cell and one primary spermatogonial cell (Figure 5). The spermatogonial cell goes through another mitotic division followed by further mitosis in its daughter cells (four rounds of mitosis in total), ultimately resulting in a cyst of 16 primary spermatocytes (Fuller 1993). The primary spermatocytes grow and mature. During this phase many genes that were previously silenced are activated and expressed at high levels and essentially all the RNA and proteins that are needed later during spermatogenesis are produced, since transcription is in essence shut down at the onset of meiosis. In these cells the Y chromosome,
INTRODUCTION

forming large loop-like structures, is an especially prominent feature when viewed with light microscopy. The growth phase is followed by meiosis, producing a cyst of 64 haploid early spermatids. The early spermatids go through extensive differentiation, radically transforming them from round cells with ‘normal’ internal architecture to highly elongated cells with tightly packaged cell machinery. A morphologically very characteristic sub-stage in spermatid differentiation is the onion stage, where the individual mitochondria have fused and become packaged into a dense sphere, giving the cell the look of a bowl with two marbles in it, one black (the mitochondrial sphere) and one white (the nucleus). In the later part of spermatid differentiation, elongation of the spermatid takes place. During this phase the flagellar axoneme is assembled, eventually forming the main part of the sperm tail. The very last stage of spermatogenesis includes the individualization and coiling of the mature sperm (Fuller 1993).

![Figure 5. Schematic view of Drosophila spermatogenesis. Adapted from Fuller (1993).](image)

*The Y chromosome*

The *Drosophila* Y chromosome is metacentric, with a long and a short arm. Sequence homology between the Y and the X, both in genes and repetitive sequences, is extremely low, so low that it has been proposed that the Y chromosome does not evolutionarily stem from the X, but from a B-chromosome. This is supported by the fact that the majority of the few known genes on the Y chromosome have ancestral genes on the autosomes and not on the X (Carvalho 2002). A very high number of them are also genes involved in male fertility. Recent data regarding the Y chromosome in *D. pseudoobscura* suggests that the *D. melano-
INTRODUCTION

gaster Y and the D. pseudoobscura Y are not homologous. Out of 15 genes and pseudogenes on the D. pseudoobscura Y none had homologs on the D. melanogaster Y, while the six known D. melanogaster Y genes all had orthologs on D. pseudoobscura autosomes, suggesting that the D. pseudoobscura Y has arisen de novo from autosomes (Carvalho & Clark 2005).

In most Drosophila cells the Y chromosome is highly heterochromatic, but in the primary spermatocytes of the male testis the Y chromosome takes on a huge floated form. These peculiar structures are called the Y chromosome loops and have been identified as the three fertility factors ks-1, kl-3 and kl-5 (Bonaccorsi et al. 1988). The fertility factors (kl-5, kl-3, kl-2 and kl-1 on the long arm of the Y chromosome and ks-1 and ks-2 on the short arm), are gigantic genes with numerous and large introns containing repetitive sequences, but also a few smaller genes. Only a handful of genes have been identified on the D. melanogaster Y chromosome; e.g. kl-2 and kl-3 encode dynein heavy chain polypeptides and ks-1 and ks-2 have been suggested to correspond to an occludin-related gene and a coiled-coils gene respectively (Carvalho et al. 2000, 2001). The Y chromosome loops have been suggested to function as storage compartments for proteins that are expressed in large quantities in the primary spermatocytes but not needed until later in spermatogenesis (Hennig et al. 1989), but some of the proteins reported to bind to the Y have also been suggested to be involved in regulation of some of the Y genes (Marhold et al. 2002).

Testis specific gene regulation
There are two separate systems in Drosophila ensuring correct expression and translational timing of testis specific genes (White-Cooper et al. 1998). The first system, which is responsible for transcriptional activation of a large number of testis specific genes, is comprised of the meiotic arrest genes always early (aly), achintya/vismay (achi/vis) and matotopetli (topi). The first aly allele was found in a screen for mutations induced by transposable elements (Cooley et al. 1988). aly was later shown to be a homolog of a C. elegans gene, lin-9 (White-Cooper et al. 2000), a gene involved in regulating vulval cell fate specification (Ferguson & Horvitz 1989). The two tandem-repeat genes achi and vis encode TALE-class homeodomain transcription factors and are homologs of human TGIF (Ayyar et al. 2003). conr was found in an EMS-screen for male sterile mutations and encodes a novel acidic protein (Jiang & White-Cooper 2003). conr and aly are mutually dependent. topi was found in the same screen as conr and encodes a testis-specific Zn-finger protein (Perezgasga et al. 2004). The protein products of the four genes are believed to be part of a protein complex that regulates the transcription of approximately one thousand genes specifically expressed in the testes.
The second system, which is responsible for another set of testis specific genes (among them genes required for spermatid differentiation), is comprised of five TAF-homologs: cannonball (can), no hitter (nht), spermatocyte arrest (sa), meiosis I arrest (mia) and ryan express (rye). The can gene was discovered in a screen for male-sterile mutations and encodes a homolog of dTAF5 (Lin et al. 1996; Hiller et al. 2004). nht, mia, sa and rye were all found in a search for second homologs of known TAF-genes (Hiller et al. 2004). TAFs (TBP-associated factor) are proteins that together with TBP (TATA-binding protein) form the general transcription factor TFIID (Albright & Tjian 2000). nht encodes a dTAF4 homolog, sa encodes a homolog of dTAF8, mia a homolog of dTAF6 and rye a homolog of dTAF12 (Hiller et al. 2004). These are believed to form a complex that is necessary for the regulation of genes involved in terminal spermatid differentiation. One example is the twine/cdc25 phosphatase, which is necessary for proper chromosomal segregation during the transit from G2 to M in meiosis (White-Cooper et al. 1993).
AIMS

The aims of the work presented in this thesis have been:

- To clone the \( TrxT \) gene and investigate the expression patterns of the two sex-specific \( Drosophila melanogaster \) thioredoxins.

