NOTE TO USERS

This reproduction is the best copy available.
Identification and Characterization of the *tantalus* Gene from

*Drosophila melanogaster*

by

Bruce Dietrich

A thesis submitted in conformity with the requirements for the degree of Doctor of Philosophy
Graduate Department of Molecular and Medical Genetics
University of Toronto

© Copyright by Bruce Dietrich 2001
The author has granted a non-exclusive licence allowing the National Library of Canada to reproduce, loan, distribute or sell copies of this thesis in microform, paper or electronic formats.

The author retains ownership of the copyright in this thesis. Neither the thesis nor substantial extracts from it may be printed or otherwise reproduced without the author's permission.

L’auteur a accordé une licence non exclusive permettant à la Bibliothèque nationale du Canada de reproduire, prêter, distribuer ou vendre des copies de cette thèse sous la forme de microfiche/film, de reproduction sur papier ou sur format électronique.

L’auteur conserve la propriété du droit d’auteur qui protège cette thèse. Ni la thèse ni des extraits substantiels de celle-ci ne doivent être imprimés ou autrement reproduits sans son autorisation.

0-612-59020-8
Identification and Characterization of the tantalus Gene from *Drosophila melanogaster*

Doctor of Philosophy, 2001

Bruce Dietrich

Department of Molecular and Medical Genetics, University of Toronto

Abstract

One longstanding and thought-provoking question in developmental biology is how an early egg achieves asymmetry, and how this asymmetry is interpreted to produce pattern in the adult organism. By choosing to study development in the fruit fly *Drosophila melanogaster*, I had hoped to understand and contribute to the growing field of knowledge that was unraveling these basic questions. My original goal of understanding the interpretation of asymmetry has been extended to encompass the mechanisms involved in the maintenance of this asymmetry.

This thesis concerns the discovery and analysis of the tantalus (tan) gene identified in a yeast two-hybrid screen as a protein interactor for the *fushi tarazu* protein, a pair-rule gene involved in asymmetry interpretation. *tan* was also independently identified in the lab of Dr. Hugh Brock at the University of British Columbia as an interactor for the *Additional sex combs (Asx)* protein, a gene involved in asymmetry maintenance.

Approximately 50% of *Drosophila* genes appear to be unique to this species and *tan* would seem to fall into this category, as *tan* is not homologous to any other identified genes. To understand the role of *tan* during development, the expression of both the gene
and protein have been followed, and a null tan allele was created. Studies of over- and under expression of tan suggest that the gene functions in a tissue-specific manner. Interestingly, Asx also has tissue-specific activity, and a collaboration between our lab and Dr. Brock's lab has demonstrated a physical and genetical interaction between tan and Asx. These results have led to the proposal that TAN acts as a tissue-specific cofactor for ASX.
Acknowledgments

I would like to thank my supervisor Henry Krause for his guidance and patience, and my supervisory committee members Brenda Andrews, Howard Lipshitz, and Arthur Roach for their ideas and encouragement. I would also like to thank Hugh Brock for his valued collaboration.

My introduction into lab life was made much easier by the generosity and invaluable assistance of John Copeland, Andrzej Nasiadka, and Robert Strome. I am very grateful for the time, effort, and friendship they provided to me while I was learning that ethanol need not be sterilized. I would also like to thank past and present lab members for their help throughout the duration of my stay in the lab. Andrew Simmonds, in particular, has been an invaluable source of computer knowledge for our lab, and could always be counted on when our computers could not be.

I would also like to thank my parents, family, and friends who provided much support and encouragement, even though the nuances of the last eight years of my scientific life escaped them. Finally, I must thank my wonderful wife, Toolika, who has been everything that one could hope for in a companion. Throughout the frustration, there was always you.
Table of Contents

Abstract ii

Acknowledgments iv

Table of Contents v

List of Tables x

List of Figures xi

List of Abbreviations xiii

Genes xiii

Other Abbreviations xiv

CHAPTER 1-Introduction 1

I. Drosophila Development 3

A) Synopsis of the Drosophila Life Cycle 3

B) Initial Asymmetry in the Oocyte 4

C) Translation of Axis Polarity 6

i.) Patterning the A-P Axis 6

ii.) Patterning the D-V Axis 9

D) From Asymmetry to Pre-pattern: Making Parasegments 12

i.) Periodicity 13

ii.) Defining Parasegments 15

E) Patterning the Larval Epidermis 17

i.) Defining Position Within the Parasegment 17

ii.) Providing Pattern Specificity to the Parasegment: the Homeotic Genes 20
II. Molecular Basis of Cell Fate Specification 22

A) The fushi tarazu Gene 22
B) The Homeobox 24

C) The Importance of the HD in FTZ Function 27
   i.) HD-Independent FTZ Activities 27
   ii.) Is the HD of FTZ Important for its Function? 31

D) Achieving Functional Specificity for HD-Containing Proteins 32
   i.) a1, α2, and MCM1 Function in Yeast 33
   ii.) extradenticle (exd) 35

III. Maintenance of Cell Fate Specification 38

A) The Polycomb and trithorax Group of Genes 38
   i.) The Polycomb Group 39
   ii.) The trithorax Group 43

B) The Additional sex combs Gene 47

IV. Thesis Outline 49

CHAPTER 2-Identification of Tantalus as a Fushi tarazu-Interacting Protein 62

Abstract 63

Introduction 63

Results and Discussion 66

Screening for Fushi tarazu interacting proteins 66
Redundancy in Tantalus function

Homeodomain binding of Tantalus

Conclusions

REFERENCES
List of Tables

Table 2-1. Results of the yeast two-hybrid screen. 81

Table 3-1. Polytene chromosome binding sites for TAN, with ASX and PC/PH sites indicated. 115

Table 3-2. Analysis of tan mutant and rescued phenotypes. 116

Table 3-3. tan and Asx interact genetically in interocellar bristle specification. 117

Table 3-4. Genetic interaction between tan and N. 118
| Figure 1-1. | The *Drosophila melanogaster* life cycle. | 51 |
| Figure 1-2. | Embryogenesis. | 52 |
| Figure 1-3. | A-P and D-V axes specification during oogenesis. | 53 |
| Figure 1-4. | Hierarchy of genes involved in A-P patterning. | 54 |
| Figure 1-5. | Domains of gap gene expression. | 55 |
| Figure 1-6. | D-V axis specification. | 56 |
| Figure 1-7. | Parasegments versus segments. | 57 |
| Figure 1-8. | *fushi tarazu* expression. | 58 |
| Figure 1-9. | Expression of *even-skipped* stripe 2. | 59 |
| Figure 1-10. | Expression pattern of selected pair-rule and segment polarity genes. | 60 |
| Figure 1-11. | Specification in the larval epidermis. | 61 |
| Figure 2-1. | Schematic diagram of the yeast two-hybrid screen. | 82 |
| Figure 2-2. | Bait constructs used in yeast two-hybrid screen. | 83 |
| Figure 2-3. | Far Western assay. | 84 |
| Figure 2-4. | Cloning *tan* from cDNA and genomic libraries. | 85 |
| Figure 2-5. | Genomic and amino acid sequence of *tan*. | 86 |
| Figure 2-6. | Polytene in situ hybridizations. | 87 |
| Figure 2-7. | Northern blot analysis. | 88 |
| Figure 2-8. | Expression pattern of *tan*. | 89 |
| Figure 2-9. | An epitope tagged version of TAN localizes to the nucleus. | 90 |
Figure 2-10. TAN binds DNA. 91

Figure 3-1. GST pull-down experiment. 119

Figure 3-2. Western blots. 120

Figure 3-3. TAN binding to polytene chromosomes. 121

Figure 3-4. In vivo detection of TAN. 122

Figure 3-5. P-element mutagenesis of tan. 123

Figure 3-6. Adult defects in tan^2 homozygotes. 124

Figure 3-7. Ectopic and over expression of tan disrupts sensory lineages. 125

Figure 3-8. Ectopic expression of tan in the wing margin. 126

Figure 4-1. Model of TAN function. 135
List of Abbreviations

Genes

ANT-C  Antennapedia complex
Antp  Antennapedia
Asx  Additional sex combs
bcd  bicoid
brm  brahma
BX-C  bithorax complex
cac  cactus
cad  caudal
Dfd  Deformed
dl  dorsal
dpp  decapentaplegic
EGFR  Epidermal growth factor receptor
en  engrailed
eve  even-skipped
exd  extradenticle
E(z)  Enhancer of zeste
ftz  fushi tarazu
grk  gurken
gt  giant
h  hairy
hb  hunchback
hh  hedgehog
hkb  huckebein
hth  homothorax
kni  knirps
Kr  Krüppel
luc  luciferase
MCMI  Minichromosome maintenance
Meis  Myeloid ecotopic insertion site
N  Notch
nos  nanos
odd  odd-skipped
opa  odd-paired
otd  orthodenticle
phx  pre-B cell homeobox 1
Pc  Polycomb
ph  polyhomeotic
pho  pleiohomeotic
prd  paired
rhomboid
Sex combs reduced
Serrate
Sloppy paired 1 & 2
snail
spätzle
Suppressor of variegation 205
tantalus
Transforming growth factor-α
Transforming growth factor-β
tailless
torpedo
torso
trithorax
twist
Ultrabithorax
veinlet
wingless

Other Abbreviations

AEL
After egg laying
AF-2
Activation function-2
A-P
Anterior-posterior
CNS
Central nervous system
D-V
Dorsal-ventral
EST
Expressed sequence tag
ETP
Enhancers of trithorax and Polycomb
GOF
Gain of function
HD (ΔHD)
Homeodomain (deleted Homeodomain)
HOM-C
Homeotic complexes
HP1
Heterochromatin protein 1
HS
Heat shock
hsp70
Heat shock promoter 70
LOF
Loss of function
PcG
Polycomb group
PEV
Position effect variegation
PNS
Peripheral nervous system
PRE
PcG response element
SOP
Sensory organ precursor
TRE
Trithorax response element
trxG
Trithorax group
WT
Wild-type
CHAPTER 1

Introduction

A great success of developmental biology has been the fusion of the long history of embryological observation with the power of molecular biology. In particular, scientists have pushed classical genetics out of the realm of mathematics and into the realm of cell biology, whereby the developmental function of the gene could be deduced. Without question *Drosophila melanogaster* has been a key experimental organism that has allowed this to occur. While the groundbreaking study of phage and bacteria in the middle of the last century allowed us to understand the nature of the gene, *Drosophila* has led the way in extending the concept of the gene to an understanding of development and, ultimately, an explanation of the evolution of animal complexity.

The contribution of *Drosophila* research began over 80 years ago, and had immediate impact when Thomas Hunt Morgan and his students provided the first definitive evidence for the chromosomal theory of inheritance. The recent sequencing of the *Drosophila* genome (Adams et al., 2000), a monumental achievement surpassed only by the sequencing of the human genome, provides a fulfilling bookend to Morgan’s feat. Most interestingly, the sequencing projects have emphasized the similarity between *Drosophila* and human genes, ensuring that *Drosophila* research will not only continue to shed light on the evolution of life on earth but also throw insight onto what we, rightly or wrongly, have decided to be the most important result of that evolution: ourselves.

In the last 20 years, development of the *Drosophila* embryo has been dissected at
the genetic and molecular level to give us one of the most complete and elegant explanations of how complex patterns derive from a single cell. This advance was initiated by several large scale genetic screens (Nüsslein-Volhard and Wieschaus, 1980; Jürgens et al., 1984; Nüsslein-Volhard et al., 1984; Wieschaus et al., 1984; Schüpbach and Wieschaus, 1986; Schüpbach and Wieschaus, 1989) which sought to determine how the anterior-posterior (A-P) and dorsal-ventral (D-V) axes are designated, and how segments are formed and specified once these axes are demarcated.

This Introduction is divided into three sections. The first section outlines early developmental events in the Drosophila embryo leading to the patterned larva. This description incorporates the work of hundreds of researchers and represents a monumental achievement in the study of development; for the first time in any organism a detailed knowledge of the events leading from oocyte selection to adult pattern has been obtained. Although some of the specifics are unique to Drosophila, many fundamental underlying principles have been deduced. These include universal signaling cascades and morphogens responsible for translating axis information into pattern. In the next two sections I review the events surrounding asymmetry interpretation and maintenance in greater detail: the second section focuses on the role of the ftz protein in interpreting asymmetry, while the third section discusses the Asx gene in the context of global mechanisms used to maintain states of transcription.
I. *Drosophila* Development

A) Synopsis of the *Drosophila* Life Cycle

The Krause laboratory focuses on the events of early development (reviewed in Wolpert et al., 1998), providing a fitting place to enter the life cycle of *Drosophila* (Figure 1-1). After the egg is fertilized the embryo undergoes several rapid rounds of mitotic division without cytoplasmic cleavage (Figure 1-2). The nuclei then migrate to the periphery of the egg and become cellularized by membranes pinching in from the egg surface. The migration of nuclei towards the surface begins after 9 mitotic divisions and by the end of 13 divisions, three hours after egg laying (AEL), cellularization is complete. During this time, the germ anlagen are specified from a group of approximately 15 nuclei derived from the posterior end of the blastoderm embryo. As *Drosophila melanogaster* is a long germ band insect, all future segments are defined by the end of the blastoderm stage.

During gastrulation, the epithelial layer from the blastoderm embryo gives rise to the mesoderm, ectoderm and, at the anterior and posterior ends of the gastrulating tissue, the endoderm. Gastrulation is followed by an extension and retraction of the germ band around the posterior end of the embryo, to the dorsal side. By the end of retraction the three head, three thoracic, and eight abdominal segments are clearly demarcated. During this time the imaginal discs, which will eventually give rise to the adult structures, are also set aside. Approximately 24 hrs AEL the larvae hatches, and after two more larval instars pupation occurs. During pupation the imaginal discs differentiate into the adult structures of the fly, which ecloses at day 10 AEL.
B) Initial Asymmetry in the Oocyte

The mother plays a major role in producing the asymmetry that specifies the axes of the developing oocyte during oogenesis. Initially, an asymmetric division of a germ-line stem cell in the germarium of the female ovary produces a new stem cell and a cystoblast (reviewed in Grüntert and St. Johnston, 1996). The cystoblast undergoes four mitotic divisions without cytokinesis, resulting in 16 cells attached by ring canals. Only two of these cells will have four such canals and both become pro-oocytes and initiate meiosis. However, only one is selected to become the oocyte. This selection may involve the asymmetric segregation of cellular components during the initial division of the cystoblast (de Cuevas and Spradling, 1998). The other pro-oocyte and remaining 14 cells become nurse cells, which deposit large quantities of RNA and protein into the egg. During the movement of the 16 cell cyst through the germarium it becomes enveloped by somatic follicle cells, key components of polarity formation. After the oocyte is selected, the first sign of asymmetry is increased levels of the cell adhesion molecule, DE-cadherin, in posterior follicle cells (Godt and Tepass, 1998; González-Reyes and St. Johnston, 1998). DE-cadherin upregulation causes migration of the oocyte to the posterior end of the egg chamber.

The posterior position of the oocyte and oocyte nucleus within the cyst allows the posterior end of the egg to be specified first (reviewed in Ray and Schüpbach, 1996; van Eeden and St. Johnston, 1999); (Figure 1-3). The oocyte signals to the surrounding follicle cells by releasing a ligand encoded by gurken (grk), a member of the transforming growth factor-α (TGF-α) family (González-Reyes et al., 1995; Roth et al., 1995). GRK
binds to the posteriorly located follicle cells through the Torpedo (TOP) receptor, a homologue to the epidermal growth factor receptor (EGFR). Activated TOP then sends a return signal to the oocyte, which causes a major change in the polarity of the oocyte microtubule network (Lane and Kalderon, 1994; Ruohola et al., 1991). This polarity reversal is the key event that specifies the A-P axis: it is required for the localization of mRNAs that will determine the anterior end of the embryo (see below), and also for the re-positioning of the oocyte nucleus to the anterior end of the egg, an event essential for D-V axis determination (González-Reyes et al., 1995; Roth et al., 1995).