- To create mutants in the \( Drosophila melanogaster \) Trx-2 gene and investigate the mutant phenotypes and potential redundancy.

- To investigate the conservation of TrxT, DHD and Trx-2 and the \( TrxT – dhd \) gene region in other insect species.

- To create mutants in the \( Pof \) gene and investigate whether POF’s localization to chromosome 4 is conserved in evolution.

- To investigate POF’s expression and function in the male germline.

- To investigate the connection between Zeste and thioredoxins.
MATERIALS AND METHODS

The methods used in the papers are described in the material and methods section of each paper.

METHODS USED IN THIS THESIS

In vitro generation of zeste mutants

Mutated versions of zeste, in which the two cysteine-encoding codons were replaced with serine-encoding codons, were constructed using a Quick-change Site-directed Mutagenesis kit, according to the instructions of the manufacturer (Stratagene), and the primers 5’-GAACAACTGCAGAAGTCTTCCCTTTCA TGAGG-3’, 5’-CCTCATGAAGTGGAGACTTCCGCAGTTGTTC-3’, 5’-GC TGAACCGGCTCAGGAAGCAGCAG-3’ and 5’-CTGCTGCTCCTTTGAG CGCAGTTCAGC-3’.

Protein aggregation experiments

Proteins were translated in vitro using a TNT® Quick Coupled Transcription /Translation System (Promega) according to the manufacturer’s protocol. 35S-labeled methionine was used to label and detect the proteins. Two types of experiments were performed. In the first, 6 µl of Zeste was mixed with 4 µl of either mock (i.e. a sample with no added DNA), DHD, TrxT or Trx-2 and 5 µl buffer (50 mM NaCl, 20 mM Tris-HCl, pH 7.8). In the second, 8 µl of Zeste, Zc520s, Zc571s or Zc520sc571s was mixed with 4 µl samples of either mock or DHD and 3 µl buffer, and 4 µl Zc520s was mixed with 4 µl Zc571s, 4 µl mock or DHD and 3 µl buffer. All samples were incubated for 90 min and then centrifuged for 10 min at 16000g at 4°C. The resulting pellets and supernatants were separated into different tubes. Each pellet was washed with 30 µl buffer and centrifuged as previously. The buffer was removed and the pellet was dissolved in 30 µl 1 × loading buffer (4 M Urea, 100 mM Tris-HCl pH 7.6, 2 % SDS, 5 % BME, 5 % Ficoll). The supernatant was mixed 1:1 with 15 µl 2 × loading buffer. All samples were denatured at 100°C for 5 min and then separated on SDS-PAGE gels. The gels were fixed in fixation buffer (1:2:7 acetic acid: methanol: dH2O) for 5 min, dried in a GelDryer at 80°C for 60 min, then exposed and the labeling in Zeste bands was quantified using a PhosphorImager.
RESULTS AND DISCUSSION

PAPERS I-V

*TrxT* encodes a male-specific thioredoxin (Paper I)

The female-specific thioredoxin gene *dhd* was for many years the only known sex-specific thioredoxin in *Drosophila* (Salz et al. 1994), until we published the cloning and characterization of *TrxT*, a male-specific thioredoxin gene (Paper I). We found that the *TrxT* gene is located right next to the *dhd* gene, with only 471 bp separating them. Salz and colleagues had previously shown that *dhd* encodes a thioredoxin specifically expressed in the female ovaries and that flies null mutant for *dhd* are female sterile. When we tested flies that were null mutants for *TrxT* we found that they were both viable and fertile and we have not been able to find any other obvious phenotypes in these flies. With a number of different Northern blots we could show that *TrxT* is expressed exclusively in males, starting in larvae and culminating in the adult, and that expression of both *TrxT* and *dhd* is not dependent on correct X/Y chromosome dosage, but is dependent on proper development of the gonads (since neither *TrxT* nor *dhd* is correctly expressed in intersex animals or pseudomales).

*TrxT* locates to *Y*-chromosome loops in primary spermatocytes

To examine TrxT and DHD expression *in vivo* we constructed fly strains that carried *TrxT-EYFP* and *DHD-ECFP* expressing constructs and examined live preparations of their testes and ovaries. We found that both TrxT and DHD are predominantly nuclear. DHD specifically locates to nurse cell nuclei and the oocyte karyosome. TrxT can first be seen in nuclei of primary spermatocytes (the growth stage between mitosis and meiosis) and persists in the nucleus until spermatid individualization. In primary spermatocytes TrxT locates specifically to the two *Y*-chromosome loops corresponding to the *ks-1* and *kl-5* fertility factors. *Y*-loops have among other things been suggested to function as storage compartments for proteins needed in later stages of spermatogenesis (Hennig et al. 1989). This could mean that the TrxT binding to *ks-1* and *kl-5* is a storage phase and that TrxT is needed at a later stage, for instance during spermiogenesis (the transformation of post-meiosis cells to finished sperms). A possible role for TrxT could be to ensure correct folding of certain proteins or to assist in dissolving protein complexes that are stabilized by disulfide bonds. Another possibility is that TrxT is involved in regulating the *Y* chromosome fertility factors. dMBD2/3, a protein believed to be involved in epigenetic silencing, has for instance also been shown to bind to at least one of the *Y*-loops (Marhold et al. 2002). Yet another possible role for TrxT in the testis could be to protect against reactive oxygen species.
RESULTS AND DISCUSSION

(ROS) in the gonads. In humans, oxidative stress has been suggested as one of the major reasons for male infertility and loss of sperm motility (Aitken & Baker 2006). This function would however most likely have to be carried out prior to or at the latest during sperm individualization, since TrxT does not seem to be present in or on mature sperm, but can be found in excessive amounts in the waste bag, where all material not needed in the mature sperm ends up (results not shown).

Trx-2 mutants have shorter life spans (Paper II)

Since TrxT null mutants have shown no obvious phenotype, we suspected that there might be redundancy between TrxT and the third known Drosophila thioredoxin, the ubiquitously expressed Trx-2, which is also expressed in nuclei of primary spermatocytes, as we have previously shown (Paper I). Trx-2 has previously been biochemically characterized and crystallized (Bauer et al. 2002; Wahl et al. 2005), but hardly anything was previously known about Trx-2’s function in vivo. To investigate the possible redundancy and to learn more about what Trx-2 is doing in the fly body we constructed deletion mutations in the Trx-2 gene. We recovered fourteen deletion mutants and one revertant. The mutant alleles produce transcripts, but no proteins that can be recognized by our Trx-2 antibody. All of the mutants are probably null mutants for Trx-2, but one of them is a guaranteed null mutant since it lacks the region that codes for the thioredoxin active site. When examining the mutant phenotypes we found that Trx-2 mutant flies are viable but have shorter life spans than control flies. The mutants are equally sensitive to oxidative stress as control flies, but over-expression of Trx-2 mediates an increased resistance to oxidative stress. When we examined double mutant flies (TrxT; Trx-2 and dhd; Trx-2) we found that the double mutants had life spans similar in length to control flies, but they were slightly more sensitive to oxidative stress.

It is very striking that one can remove three major thioredoxins in Drosophila without causing lethality or other similarly strong phenotypes, especially since Drosophila lacks glutathione reductase and consequently should be completely reliant on the thioredoxin system for reduction of glutathione and for other common glutaredoxin-system functions (Kanzok et al. 2001). The strong dependence on the thioredoxin system is made apparent by the lethality seen in Drosophila thioredoxin reductase (TrxR-1) mutants (Missirlis et al. 2002). The entire thioredoxin system does however not seem to be as essential as TrxR, since both Trx-2, TrxT and DHD can be removed without any ensuing lethality. Neither are Trx-2 mutants any more sensitive to oxidative stress than controls, but the fact
that the double mutant is slightly more sensitive to oxidative stress could perhaps suggest that either TrxT or DHD, or both, are in some way connected to protection against oxidative stress in the gonads. Considering this, it might be surprising that a knock-down of the cytosolic *Drosophila* thioredoxin peroxidase TPx-1 (using RNAi) causes increased sensitivity to oxidative stress (Radyuk *et al.* 2003), while over-expression of TPx mediates resistance to oxidative stress (Missirlis *et al.* 2003). The over-expression results for *Drosophila* TPx are however in line with our over-expression results for Trx-2. Over-expression of thioredoxins has also in many other cases been shown to increase resistance to oxidative stress: over-expression of DHD has recently been shown to increase resistance to paraquat, a substance that is commonly used to induce oxidative stress (Chen *et al.* 2004). Similar results have been shown in both mammalian cell-lines and mice (Mitsui *et al.* 2002; Byun *et al.* 2005).

**Thioredoxin redundancy or specialized functions?**

The negative effect of Trx-2 deletion on life span but not on oxidative stress could suggest that there is extensive redundancy between different thioredoxins or thioredoxin-like proteins in *Drosophila*, but it could also mean Trx-2 has specialized functions with a stronger connection to systems affecting life span than to systems affecting oxidative stress tolerance, *i.e.* Trx-2 is perhaps not at all responsible for reducing TPx. This would however mean that another thioredoxin is responsible for reducing the cytosolic TPx. Neither TrxT nor DHD could be candidates for this since both are exclusively expressed in the gonads. This leaves us with the other thioredoxin-like proteins identified in our database search. The two major candidates, based on active site amino acid sequence, would be Txl and the protein encoded by CG13473. Both are possible, but Txl seems less likely since it has a large extra domain that equals the thioredoxin domain in size. The Txl ortholog in humans has recently been suggested to have a function in sugar starvation stress response, suggesting a quite different role for Txl in the cell machinery (Jiménez *et al.* 2006). Essentially nothing is known about CG13473, as can be guessed by the anonymous designation. It would be very interesting to create and analyze mutations in CG13473 and to investigate the expression profile of the CG13473 encoded protein. It could also be very informative to make double mutants of CG13473 and Trx-2 and analyze their response to oxidative stress. It should also be mentioned that two of the identified thioredoxin-like proteins are suggested glutaredoxins/-related proteins: clot is a glutaredoxin-related protein that is necessary for reduction of glutathione in the production of drosopetins, the red pigments of the *Drosophila* eye (Giordano *et al.* 2003), and Grx-1 has been classified as a glutaredoxin on Flybase (http://flybase.org), by sequence identity alone. Grx-1 could very well act as reducer for a number of the other
RESULTS AND DISCUSSION

peroxiredoxins that have been found in *Drosophila* (Rodriguez et al. 2000; Radyuk et al. 2001).