Microtubule reorganization allows the oocyte nucleus to move to a random anterior margin location (Theurkauf et al., 1992). Once there, GRK signals again to the overlying follicle cells through TOP to specify the dorsal side (Neuman-Silberberg and Schüpbach, 1994; Schüpbach, 1987). TOP activation in dorsal follicle cells prevents the expression of pipe, a gene required for modification of the Spätzle (SPZ) ligand, which in turn, is responsible for ventral fates (Morisato and Anderson, 1994; Sen et al., 1998). If the GRK/TOP pathway is inactivated during dorsal specification, active ventral ligand is produced everywhere, resulting in a ventralized embryo.

To summarize, axes designation in the oocyte uses DE-cadherin based cell adhesion to posteriorly locate the oocyte and oocyte nucleus within the 16 cell cyst. This positioning allows the GRK/TOP signaling cascade to reorganize the microtubule network and produce an A-P polarity in the egg. Microtubule reorganization also causes the nucleus to migrate to an anterior margin location where signaling specifies the D-V axis by limiting the release of the ventral-inducing ligand SPZ to the future ventral side.
Although the GRK signaling pathway nicely explains the formation of both axes, the cause of the original asymmetry in the posterior and dorsal follicle cells, allowing them to react differentially to the GRK ligand, is still not known (González-Reyes and St. Johnston, 1998). How the follicle cells achieve this asymmetry is one of the few unanswered questions of axis determination.

C) Translation of Axis Polarity

In order for the newly established oocyte polarity to lead to patterning in the egg, several maternal gene products must act as morphogens. Morphogens are defined as factors capable of specifying different cell fates at different concentrations along an axis. The positional values created by these morphogens in Drosophila are used to produce the limited spatial expression of the first zygotic genes along the axes. These initial zygotic genes are the first in a hierarchy of genes whose expression will translate the initial asymmetry in the A-P and the D-V axes into segments and germ layers (ectoderm and mesoderm), respectively (Figure 1-4).

i.) Patterning the A-P Axis

A-P specification requires three systems, unlike the D-V axis which requires only one (reviewed in St. Johnston and Nüsslein-Volhard, 1992). The anterior, posterior, and terminal systems are initiated independently, but work together to provide pattern to the A-P axis.

In the anterior system bicoid (bcd) mRNA is localized to the anterior end of the oocyte as a result of the microtubule reorganization events that occurred earlier during oogenesis (Berleth et al., 1988; Frigerio et al., 1986; St. Johnston et al., 1989), and is
translated after the egg is laid to produce a gradient of protein extending to the middle of the embryo (Driever and Nüsslein-Volhard, 1988a; Driever and Nüsslein-Volhard, 1988b); (Figure 1-4). BCD contains a homeodomain (HD) DNA binding motif (Berleth et al., 1988; Frigerio et al., 1986) and functions as a morphogen in anterior patterning (Driever and Nüsslein-Volhard, 1988b).

BCD regulates several gap genes, including *hunchback* (*hb*) and *orthodenticle* (*otd*), in the anterior end of the embryo. Gap genes are the first zygotic genes to be activated and encode transcription factors which pattern large regions of the embryo (Nüsslein-Volhard and Wieschaus, 1980); (Figure 1-5). The *hb* enhancer contains both strong and weak binding sites for BCD (Driever and Nüsslein-Volhard, 1989; Driever et al., 1989; Struhl et al., 1989) and, consequently, is expressed in a broad domain that mimics BCD expression (Tautz, 1988). The *otd* enhancer, on the other hand, has only weak BCD sites and is only activated in the most anterior regions of the embryo where BCD concentrations are highest (Gao and Finkelstein, 1998). In this way, BCD acts as a morphogen by limiting individual gene expression to specific regions along the A-P axis, based on the affinity of BCD for each target gene enhancer.

HB, also a transcription factor, appears to be the primary morphogen required to specify thoracic and abdominal fates (Hülskamp et al., 1990; Schulz and Tautz, 1994; Struhl et al., 1992; Wimmer et al., 2000), and acts in conjunction with BCD (which may be redundant) to directly regulate transcription of several additional gap genes (reviewed in Rivera-Pomar and Jäckle, 1996). For example, the gap gene *Krüppel* (*Kr*) is required in thoracic and abdominal segments (Nüsslein-Volhard and Wieschaus, 1980) and has its
expression boundaries set by both BCD and HB (Gaul and Jäckle, 1989; Hoch et al., 1992; Hoch et al., 1991; Hülskamp et al., 1990). Kr is activated at low concentrations of BCD and HB found in the middle of the embryo, but is repressed in more anterior regions. The anterior repression of Kr is presumably due to repressor proteins encoded for by other gap genes, like giant, which are activated by higher levels of BCD. The posterior limits of Kr expression are set by yet other gap genes activated by the posterior system. Therefore, because of the different affinities of BCD and HB for target gene enhancers and the ability of gap genes to cross-regulate their expression, several broad domains of gap gene activity can be established within the anterior region of the embryo.

Patterning in the posterior end of the embryo requires a posterior to anterior gradient of nanos (nos) protein (Gavis and Lehmann, 1992; Wang and Lehmann, 1991) to inhibit translation of maternally provided hb mRNA (Hülskamp et al., 1989; Irish et al., 1989; Struhl et al., 1989). In addition to its zygotic expression hb mRNA is deposited ubiquitously by the mother in the embryo, and it is the ubiquitous hb mRNA in the posterior end that NOS must counteract. Suppression of maternal HB in the posterior of the embryo allows the Caudal (CAD) transcription factor to function. CAD is found in a gradient emanating from the posterior end and is responsible for activating several gap genes in the posterior region (Rivera-Pomar et al., 1995), similar to the role of BCD in the anterior end. Interestingly, cad mRNA is also ubiquitously provided by the mother and has its gradient formed by the inhibition of its translation in the anterior end (Mlodzik et al., 1990). Surprisingly, this inhibition is performed by BCD, as BCD uses its HD to bind to 3’UTR sequences in the cad mRNA to prevent translation (Dubnau and Struhl,
Finally, in the termini, the most anterior and posterior follicle cells secrete a ligand for the Torso (TOR) receptor (reviewed in Duffy and Perrimon, 1994). TOR, a member of the tyrosine kinase class of receptors, is zygotically expressed and found throughout the surface of the embryo (Casanova and Struhl, 1989; Sprenger et al., 1989). The follicle cells at the poles retain the ligand in the perivitelline space until after fertilization, at which time it is released and binds to the receptor. The downstream result of TOR activation is derepression of the gap genes \textit{tailless} and \textit{huckebein} in the terminal regions (Liaw et al., 1995; Paroush et al., 1997; Rusch and Levine, 1994). These proteins limit expression of other gap genes in the termini of the embryo and pattern the terminal acron and telson structures, with BCD overlap making anterior structures different from posterior (Pignoni et al., 1990; Weigel et al., 1990).

To summarize, localized expression of BCD and NOS is required to effect the designation of the A-P axis by producing opposite gradients of the HB and CAD transcription factors. BCD/HB and CAD extend from the anterior and posterior poles, respectively, of the embryo and provide positional information along the axis. This positional information, in conjunction with signals from the terminal system, is interpreted to produce the specific aperiodic expression domains of the individual gap genes. The gap genes then cross-regulate each other, refining their expression domains into sharp on-off patterns.