**TrxT, dhd and Trx-2 are conserved in Drosophilids (Paper III)**

In our first thioredoxin-study we found that the TrxT, DHD and Trx-2 proteins in *D. melanogaster* have a fairly high amino acid sequence identity to each other and we have also found that TrxT and dhd, the two sex-specific genes are located right next to each other, transcribed in opposite directions (Paper I). We were intrigued, both by the likeness of the three proteins and the unusual gene organization of TrxT and dhd and therefore decided to look for the three thioredoxin genes in other insect species. We found that all three thioredoxins had conserved putative orthologs in all of the twelve *Drosophilid* species in the genome database (Paper III). Additionally, putative orthologs for Trx-2 were found in the more distantly related species *Anopheles gambiae* (mosquito) and *Tribolium castaneum* (red flour beetle). We also found that the organization of TrxT and dhd is conserved in the Drosophilids, with two interesting exceptions. In one species, *D. ananassae*, the dhd gene seems to have been subjected to retro-transposition at some point in time, since dhd is located more than 1 Mbp away from TrxT and the neighboring genes snf and CG4198 in this species. In the second species, *D. willistoni*, dhd also seems to have been subjected to retro-transposition since it is located immediately downstream of TrxT in this particular species (inserted between snf and TrxT), as opposed to upstream in all the other cases. Additionally, two extra copies of the dhd gene are present in the *D. willistoni* genome, suggesting more retro-transposition events. These results, combined with data from phylogenetic comparisons of the three thioredoxins, showing that TrxT and DHD cluster together despite the fact that TrxT shares the highest sequence identity with Trx-2, clearly suggest that Trx-2 is the ancestral gene, which at some point gave rise to TrxT, which in turn gave rise to dhd, most likely through retro-transposition.

**TrxT has an extremely variable C-terminal domain**

When comparing the twelve *Drosophila* TrxT orthologs we realized that the C-terminal domain (an additional domain after the common thioredoxin domain) of TrxT is hyper-variable, *i.e.* has very low amino acid sequence identity between different *Drosophila* species. A thioredoxin-like protein in *Drosophila* that also has an additional domain is Txl, with an extra domain of roughly 190 amino acids. This domain is, contrary to TrxT’s extra domain, as conserved as Txl’s thioredoxin domain. In fact, the entire Txl protein is conserved, from *C. elegans*, to *Drosophila*, to humans (Miranda-Vizuete & Spyrou 2000). The extra domains of Txl orthologs have no homologies to other known protein domains, but human Txl
has been implicated in cell response to glucose starvation in a human cell line (Jiménez et al. 2006). Considering the large difference in conservation level of the extra domains of TrxT and Txl, and considering how different their extra domains are in amino acid sequence, it seems unlikely that these two domains should be related in function. There are, however, also other known thioredoxins with extra domains, especially in humans. Many of these are exclusively expressed in the male testis and bind specific compartments of the mature sperm tail (Miranda-Vizuete et al. 2004). Many of the extra domains of the mammalian sperm specific thioredoxins have however been determined to be nucleosidediphosphate kinase (NDPk) domains.

The very high variability of the extra domain of TrxT would suggest that the part of TrxT that encodes the extra domain is evolving at a very high rate, while the rest of the gene evolves at neutral rate. Since the extra domain of TrxT is so extremely variable, it can not be accurately aligned and the evolutionary rate can therefore not even be determined. A certain type of genes that are often highly variable, even between closely related species, are genes involved in sexual conflict and male-female co-evolution (Swanson & Vacquier 2002). A good example of this type of quickly evolving genes, is the seminal protein Ovulin (also referred to as Acp26Aa) (Tsaur & Wu 1997; Wong et al. 2006). Ovulin is expressed in the male accessory gland, secreted into the sperm soup and transferred to the female together with the sperm during copulation (Herndon & Wolfner 1995). Inside the female oviduct Ovulin is processed by components supplied by both male and female (Park & Wolfner 1995). The processed protein then stimulates egg-laying (Heifetz et al. 2000). The hyper-variability of TrxT’s C-terminal domain suggests that TrxT might have a connection to sexual conflict, possibly having a role in regulating seminal fluid proteins or protecting them from premature cysteine oxidation. A possible role for TrxT in regulation or protection of seminal fluid proteins would however require TrxT either to be incorporated into the sperm tail or to be transferred to the sperm soup in some other way. We have so far no indications that that is the case, but it would certainly be an interesting question to pursue.

Regulatory motifs in the region between TrxT and dhd
With our TrxT-EYPF and DHD-ECFP constructs we have previously shown that the short region between TrxT and dhd is enough to ensure correct expression of both genes (Paper I). This means that the regulatory sequences of TrxT and dhd essentially have to be situated somewhere within this region. We decided to investigate this by searching for conserved motifs in the 150 first bases upstream of TrxT and dhd in different Drosophila species. We found two conserved candidate motifs that were specific for TrxT and three that were specific for dhd.
When comparing the motifs with upstream regions of other testis and ovary specific genes we could essentially not find the motifs in any of the examined genes. This could suggest that the TrxT-dhd gene pair is unique when it comes to their regulatory motifs and that their regulatory motifs might be specific for these particular genes. It would however be quite surprising since Perezgasga and coworkers (2004) have shown that transcription of TrxT is regulated by the meiotic arrest genes topi and comr. Both of these genes are believed to be part of a complex with proteins encoded by the meiotic arrest genes aly and achi/vis (White-Cooper et al. 2000; Ayyar et al. 2003). This complex regulates the transcription of approximately one thousand genes in the testes (Perezgasga et al. 2004). Nothing is known, so far, about what genes might be regulating dhd. It should be stressed that it is possible that the true regulatory motifs of TrxT and dhd are not at all identical to the conserved motifs that we have found. The true motifs might for example be too degenerated between the species for us to find them with this method. It is also possible that transcription factors regulating these genes are varying between different Drosophilids, even though we see the same expression pattern of TrxT and dhd in all of the species we have examined.