\textit{ii.) Patterning the D-V Axis}

Although the D-V axis is specified using a different initial mechanism (signal
transduction) than the A-P axis (localized determinants) the end result is the same: the production of a morphogen to supply positional information along the axis (reviewed in Morisato and Anderson, 1995; Rusch and Levine, 1996). This morphogen is encoded by the *dorsal* (*dl*) gene (Roth et al., 1989), a homologue of the $NF-\kappa B$ gene of the IL-1R-$NF-\kappa B$ pathway found in both the plant and animal kingdoms (reviewed in Belvin and Anderson, 1996); (Figure 1-6). DL is an outstanding example of the ideal morphogen as it acts as both a repressor and activator, and induces expression of genes which themselves limit DL function. DL is ubiquitously provided by the mother but is sequestered in the cytoplasm through its association with the protein product of the *cactus* (*cact*) gene, CACT (Kidd, 1992; Wasserman, 1993; Whalen and Steward, 1993). CACT repression of DL function is lifted by signaling of the ventral ligand Spätzle (SPZ) to the ubiquitously expressed receptor encoded by *Toll* (Hashimoto et al., 1991; Hashimoto et al., 1988). The result of this signaling is the release of DL from CACT, allowing DL to translocate to the nucleus.

As active SPZ extends in a graded fashion from ventral regions, a similar graded distribution of nuclear DL is produced. This graded distribution allows DL to play a role in specifying all fates (four in total) along the D-V axis, either directly or indirectly. In ventral regions high levels of DL cause the expression of *twist* (*twi*) and *snail* (*sna*), two genes required for gastrulation and mesoderm specification (Ip et al., 1992; Jiang et al., 1991; Pan et al., 1991; Simpson, 1983; Thisse et al., 1987; Thisse et al., 1991). The enhancer regions of these genes have low affinity sites for DL, limiting their expression to ventral regions.
Following its activation by DL, SNA acts as a direct transcriptional repressor by limiting the expression domains of other DL target genes to lateral regions (Boulay et al., 1987; Ip et al., 1992; Kosman et al., 1991; Leptin, 1991). For example, *rhomboid (rho)* is required specifically in ventral-lateral regions to specify neuroectodermal fate. Even though *rho* has high affinity sites for DL and DL coactivators in its enhancer (Gray et al., 1994; Jiang and Levine, 1993), its expression is limited to ventral-lateral regions because this is the only location where the DL activator is present and SNA, its repressor, is not.

DL is not expressed in dorsal regions but plays a role similar to BCD and NOS by repressing the expression of dorsalizing genes in ventral regions. The primary dorsalizing signal is the morphogen encoded by *decapentaplegic (dpp)*, a homologue of the transforming growth factor-β (TGF-β) family of cytokines (Irish and Gelbart, 1987; Padgett et al., 1987). DL uses corepressors to directly down-regulate transcription of *dpp* and several other dorsal genes in the ventral regions of the zygote, thereby limiting their domains of expression and activity to dorsal regions (Huang et al., 1995; Lehming et al., 1994).

Summarizing, TOLL receptor activation by ventral SPZ results in a gradient of nuclear DL localization extending from the ventral surface. At high DL levels the mesoderm specifying factors TWI and SNA are induced. SNA then limits the expression of genes under DL control, such as *rho*, to ventral-lateral regions, thereby specifying neuroectodermal fates. DL also acts as a repressor of dorsalizing factors, like DPP, by directly down-regulating their expression in ventral and ventral-lateral regions. This allows DPP to specify amnioserosa (high DPP) and ectodermal fates (low DPP).
Examination of the underlying mechanism of axes formation in *Drosophila* reveals that only one determinant need be localized in order to initiate a complicated patterning process. Although posterior NOS functions to inhibit maternal *hb* mRNA translation in the posterior end, in *hb* maternal (*hb*mut) mutants NOS function is dispensable and fertile adults can be obtained in a *hb*mut/nos double mutant (Hülskamp et al., 1989; Irish et al., 1989; Struhl et al., 1989). Therefore, one localized morphogen (BCD) can pattern the A-P axis, excluding the terminal regions. The BCD gradient patterns the anterior end and also produces a gradient of the posterior determinant CAD (HB can also prevent anterior CAD function by an unknown mechanism). DL acts in a similar fashion by acting as a morphogen to pattern ventral fates and by preventing the ubiquitous expression of the dorsal morphogen DPP.

D) From Asymmetry to Pre-pattern: Making Parasegments

The activities of BCD/HB, CAD, DL, and DPP have provided the egg with asymmetries (limited zygotic gene expression) that can now be used to pattern the organism. For the sake of brevity and relevance, I will focus on the events involved in patterning the body trunk in the A-P axis. Although the differentiation mechanisms used in each axis have similarities, a major difference between the two is the production of periodicity along the length of the A-P axis by the aperiodic gap and maternal genes.

Periodicity in the A-P axis is first visible as the alternating striped expression patterns of the pair-rule genes. The pair-rule genes represent the next wave of genes after the gap genes and will produce the first visible sign of segmentation in the embryo, the parasegments (reviewed in Martinez-Arias and Lawrence, 1985); (Figure 1-7). The 14
parasegments are preludes to the future segmental divisions prominent in larvae and adults, and are unique in that they display an early lineage restriction. The "compartments" produced by this lineage restriction prevent cell mixing and are similar to rhombomeres and somites found in vertebrates (reviewed in McGinnis and Krumlauf, 1992). In Drosophila, these compartmental boundaries act as organizing centres for future patterning events in both larval and adult structures (reviewed in Ingham and Martinez-Arias, 1992; Lawrence and Struhl, 1996). As development of all higher organisms consists of the sequential division of the egg into "domains" of differentiation, the discovery and analysis of parasegments has offered a simple system to understand how regions of an embryo become restricted in their capacity to differentiate, immune from competing influences in the egg.

i) Periodicity

The protein expression pattern of the pair-rule gene fushi tarazu (ftz) (Kaufman et al., 1980; Wakimoto and Kaufman, 1981) at cellular blastoderm is a remarkable image (Carroll and Scott, 1985; Karr and Kornberg, 1989). Merely two hours after fertilization seven sharp ftz stripes in alternating parasegments along the A-P axis of the egg are visible (Figure 1-8). The segmented nature of Drosophila implied that regulators of differentiation would develop periodic patterns of expression during development, but one marvels at how early in development this expression is achieved and required.

The periodicity of the pair-rule genes results from the regulatory abilities of the overlapping maternal and gap proteins (reviewed in Pick, 1998). Initiation of even-skipped (eve) expression, another pair-rule gene, by the maternal and gap proteins has
been well studied and serves as a model for the activation of other pair-rule genes. The periodic expression of *eve* is complex and involves the use of stripe-specific enhancers, combined-stripe enhancers, and seven-stripe enhancers that act following the initiation of a broad band of very weak *eve* expression in the trunk of the embryo.

In the case of the stripe-specific enhancers, activation of *eve* stripe 2 expression is the best understood (Figure 1-9). Stripe 2 expression requires activation by maternal BCD and zygotic HB and is counter balanced by the repressive effects of the gap proteins Giant (GT) and Krüppel (KR) (Small et al., 1992; Small et al., 1991). The peaks of GT and KR expression are offset, allowing BCD and HB to activate *eve* expression in a limited domain between the two repressors. The binding sites for all four proteins are partially overlapping, and it appears that competition for binding and local repression by quenching occurs (Arnosti et al., 1996; Gray et al., 1994). This elegant mechanism, sometimes using different players, provides periodicity for other *eve* stripes (see Fujioka et al., 1999) and is probably used for achieving periodicity in the expression domains of other pair-rule genes. All *eve* stripes, however, are not individually controlled as composite-stripe enhancers (eg. stripes 4 and 6) appear to exist (Fujioka et al., 1999). Following activation of the early stripe enhancers a single late enhancer is activated which functions to increase the level of *eve* expression in all seven stripes and to refine the early broad stripes into sharper domains. The late enhancer requires the auto-regulating ability of EVE and the repressor effects of several other pair-rule gene products (Frasch and Levine, 1987; Fujioka et al., 1995; Fujioka et al., 1996).