**Pof in genus Drosophila suggests autosome-specific gene regulation (Paper IV)**

For many years it has been widely known that the male X chromosome in Drosophila is “painted” by the components of the dosage compensation complex. This phenomenon of “chromosome painting”, where a protein or proteins bind a chromosome along its entire length, was for a long time believed to be exclusive for the X chromosome, until it was shown that the 4th chromosome in D. melanogaster is painted by the protein encoded by the Painting of fourth gene (Larsson et al. 2001). In Drosophila busckii, a related species where the F-element (the general designation for 4th chromosome corresponding element in other fruit flies) is part of the X, POF was shown to bind the entire X chromosome in males, suggesting a connection of Pof to the dosage compensation complex in D. busckii. To investigate this connection further we stained male and female polytene chromosomes from Scaptodrosophila lebanonensis and from several species in genus Drosophila with antibodies against POF, MSL3 and acetylated H4K16 (Paper IV). We find H4K16 acetylation (a histone modification specific to the DCC-component MOF) on the male X in all of the examined species and MSL3 staining in all species except S. lebanonensis and D. busckii. The fact that H4K16 acetylation is found in all of the examined species suggests that a similar dosage compensation system is used, even though MLS3 can not be recognized by our antibody in two of the species. In the distantly related S. lebanonensis, where the
RESULTS AND DISCUSSION

F-element is part of the X just as in *D. busckii*, we find no POF staining. Neither can we find any POF staining in *D. willistoni*, where the F-element is part of one of the autosomes. Staining of *D. busckii* polytene chromosomes reveals that POF colocalizes with H4K16 acetylation, strengthening the hypothesis that POF is connected to dosage compensation in certain species. In all examined species where the F-element exists as a unique dot chromosome (as in *D. melanogaster*) POF binds to the F-element in both sexes, but in two cases (*D. ananassae* and *D. malerkotliana*) POF also binds to the male X. Double staining with antibodies against POF and MSL3 in *D. ananassae* reveals that the two proteins colocalize on the male X, while POF is the only one that stains the F-element. To determine whether the POF’s chromosome-specificity depends upon the protein sequence of POF in that particular species, the *Pof*-gene from *D. ananassae* was chosen and a *DaPof-EYFP* construct was made and injected into *D. melanogaster* embryos. *DaPOF-EYFP* stains the 4th chromosome, but it does not locate to the male X in *D. melanogaster*. This ultimately suggests that POF has had a connection to, and has possibly even been part of, the dosage compensation complex during *Drosophila* evolution and that POF targeting to a specific chromosome is determined by an external source, for instance another protein or a non-coding RNA, and not by the POF protein sequence *per se*.

Deletion mutants in the *Pof* promoter region are lethal

Staining of POF in various different cell types revealed that POF is located to the 4th chromosome in generally all cells (Paper IV). To investigate whether POF has an essential function we created deletion mutations in the *Pof* gene region using P-element excision. All deletions in the *Pof* promoter region resulted in lethality and female sterility (as seen in germline clones), which led us to draw the conclusion that *Pof* is essential for fly viability. This is however not the case. At the time when the mutants reported in Paper IV were created, it was not known that a second gene is located in the *Pof* promoter region. Later releases of the *D. melanogaster* genome annotation indicated that a gene designated *CG33228* was located in this region. The *CG33228* gene organization and mRNA sequence was characterized and confirmed by Johansson and colleagues (2007). They constructed new deletion mutants in the coding region of *Pof* and could conclude that *CG33228* deletion mutants are lethal and female sterile, while *Pof* deletion mutants are viable. This may at first have seemed as a setback, but when they crossed haplo-4th flies with Pof-mutant flies they found that a complete lack of POF combined with only one copy of chromosome 4 leads to lethality. This phenotype could be rescued by a transgenic Pof-construct (Johansson et al. 2007). These findings are very interesting since the haplosufficiency of *Drosophila’s* 4th chromosome has previously been explained by its small size and low number of genes. The results clearly suggest that *Pof* has an important function in regulating
RESULTS AND DISCUSSION

the 4th chromosome in Drosophila. It also adds credibility to the hypothesis that POF has an evolutionary connection to the MSL complex.

Pof in spermatogenesis (Paper V)

Previous publications on Painting of fourth have indicated that expression of Pof RNA and POF protein is high in the male gonads (Larsson et al. 2001). We were intrigued by this since the X chromosomal MSL-complex is not present in the male germline, even though it has been shown that the male X is dosage compensated also in the germline (Rastelli & Kuroda 1998; Gupta et al. 2006). To look for POF expression in the male germline we stained male testes with an antibody against POF and also looked at testes from males carrying a POF-EYFP expressing construct. We found that POF expression is concentrated to a few small foci in early primary spermatocyte nuclei and that the number of foci increase the older the primary spermatocytes get and POF staining eventually becomes evenly distributed within the nucleus (Paper V). The POF localization is lost during meiosis, but returns during the spermatid maturation stages, culminating with a strong staining in the highly condensed elongated spermatid nuclei, just prior to individualization. No POF staining is seen in mature sperms. It is still unknown what structures POF might be binding to in the primary spermatocytes and in the elongated spermatid nuclei. The 4th chromosome is generally believed to localize together with the X chromosome within the spermatocyte nuclei, but it is also possible that the 4th chromosome is de-heterochromatinized, just as the normally heterochromatic Y chromosome that forms large decondensed structures (the Y-loops) in spermatocyte nuclei.

Pof upregulates chromosome 4 genes in the male testes

It is tempting to speculate on what function Pof might be serving in the male germline. A function for Pof in regulating genes on the 4th chromosome in the male germline is obviously the first thing that comes to mind. This idea might not seem very controversial at first glance, considering previous results on Pof, but when considering that the MSL-complex is not expressed in the testes, it becomes increasingly so. It is also possible that POF is regulating genes on the X chromosome, the Y chromosome or particular groups of genes that are specifically regulated in a certain manner in the testes etc. To test this idea we performed a microarray analysis on testis tissue from wildtype and Pof mutant males. The genes examined were categorized into twelve groups: six chromosomes (X, 2L, 2R, 3L, 3R and 4), heterochromatic chromosome regions (Xh, 2h, 3h and Yh), genes with unknown location and a subset of testis specific genes. We found that a lack of Pof in the male germline causes a global down-regulation of genes on the 4th.
RESULTS AND DISCUSSION

chromosome when compared to levels in wildtype testes. We could find no significant up- or down-regulation of any of the other examined groups. These results are completely in line with the microarray findings of Johansson and co-workers (2007), where they show that loss of Pof causes a global down-regulation of the 4th chromosome genes in first instar larvae. Intriguingly, they also show that a loss of heterochromatin protein-1 (HP1) causes a global up-regulation of genes on chromosome 4 and that binding of POF and HP1 to chromosome 4 is interdependent (Johansson et al. 2007). This suggests a balancing mechanism to ensure correct expression levels of chromosome 4 genes, involving POF and HP1 as partners with opposite functions: up-regulation performed by POF and down-regulation performed by HP1. The proposed balancing mechanism is reminiscent of the proposed dosage compensation mechanisms in mammals and C. elegans, where initial global up-regulation of the X chromosomes is balanced by inactivation of one X in females and down-regulation of both X:s in hermaphrodites respectively. This similarity gives more credibility to the idea that the twofold up-regulation of the X chromosome in Drosophila males performed by the MSL-complex is being balanced by a, so far unknown, protein system or complex.

An interesting phenomenon specific to the X chromosome in the male germline is the condensation and/or inactivation of the X chromosomes earlier than the autosomes prior to meiosis. This phenomenon is called precocious X inactivation and has been observed in many organisms, including C. elegans, Drosophila and mammals (Lifschytz & Lindsley 1972; Wu & Xu 2003). Precocious X inactivation seems to be limited to organisms where males are the heterogametic sex, since it has not been observed in either birds or butterflies, where it is the females that are heterogametic (Wu & Xu 2003). Interestingly, in Drosophila, translocations between the X and the autosomes result in dominant male sterility in more than 75% of the cases, but no male sterility is seen in translocations between the 4th and X (Lindsley 1965; Lindsley & Tokuyasu 1980). It has been proposed that X:A specific sterility is caused by a disturbance in the precocious X inactivation by the translocated chromosome. The disturbance could either be that the translocated autosome becomes inactivated too soon or that the precocious inactivation of the translocated X is delayed or stopped. The reason why X:4 translocations do not cause male sterility has been believed to be either that the 4th is too small to influence the precocious inactivation or that the 4th has different regulatory properties than the other autosomes (Lindsley & Tokuyasu 1980). The 4th chromosome is decidedly more heterochromatic than the other autosomes, and so the heterochromatic state of the 4th might be too similar to the inactivated state during meiosis to be able to exert any activating influence over the translocated X. But what of the translocated 4th? Since POF is up-regulating chromosome 4 genes also in the male germline, it is definitely possible that POF protects against the
RESULTS AND DISCUSSION

inactivating effect of the translocated X. This hypothesis would be very interesting to test by introducing X:4 translocations in a Pof mutant background. If the X:4 males are sterile in Pof mutant background it would suggest that POF protects the 4th chromosome from improper inactivation. This would add even more support for the notion of POF as an important player in chromosome-wide gene regulation.