Although a hierarchy within the pair-rule class was originally proposed based on
genetic interactions between members, the situation appears to be more complicated (see Pick, 1998). This model proposed that "primary" pair-rule genes like eve, hairy (h), and runt interpret gap gene cues and then were directly responsible for providing the striped pattern for the "secondary" pair-rule genes such as fitz because composite seven-stripe, but not stripe-specific, enhancers were found for fitz (Hiromi and Gehring, 1987; Hiromi et al., 1985). Also, early results suggested that the secondary pair-rule genes did not play a role in regulating the expression of the primary genes. Recent studies have indicated that this model is too simple (Nasiadka and Krause, 1999; Pick, 1998; Saulier-Le Drean et al., 1998; Yu and Pick, 1995) and suggest that maternal and gap proteins play an important role in the early specification of stripes for most pair-rule genes, with cross-regulatory interactions between pair-rule genes refining and properly positioning the stripes.

ii) Defining Parasegments

As there are numerous pair-rule genes with partially overlapping expression domains, it was not immediately apparent where the 14 parasegments would arise. Assistance came from the realization that the parasegmental boundaries form between stripes of engrailed (en) and wingless (wg) expression (reviewed in Ingham and Martinez-Arias, 1992), two members of the segment polarity gene family. wg is expressed in the posterior domain of one parasegment while en is expressed in the anterior region of the next parasegment. It is now clear that the overlapping expression domains of the pair-rule genes act as a combinatorial code to delimit the domains of en and wg expression (DiNardo and O'Farrell, 1987; Ingham et al., 1988); (Figure 1-10).

The 14 en stripes form at the anterior borders of the alternating eve (odd-
numbered parasegments) and fitz (even-numbered parasegments) stripes (Harding et al., 1986; Howard and Ingham, 1986; Macdonald et al., 1986). Although stripes of EVE and FTZ are broad at the time of en initiation (Frasch and Levine, 1987; Harding et al., 1986), en expression is limited to a few cells in the anterior of each parasegment (Fjose et al., 1985; Kornberg et al., 1985). This important limitation on en expression is created by the combinatorial code of pair-rule genes: odd-numbered en stripes require the overlap of EVE and Paired (DiNardo and O'Farrell, 1987; Morrissey et al., 1991). In contrast, even-numbered en stripes require the overlap of FTZ and Odd-paired (Benedyk et al., 1994; DiNardo and O'Farrell, 1987). Moreover, en expression in the middle of the parasegment is repressed by other pair-rule proteins (Cadigan et al., 1994; DiNardo and O'Farrell, 1987; Manoukian and Krause, 1992).

The early broad expression of EVE and FTZ initially acts to repress wg expression in the posterior of each segment (Ingham et al., 1988). However, as development proceeds the expression of eve and fitz decays in the posterior region of the parasegment and it is at this time that the appropriate activators turn wg expression on there (Baker, 1987; Benedyk et al., 1994; Ingham et al., 1988). Additionally, other genes, like naked and sloppy-paired, provide polarity to the parasegment by limiting the regions where en and wg expression can be activated and maintained (Cadigan et al., 1994; Mullen and DiNardo, 1995).

Pair-rule gene expression is initiated before cellular blastoderm (~ 2hrs AEL) and within an hour the parasegments have been defined by the juxtaposition of en and wg stripes. The importance of the parasegments, as I discuss below, is in their organizing
ability; the borders of the parasegments act as sources of morphogenic activity to specify different cell fates within the parasegments.

E) Patterning the Larval Epidermis

The segmented A-P axis will give rise to several structures, including the larval epidermis and the adult imaginal discs. Again, for simplicity, I will focus on only one of these tissues, the larval epidermis. The larval epidermis consists of 3 thoracic and 8 abdominal segments (the head segments are involuted), with each segment represented by a unique denticle pattern (Figure 1-1A). In most segments the cuticle consists of rows of denticles in the anterior half and naked cuticle in the posterior half. What also makes the denticle pattern of the cuticle interesting is that not only are the individual segments unique, in terms of the bristle pattern, but there are also differences in denticle morphology within each segment as one moves from the anterior to posterior end of the segment. The existence of unique segments and polarity within each segment offers a simple system to understand complex patterning interactions. For example, how are individual cell fates (the different denticles) established within the framework of a larger defined field (the parasegment) which itself must be given a unique identity (the different segments)?

i) Defining Position Within the Parasegment

Lawrence has proposed that the different ventral cuticle patterns observed across the segment could be achieved by a morphogen gradient specifying different bristle types at different morphogen concentrations (Lawrence and Sampedro, 1993). In this model the function of the segment polarity genes is solely to establish and maintain the
parasegmental borders, borders which are essential to maintain a “source” and “sink” for the morphogen. Other models, in contrast, suggest that the overlapping domains of segment polarity gene expression contain enough information to pattern the approximately seven different individual cell types that exist within a parasegment (Bejsovec and Wieschaus, 1993; Martinez-Arias et al., 1988). Recent results suggest that both methods - morphogen gradients and combinatorial codes - are used to pattern the ventral epidermis of the larvae.

Denticle specification occurs several hours after the parasegmental boundaries are established, and until this specification can occur the integrity of the parasegmental boundaries must be maintained (reviewed in DiNardo et al., 1994). Since expression of the pair-rule genes decays around the end of germ band extension (~4 AEL), a mechanism is required to maintain expression of \( en \) and \( wg \). In fact, they maintain each other's expression during the early phase of their expression, while at later stages this dependence is lifted. \( EN \) contains a homeodomain (Poole et al., 1985) and acts as a transcription factor, while \( WG \) is a secreted glycoprotein that participates in signaling events (Cabrera et al., 1987; Rijsewijk et al., 1987). During the early phase of their expression, \( WG \) signals across the parasegmental boundary to the \( en \) expressing cells and activates a signaling cascade that results in the continued expression of \( en \) (Bejsovec and Martinez-Arias, 1991; Cumberledge and Krasnow, 1993; DiNardo et al., 1988; Martinez-Arias et al., 1988). The \( en \) expressing cell then produces another signal, encoded by \textit{hedgehog} (\textit{hh}), which signals back to the \( wg \) expressing cell to maintain \( wg \) expression (Ingham et al., 1991; Ingham, 1993). This reciprocal signaling occurs locally and
consolidates the parasegmental boundary until the process of cell specification can occur (Figure 1-11B).

Although proper denticle specification was known to require both WG and HH signaling (Bejsovec and Wieschaus, 1993), two additional signaling systems have been identified which are also essential for proper bristle patterning to occur. These systems act downstream of the WG and HH signals: Veinlet (Rhomboid) signals through the EGFR pathway, and Serrate signals through the Notch pathway (O'Keefe et al., 1997; Szüts et al., 1997; Wiellette and W., 1999); (Figure 1-11C-E).

Initially, a bipartite signal involving WG and HH is required during the early stages of denticle specification (Alexandre et al., 1999; Gritzan et al., 1999; Moline et al., 1999; Sanson et al., 1999). These molecules, however, are blocked at the parasegment boundary, resulting in unidirectional signals traveling away from the boundary. Repressive signaling by WG anteriorly and HH posteriorly, within the same parasegment, sets the boundary of Serrate (Ser) expression. HH and SER then act as activators for veinlet (ve) in non-Ser expressing cells, while WG repression of ve limits its expression to more anterior regions of the parasegment. Through this cross-regulatory network of signaling cascades, a combinatorial code of molecules has been established which could potentially provide enough specificity to establish the identities of all denticles (Alexandre et al., 1999). WG signaling in the posterior of the parasegment specifies naked cuticle, while the overlapping domains of EN, HH, SER, and VE lead to the specification of different denticles within the anterior of the parasegment.
ii) Providing Pattern Specificity to the Parasegment: the Homeotic Genes

The activities of the WG and HH signaling pathways have led to the specification of cell fates within the parasegment; individual cells now know their location along the axis of a parasegment. The responsibility for providing segment-specific patterns (larval and adult head, thoracic or abdomen) to these cells lies with the genes of the homeotic complexes (HOM-C). Interestingly, the HOM-C acts analogously to WG/HH signaling described above in that HOM-C genes specify positional values along the length of the A-P axis, similar to the positional values specified along the length of the individual parasegments by WG/HH signaling. It is important to stress that HOM-C genes do not make, for example, abdominal-specific structures, but they tell the organism where such structures should be located.