NEW RESULTS

Thioredoxins can break Zeste aggregates

All the three Drosophila thioredoxins have previously been shown to be present in nuclei of the germline (Paper I). We were interested in the possibility that thioredoxins may play a role in cellular memory or gene regulation in the Drosophila germline, and since TrxT was first identified in a yeast-two-hybrid screen for Zeste-interacting proteins (Chen 1992), we decided to investigate the possible interactions between Zeste and thioredoxins. Pirrotta and colleagues have shown that Zeste forms large aggregates both in vitro (Bickel & Pirrotta 1990) and in vivo (Chen & Pirrotta 1993a), and its aggregation is believed to be an important element of the mechanism underlying many of the zeste mutant phenotypes seen in vivo. In our initial assays of the interactions between thioredoxins and Zeste, TrxT was shown to dramatically increase the mobility of a Zeste/DNA complex when separated by gel electrophoresis (results not shown). This led us to speculate that thioredoxins might be able to break Zeste aggregates. Our approach to test this hypothesis was to translate zeste and the three Drosophila thioredoxins in vitro, mix Zeste with TrxT, DHD, Trx-2 or a mock sample separately, and to test the aggregation capacity of Zeste under these four different conditions in a pelletation assay previously developed to monitor Zeste aggregation (Bickel & Pirrotta 1990; Rosen et al. 1998). In this assay, aggregated Zeste protein will form a pellet, while non-aggregated Zeste remains in the supernatant. We found that DHD could clearly break Zeste aggregates in vitro, as seen from the low amount of Zeste in the pellet, while both TrxT and Trx-2 had much weaker effects in this assay (Figure 7A).

The C-terminal hydrophobic region of Zeste has been shown to be necessary for aggregation of the protein (Chen et al. 1992; Chen & Pirrotta 1993a-b). This hydrophobic region is predicted to have an alpha-helical structure and three leucine zippers, two of them aligned in a row on one side and one on the opposite side. Aggregation of Zeste has been proposed to involve an initial dimerization followed by multimerization of the dimers. There are only two cysteines in this hydrophobic region, one situated two amino acids from the C-terminus, the other
situated approximately in the middle of the hydrophobic region, at the end of one of the leucine zippers (Figure 6). We hypothesized that the two cysteines located in this hydrophobic region are involved in the aggregation process by forming disulfide bonds, thereby stabilizing the aggregation, and that the thioredoxins can break Zeste aggregates by reducing these disulfide bonds.

![Figure 6](image-url) Hypothetical aggregation model of Zeste’s hydrophobic region. Two Zeste molecules dimerize at the double leucine zipper. The C-terminal cysteine could act as a stabilizer. Zeste dimers form aggregates through multimerization at the other leucine zipper, where the second cysteine could act as another stabilizer.

To test this hypothesis we created zeste clones that were mutated at one or both of the positions encoding these cysteines. We then performed the same pelletation assay as above, but this time with the four different versions of Zeste (wildtype Z, Zc520s, Zc571s and Zc520sc571s) and with only a mock sample or DHD added. Both mutant Zeste forms aggregated much more weakly than the wild type protein when mock sample was added, and their low degrees of aggregation could be completely inhibited by adding DHD. When both single mutants were mixed in a sample, aggregation was also weaker than wild type, but stronger than with either of the single mutants alone. However, the strongest inhibition was observed in assays with the double mutant, in which virtually no aggregation was detected (Figure 7B). We conclude that thioredoxins can affect Zeste aggregation and that the two cysteines c520 and c571 are important for the Zeste aggregation process, which provides a possible mechanistic link to the thioredoxins. Further support and proof for the hypothesized aggregation model remains to be shown.
Figure 7. Zeste – thioredoxin interactions. A. Thioredoxins can break Zeste aggregates \textit{in vitro}. Proteins were translated \textit{in vitro} and used in a pelleting assay separating the aggregated form of Zeste in the pellet (P) from the non aggregated form in the supernatant (S). The amount of Zeste in the pellet is indicated as percent of total. The panels are PhosphorImager captures of SDS-PAGE gels. Added thioredoxin or mock protein is indicated. B. Zeste proteins with cysteine substitutions aggregate less than wild type Zeste \textit{in vitro}. The same method was used as in A.
CONCLUSIONS

- *TrxT* encodes a testis-specific thioredoxin that specifically locates to the Y-chromosome loops *ks-1* and *kl-5* in primary spermatocytes.

- *dhd* encodes an ovary-specific thioredoxin that specifically locates to nurse cell nuclei and the oocyte karyosome.

- *Trx-2* mutants are viable but have shorter life spans than control flies.

- *Trx-2* mutants are as sensitive to oxidative stress as control flies, but over-expression of *Trx-2* mediates an increased resistance to oxidative stress.

- All three thioredoxin genes can be removed without causing lethality or any other strong phenotypes.

- The genomic organization of *TrxT* and *dhd* is conserved in ten out of twelve Drosophilid species.

- The C-terminal domain of *TrxT* is hyper variable.

- POF binding to the F-element (chromosome 4 in *D. melanogaster*) is conserved in genus *Drosophila*.

- POF colocalizes with MSL-3 and H4K16ac in species where POF binds the male X.

- POF is present in the nuclei of primary spermatocytes, but also in nuclei of maturing spermatids.

- Lack of POF in the male germline causes a global down-regulation of chromosome 4 genes.

- Thioredoxins can break Zeste aggregates *in vitro* and Zeste proteins that lack the two cysteines in the hydrophobic domain have a very low aggregation capacity.
SUMMARY IN SWEDISH


**Gen-namnet Painting of fourth** härstammar från upptäckten att POF binder till ("målar") *Drosophilas* kromosom 4. Jag visar i min avhandling att POFs bindning till den fjärde kromosomen är bevarad i olika *Drosophila*-arter och att POF kolokaliserar med både ett protein och en histon-modifiering, som är förknippade med doskompensation, i arter där POF också binder till hanens X-kromosom. POF uttrycks överallt i både honor och hanar, men i väldigt höga nivåer i hanens testiklar. Jag visar här att POF finns i cellkärnan hos primära spermatocyter, men också i kärnan på mognande spermatider, och att avsaknad av POF in hanens könsceller orsakar en global nedreglering av gener som ligger på kromosom 4. Kombinationen av mina POF-resultat tyder på att POF har en viktig funktion i det första kända fallet av genreglering av en hel autosomal kromosom.
ACKNOWLEDGEMENTS

I would like to thank all the people who have contributed in some way to this thesis, or been important to me and the journey that has brought me to this point in time:

Janne, my wonderful supervisor. You are a brilliant scientist AND you have a lovely family and a social life! You are of a rare kind. ;-) You have done a fantastic job through all my PhD-student years, giving me a kick when I needed it, supporting me whenever I was low and letting me spread my wings just at the right time. Thank you also for all your good advice regarding this thesis. I don’t think I could have had a better supervisor than you!