HOM-C genes exist within two complexes: the Antennapedia complex (ANT-C), which contains 5 genes responsible for specifying head and thoracic identities anterior to parasegment 5, and the bithorax complex (BX-C), which contains 3 genes responsible for specifying thoracic and abdominal identities posterior to parasegment 4 (reviewed in Morata, 1993). Expression of these complexes is required throughout development, and is initiated very early during embryogenesis at the time when the gap and pair-rule genes are active. In particular, the BX-C gene Ultrabithorax (Ubx) is repressed in anterior regions of the embryo by hunchback protein (White and Lehmann, 1986) while the fitz pair-rule protein activates Ubx expression outside areas of Hunchback repression (Ingham and Martinez-Arias, 1986; Müller and Bienz, 1991; Müller and Bienz, 1992).

The number of primary HOM-C genes (8) is not enough to pattern all 14
parasegments individually, necessitating the need for a combinatorial code between different members to pattern at least some of the individual parasegments. This is most clearly understood for the abdominal regions where the combination of *Ubx* and *abdominal-A* is required to pattern parasegments 7-9, while *Ubx* alone is responsible for patterning parasegment 6. Although progress has been made in finding HOM-C target genes in the imaginal discs (Weatherbee et al., 1998), targets for patterning the larval epidermis remain largely unknown, but experiments are beginning to determine which genes the HOM-C regulate and how the 8 HOM-C genes work together to pattern the 14 parasegments (Casares et al., 1996; Castelli-Gair and Akam, 1995; Li et al., 1999).

It should not go unnoticed that the embryo uses the same genes (*eve* and *fzd*) to define the borders of the parasegmental divisions (determined by *en* and *wg* expression) and to delimit the domains of expression of genes (HOM-C) required to pattern these parasegmental divisions. Importantly, in terms of evolution, these processes are not linked directly but exist in parallel. The WG/HH system and HOM-C genes converge during later stages of development to produce the pattern of the parasegment; the WG/HH system specifies position along the parasegment, while the HOM-C specifies position along the A-P axis, producing segmental identity (thorax or abdomen). The parallel nature of the WG/HH signaling system and HOM-C specification function allows for evolutionary change by the HOM-C, independent of the WG/HH signaling system. The integrity of the parasegment is maintained by the WG/HH signaling system, leaving HOM-C genes free to evolve and possibly alter their targets. Patterning of the different segments can then change without affecting the cellular coordinates within each
parasegment.

F) Summary

The aspects of Drosophila development described here are remarkable in that, for the first time in any organism, one can follow a developmental process uninterrupted from the initial specification of the oocyte to the final specification of pattern in the developed organism. Although great strides have been made there are still obvious questions to be addressed, but the synthesis of genetics and embryology is now well under way. I have only discussed pattern in the larval epidermis, which is somewhat simplified being basically a two-dimensional field. However, the more complicated genetic interactions underlying three-dimensional structures such as wings and legs are also well understood and utilize similar principles (reviewed in Lawrence and Struhl, 1996). Truly, Drosophila melanogaster has provided an exceptional system to understand basic developmental questions.

II. Molecular Basis of Cell Fate Specification

A) The fushi tarazu Gene

The fushi tarazu (fitz) locus was identified in the Kaufman lab by screens for non-complementing mutations of a deletion for the Antennapedia region (Lewis et al., 1980a; Lewis et al., 1980b) and independently by the Nüsslein-Volhard and Wieschaus screen for segmentation mutants (1980). Characterization of the fitz mutant embryonic phenotype revealed a deletion of approximately half the number of segments in unhatched first instar larvae (Jürgens et al., 1984; Wakimoto and Kaufman, 1981), and experiments using a temperature sensitive allele revealed that the critical period of fitz activity required to
produce wild-type segments in the larvae is the 2-4hr period (cellular blastoderm) AEL (Wakimoto et al., 1984). This early role for FTZ in segmental patterning was confirmed when cloning (Bender et al., 1983; Garber et al., 1983; Kuroiwa et al., 1984; Scott et al., 1983; Weiner et al., 1984) and in situ analysis revealed that the 7 stripe expression pattern of ftz during cellular blastoderm (Hafen et al., 1984) corresponded to those regions deleted in the larval cuticle of ftz mutants (see Carroll and Scott, 1985; Martinez-Arias and Lawrence, 1985). Besides this early stage of ftz function, it is also required during neurogenesis in every segment (Doe et al., 1988) and it is also expressed in the developing hindgut (Krause et al., 1988), although its role there has not been addressed.

ftz encodes a 413 amino acid protein and contains a homeobox motif similar to that found in many homeotic genes (Kuroiwa et al., 1984) (Laughon and Scott, 1984; McGinnis et al., 1984a; McGinnis et al., 1984b; Weiner et al., 1984). FTZ uses its homeodomain (HD) to directly bind its own enhancer (Schier and Gehring, 1992; Schier and Gehring, 1993) and that of en (Desplan et al., 1988; DiNardo and O'Farrell, 1987; Howard and Ingham, 1986) to directly increase their levels of expression during cellular blastoderm (Florence et al., 1997; Nasiadka and Krause, 1999). FTZ also directly increases the expression of several homeotic genes in the ftz-dependent segments at this time (Ingham and Martinez-Arias, 1986; Ish-Horowicz et al., 1989; Müller and Bienz, 1992), while negatively regulating wg expression (Ingham et al., 1988; Ish-Horowicz et al., 1989), also directly (Copeland et al., 1996; Nasiadka et al., 2000; Nasiadka and Krause, 1999). An activating role for FTZ in the embryo is consistent with assays in tissue culture and yeast cells where FTZ acts a transcriptional activator, dependent on HD
binding sites and the HD of FTZ (Fitzpatrick and Ingles, 1989; Han et al., 1989; Jaynes and O'Farrell, 1988; Ohkuma et al., 1990; Winslow et al., 1989). However, it has yet to be determined how FTZ acts as a repressor.

As described in the Introduction, FTZ function is fundamental in defining the parasegmental boundaries and producing segment-specific expression patterns for the homeotic genes. However, a rather surprising discovery suggested that the HD of FTZ is not required for it to perform the majority of its functions in patterning the epidermis (Fitzpatrick et al., 1992). Before discussing this novel finding in greater detail it is appropriate to review the HD motif, revealing why this novel activity for FTZ is so surprising.

B) The Homeobox

The term homeosis was first used by William Bateson to describe mutations that transformed one particular segmental or metameric structure of an organism into the identity of another (Bateson, 1894). Genetic analysis in *Drosophila* over the last 80 years has identified several loci which produce similar effects, most notably genes of the BX-C and ANT-C described by Lewis (1978) and Kaufman (1980), respectively. When these genes were cloned and found to contain a similar 60 amino acid motif, the DNA sequence was appropriately named the “homeobox” and the amino acid motif a “homeodomain” (McGinnis et al., 1984a; McGinnis et al., 1984b; Shepherd et al., 1984). This HD motif was also found in MATa1 and MATα2 proteins required for cell-type switching in yeast and was reminiscent of the helix-turn-helix motif found in several repressor proteins from λ phage (Laughon and Scott, 1984), suggesting that the homeobox was highly conserved
during evolution (Shepherd et al., 1984). Indeed, the homeobox has been found in all metazoa that have been analyzed and has provided new insights into the process of evolution (reviewed in Carroll, 1995).