Per, my fellow trekker (or is it trekkie?). It was as small step for humanity, but a great step for GroupJL when you joined us. ;-) It has been great to have someone to share my movie and music taste with! You are one of the most amusing people I know, and a fine scientist.

Anna-Mia, my lab-pal, my office-buddy, my IKSU-mate. From the moment you set your foot in our office I knew you would fit right in. You are a true friend and I have appreciated our friendship very much! And I will definitely be dropping by some time in the future to check out your beautiful house. ;-)  

Karin, our fly-wizard. Thank you for your excellent fly work and great party spirit. Work would have been both less fun and more cumbersome without you.

Project students and past members of group JL: Thomas, Mikael, Maria, Carolina, Elin, Theresa and the latest addition to group JL, Lina. You all contributed to the positive atmosphere in group JL. I would like to especially thank Theresa for performing the major part of that huge fertility-experiment. You did a fantastic job! (And thank you Lina for solving my apartment renting problem!)

Thanks to Per, Simon and Åsa for constructive criticism and reading of this thesis.

Anssi, thank you for sharing your never ending knowledge and personal views on science, religion and everything else and thank you for all the invaluable help I have received during the years!

Liz and Åsa for constructive criticism and encouragement during our post-Floortalk-committee meetings. And thank you for the fact that you were there as “safety nets” if I ever would have needed it.

All the UCMPers: Our fellow fly groups RHP (thank you Caroline for making me feel welcome into ‘the Clan’) and DH, the Worm group and the two X-Ray groups (Anders O and Malin LP, thanks for nice company in the office). Thanks for making UCMP such an incredibly fun and friendly place to work! I want to especially thank Maria Westling for your invaluable help with everything during my years at UCMP.
Everybody at the old genetics department: Anna, Bettan, Gunilla (tack för superbra service med flugmaten), Helena, Jesper, Kerstin, Magnus, Marcus, Marianne, Mikael, Sa, Stefan (tack för att jag fick dela rum med dig när jag skrev avhandlingen och för att du ställer upp som toastmaster på festen) and everybody else. Thanks for all the incredibly fun genetics parties and nice company during lunches and coffee breaks in the ‘good old days’!

Jenny, Åsa och Elisabeth mina kompisar från biologprogrammet, ni gjorde grundutbildningen väldigt trevlig! Vi hade nog alla fyra en ganska omvändande tid då. ;-) Tack särskilt till Jenny och Jörgen för att jag fått hyra er lägenhet, det har varit super!

Anki, jag vet inte om det var helt optimalt för betygen och studiemoralen att vi hängde ihop som ler och långtarm på gymnasiet, men det var desto bättre för själern. Tack för att jag fått lära känna Bertram och Elliot! Du kommer alltid vara mina allra bästa vän.

ACC-folket, tack för alla trevliga tillställningar, av alla de slag. Jag vet, jag borde ha skrivit min avhandling i LaTeX. ;-) Nya och gamla vänner i Fjollträsk, tack vare er känns det inte så läskigt att flytta söderut.

Ulrika Östh, min klassföreståndare på högstadiet. Du har en viktig del i att jag alls har kommit så långt. Men minns du den där fäniga uppsatsen jag skrev? Sometimes life is stranger than fiction.

Daniel, jag lärde mig mycket om mig själv och världen under vår tid tillsammans. Jag är väldigt glad att vi förblivit goda vänner och jag kommer sakna våra trevliga sci-fi-kvällar. Är du säker på att du inte ska flytta söderut du med? ;-) 

Karin och Bertil, ni har verkligen fått mig att känna mig som en i familjen. Tack för all hjälp! Jag ser fram emot många besöker av er och oss, i Stockholm, Umeå och Bakviken.

Mamma och Pappa, tack för att ni alltid ställer upp för mig och att ni stöttar mig och låter mig vara den jag är! Hille och Jonas, vi hör av alldeles för sällan. Nu när jag flyttar lite närmare hoppas jag vi kan hälsa på varandra oftare. Jag älskar er allihop! Tack oxå till övriga släkten för att ni bidragit till att göra min uppväxt lycklig och ljus!

Och framförallt min egen personliga isbjörn. Jag vet att du var lite ledsen för mig och hjälpa mig med alla små vardagsbestyr som jag inte hade tid med när jag skrev avhandlingen. Men vet du, det räckte med att du fanns där i tanken (och några telefonsignaler bort). Din kärlek får mig att tro på min egen förmåga. Nu är snart mer än tre års väntan äntligen över... Jag tokålskar dig! ;-) 

Live Long and Prosper!
REFERENCES


REFERENCES


REFERENCES


REFERENCES


REFERENCES


REFERENCES


REFERENCES


REFERENCES


REFERENCES


REFERENCES


REFERENCES