HD containing proteins can be divided into at least 20 different classes, with their homology based on the strong conservation of the same 7 amino acids in more than 95% of classified HDs and another 10 amino acids found in 80% of these same proteins (reviewed in Gehring et al., 1994). Remarkably, the other diverged sequences do not appear to affect the structure of the HD as the Antennapedia, Engrailed and MATα2 HDs, which are considerably diverged, form almost identical 3-D structures (Kissinger et al., 1989; Qian et al., 1989; Wolberger et al., 1991). These structure analyses revealed that the HD motif consists of three helices, the second and third of which resemble the helix-turn-helix motif of the λ repressor. Helices 1 and 2 are anti-parallel to each other and stack on top of the third helix, which is perpendicular to the first two. This third helix, also known as the recognition helix, makes contacts in the major groove of the DNA through several highly conserved residues, while minor groove contacts are made through conserved residues in the flexible N-terminal arm.

The similarity of the HD tertiary structures and the conserved amino acid sequences among different proteins suggested that they would also recognize similar DNA sequences, and this has proven to be true (reviewed in Hayashi and Scott, 1990). HD proteins recognize a similar 5’ ATTA 3’ core sequence with approximately the same affinity ($K_d = 10^{-8} - 10^{-9}$). However, some specificity is achieved by DNA contacts made by position 50, located in the recognition helix, with the two bases preceding the ATTA
core (reviewed Treisman et al., 1992). Bicoid, for example, contains a Lysine residue at this position and prefers a GG dinucleotide, while Antennapedia contains a Glutamine and prefers a CC dinucleotide (Hanes and Brent, 1989; Treisman et al., 1989). However, as the HD proteins of the ANT-C and BX-C, including FTZ, have the same residue at this position, it raises the question as to how HD-containing proteins of this class achieve specificity in vivo.

Studies of transcriptional regulation in vivo have clearly demonstrated that, regardless of how HD proteins attain specificity, the HD plays a major role (reviewed in Hayashi and Scott, 1990). In particular, HD swaps revealed that the HD is crucial in targeting proteins to the proper targets (Gibson et al., 1990; Kuziora and McGinnis, 1989; Mann and Hogness, 1990). Experiments which replaced the Deformed (DFD) HD with the HD of Ultrabithorax (UBX), for example, alters 17 of the DFD residues in the HD, none of which reside in the recognition helix (Kuziora and McGinnis, 1989). This swap, which also altered 5 residues just C-terminal to the HD, targeted the DFD-UBX hybrid to a Ubx target, Antennapedia (Antp), resulting in the activation of this gene. This result was particularly interesting because UBX normally represses Antp expression during development, suggesting that the HD of UBX targets the protein but that regulation, either positive or negative, depends on sequences outside of the HD. DFD normally autoregulates, potentially explaining the positive effect of the DFD-UBX hybrid on Antp transcription. Interestingly, the DFD-UBX hybrid was not targeted to the endogenous Dfd gene, presumably because it is not a normal target of UBX.

These results were extended in other HD swap experiments to narrow down the
regions responsible for targeting. Residues in the N-terminal arm of the HD were crucial for providing target specificity; in some cases only 5 amino acids could alter the targeting ability of the HDs of DFD, Sex combs reduced and ANTP (Chan and Mann, 1993; Furukubo-Tokunaga et al., 1993; Lin and McGinnis, 1992; Zeng et al., 1993). Interestingly, these residues are not predicted to contact DNA, suggesting the possibility that they recognize particular cofactors that assist HD proteins in target recognition. Regardless of the mechanism, the HD swap experiments support the notion that the HD and surrounding sequences are critical for providing target specificity to HD containing proteins in vivo, making the finding of homeodomain-independent activities for FTZ quite surprising.

C) The Importance of the HD in FTZ Function

i.) HD-Independent FTZ Activities

The above studies clearly demonstrate the importance of the HD for gene regulation in vivo, and similar results were expected in experiments using the ftz protein (Fitzpatrick et al., 1992). This study analyzed FTZ activity by attaching ftz transgenes to the hsp70 promoter to drive ubiquitous expression upon transient heat-shock (HS) pulses (Struhl, 1985). Expression of wild-type FTZ using this method produces an “anti-ftz” phenotype (Struhl, 1985) in which the ftz-independent segments are deleted, due to the inappropriate expression of ftz in regions that do not normally require its activity (Ish-Horowicz et al., 1989). Additionally, the HS-ftz transgene can rescue the ftz-dependent segments, as determined by the production of larval cuticle, in a ftz mutant background (Copeland et al., 1996; Hyduk and Percival-Smith, 1996). Although the anti-ftz phenotype
is still produced in these embryos due to the ectopic expression of ftz, the experiment demonstrates that the transgene can perform many of the functions of endogenous ftz and is not simply acting through endogenous ftz, an important consideration since ftz auto-regulates its expression.

Surprisingly, of all the constructs tested, the only ftz-deletion construct that could produce the anti-ftz phenotype or rescue the ftz-dependent segments in a ftz mutant was one in which the majority of the HD is missing (most of helices 1 and 3 and all of helix 2 are missing) and is incapable of binding DNA (Copeland et al., 1996; Fitzpatrick et al., 1992; Hyduk and Percival-Smith, 1996). The fact that this construct, FTZΔHD, is capable of rescuing ftz-dependent cuticle in a ftz mutant background demonstrates that the HD of FTZ is not required for it to perform the majority of its functions in the epidermis. These functions include activating en and the homeotic genes, while repressing wg. However, unlike wild-type FTZ, FTZΔHD cannot rescue a ftz mutant phenotype when its expression is solely controlled by endogenous ftz enhancer sequences (Furukubo-Tokunaga et al., 1992). This is most likely caused by the inability of FTZΔHD to auto-regulate at low concentrations, a step that may require direct FTZ binding to DNA (Schier and Gehring, 1992; Schier and Gehring, 1993). The levels produced by the hsp70 promoter appear to bypass the need for auto-regulation, suggesting that this may be the only step that requires the HD.

The finding of HD-independent activities for FTZ is consistent with observations demonstrating that its HD is not required in vivo if protein-stabilizing mutations are
present (I. Duncan, personal communication). Also, the N-terminus of FTZ, which does not include the HD, can synergistically activate transcription in tissue culture cells with the pair-rule protein Paired (PRD) (Ananthan et al., 1993). These results are consistent with the HD being required only at low levels, and strongly suggested that FTZ would interact with other proteins to carry out its functions.

Searches for FTZ-interacting proteins have revealed at least two promising candidates to date (Copeland et al., 1996; Guichet et al., 1997; Yu et al., 1997). Consistent with the results of Ananthan et al. (1993), our lab has demonstrated that FTZ and PRD interact directly in vitro through an N-terminal domain in FTZ, and that this FTZ/PRD interaction is involved in wg repression in vivo (Copeland et al., 1996). FTZ and PRD are repressors and activators, respectively, for wg expression. Therefore, FTZ may function as a direct repressor by inhibiting PRD function at the wg promoter, or FTZ may act by squelching (promoter-independent association) PRD, preventing it from binding the wg promoter. Further analysis will be required to determine which of these models is used by FTZ.

A rather exciting FTZ-interacting protein recently identified is the orphan nuclear receptor encoded by the Ftz-F1 gene (Guichet et al., 1997; Yu et al., 1997). Ftz-F1 encodes two transcripts: α, which is maternally expressed, and β, which is zygotically expressed (Lavorgna et al., 1993; Ohno and Petkovich, 1993). These isoforms differ in their N-terminal regions but contain the same DNA- and ligand-binding domains characteristic of this class of protein. α-Ftz-F1 mutants (no maternal product) have a pair-rule phenotype identical to that of ftz mutants (Guichet et al., 1997; Yu et al., 1997).
Since α-Ftz-F1 binds to the ftz enhancer (Ueda et al., 1990) it seemed likely that the pair-rule phenotype produced by the α-Ftz-F1 mutant was due to a lack of ftz expression. However, ftz expression is normal in α-Ftz-F1 mutants (Guichet et al., 1997; Yu et al., 1997). In fact, the ftz-dependent en stripes are missing and the ftz-dependent repression of wg does not occur, suggesting that α-FTZ-F1 acts as a cofactor for FTZ regulation of en and wg expression. Consistent with these findings, a reporter construct from the en gene requires juxtaposed FTZ and α-FTZ-F1 binding sites for activity (Florence et al., 1997).

α-FTZ-F1 interacts strongly with an N-terminal domain of FTZ, which includes an LXXLL motif (where X is any amino acid) (Schwartz et al., 2000). This motif, the nuclear receptor box, is required by several coactivators to bind nuclear receptors (Heery et al., 1997), and a deletion in FTZ that includes this domain prevents the cooperative regulation of en expression in vivo by FTZ and α-FTZ-F1 (Schwartz et al., 2000). The region of α-FTZ-F1 required for FTZ binding in vitro is the conserved AF-2 domain (Schwartz et al., 2000), a region in the ligand binding domain of other hormone receptors that is required to bind coactivators (Durand et al., 1994; Wurtz et al., 1996).

Most cofactors identified for nuclear hormone receptors function as repressors (in the absence of ligand binding) or activators (in the presence of ligand binding) by acting as, or recruiting, histone deacetylase or histone acetyltransferase complexes, respectively (Torchia et al., 1998). Although a ligand for α-FTZ-F1 has not been identified, the finding that α-FTZ-F1 uses its conserved AF-2 domain to bind to the conserved nuclear receptor box of FTZ (Schwartz et al., 2000) suggests that FTZ may be involved in
recruiting other factors, such as acetyltransferases, required for nuclear hormone receptor activity, a novel finding for HD-containing proteins. Additionally, the cooperative interaction between FTZ and α-FTZ-F1 provides target specificity to FTZ, since α-FTZ-F1 is required for FTZ to regulate some genes (eg. en) but not others (eg. ftz) (Schwartz et al., 2000).

**ii. Is the HD of FTZ Important for its Function?**

The data suggest that the HD is dispensable for all known ftz-dependent activities in the epidermis, although it has yet to be determined whether the HD is required in other tissues where FTZ is also expressed. On the other hand, it is has also been shown that FTZΔHD, under endogenous enhancer/promoter control, is not sufficient to rescue ftz null embryos, presumably due to the inability of FTZ to autoregulate and produce levels of FTZ required for activation of itself and other target genes. Although the implication is that FTZ only requires its HD to autoregulate at low levels (Copeland et al., 1996) there are reasons to be cautious about this conclusion.

First, HS-FTZΔHD is less efficient than HS-FTZ at inducing the anti-ftz phenotype in either ftz+ or ftz- backgrounds, suggesting that the lower levels of FTZΔHD, besides not being able to autoregulate endogenous ftz, may not be sufficient to properly regulate en and wg (Hyduk and Percival-Smith, 1996). Second, most assays for FTZ function use artificial expression methods (HS promoters) to produce protein, an important consideration since differences of less than two-fold in protein levels can have a noticeable effect on the ability of a protein to regulate transcription (Berleth et al., 1988;
Frohnhöfer and Nüsslein-Volhard, 1986; Roth et al., 1989; Thisse et al., 1991). Thirdly, it appears that FTZ can be directed to lower affinity sites in the presence of a coactivator (Florence et al., 1997; Yu et al., 1997). For example, the activation of an en reporter construct requires binding sites for both α-FTZ-F1 and FTZ in vivo (Florence et al., 1997), and FTZ binds a low affinity endogenous FTZ site cooperatively with α-FTZ-F1 in vitro (FTZ binding is increased by at least 50 fold), as long as binding sites for both proteins are present (Yu et al., 1997). These results suggest that at lower concentrations of FTZ the HD could be important for many of the FTZ-dependent activities, increasingly important if cofactors are required to target FTZ to lower affinity sites. At higher concentrations the need for DNA-bound FTZ may be bypassed, allowing FTZ to interact with the proper cofactors, like α-FTZ-F1 and PRD, to regulate transcription. Therefore, a fine line might exist between HD-dependent and HD-independent activities.

D) Achieving Functional Specificity for HD-Containing Proteins

Although several examples of how protein-protein interactions can provide target specificity to HD-containing proteins exist (see Wegner et al., 1993; Xue et al., 1993), two well studied examples provide exquisite insight. Work in yeast has demonstrated how cooperative interactions between the HD proteins of the MATα and MATα loci allow yeast to “differentiate” into several cell types. Additionally, study of the extradenticle (exd) protein from Drosophila, homologous to the α1 protein encoded by MATα, has expanded the yeast findings and demonstrates how layers of regulation can be used to specifically target HD-containing proteins to the proper targets.
i.)  $a_1$, $\alpha_2$, and $MCM1$ Function in Yeast

The yeast *Saccharomyces cerevisiae* can exist as one of three different cell types: the haploid types $a$ and $\alpha$, and the diploid $a/\alpha$ (reviewed in Johnson, 1995). In haploid cells either the $a$-specific or $\alpha$-specific genes, along with the haploid-specific genes, are activated, while in diploid cells the haploid-specific genes must be repressed. The $a$ cell appears to be the default cell type and, therefore, the $a$-specific genes must be repressed in both the $\alpha$ and $a/\alpha$ cells. This repression occurs through the cooperative interaction between the MAT$\alpha$ protein, $\alpha_2$, and the ubiquitous MCM1 protein (present in all three cell types), a non-HD containing protein. In the diploid cell $a_1$, from the MAT$\alpha$ locus, and $\alpha_2$ bind cooperatively to haploid-specific genes to repress them. These interactions have proven an exceptional model for understanding how HD proteins can be directed to specific gene targets.

All $a$-specific genes contain a 32bp operator upstream of their promoters. Within this sequence is a site that recognizes an MCM1 dimer, flanked on either side by sites for $\alpha_2$ (Keleher et al., 1988; Keleher et al., 1989). $\alpha_2$ dimerizes in solution using an N-terminal domain (Sauer et al., 1988) that is attached to the HD by a flexible region. This flexibility causes promiscuous DNA binding by $\alpha_2$, as it can recognize individual sites in different head-to-tail orientations as well as sites separated by variable spacing (Smith and Johnson, 1992). Order is given to this flexible domain by an interaction between an $\alpha_2$ dimer and an MCM1 dimer (Tan and Richmond, 1998; Vershon and Johnson, 1993), allowing MCM1 to set the proper spacing for the $\alpha_2$ sites by locking $\alpha_2$ into a set
configuration (Smith and Johnson, 1992). This increases the specificity, although not the affinity, of \( \alpha_2 \) for the 32bp operator sequence and demonstrates how \( \alpha_2 \) can discriminate between different operators by the use of a cofactor.

In diploid cells, \( \alpha_1 \) and \( \alpha_2 \) interact to form a heterodimer that cooperatively recognizes the haploid-specific operator sequence (Dranginis, 1990; Goutte and Johnson, 1988). This interaction requires a second flexible domain in \( \alpha_2 \) located C-terminal to the HD that interacts with residues between the first and second helices of the \( \alpha_1 \) HD (Mak and Johnson, 1993). Similar to the situation with MCM1, the interaction between \( \alpha_1 \) and \( \alpha_2 \) structures the flexible domain of \( \alpha_2 \) (Li et al., 1995) and potentiates the binding ability of \( \alpha_1 \) (Stark et al., 1999), which has low DNA binding specificity and affinity on its own. Interestingly, this cooperativity can occur in vivo even when the recognition helix of the \( \alpha_2 \) HD is mutated to prevent DNA binding by \( \alpha_2 \) (Vershon et al., 1995). These same mutations in \( \alpha_2 \), however, severely compromise the ability of \( \alpha_2 \) to interact cooperatively with MCM1, suggesting a different role for the DNA binding ability of \( \alpha_2 \) in each complex.

The example of \( \alpha_1 \) and \( \alpha_2 \) nicely illustrates the potential mechanisms by which HD-containing proteins can achieve specificity in vivo: a cell-type specific protein (\( \alpha_2 \)) can interact with a ubiquitous factor (MCM1) or another cell-type specific protein (\( \alpha_1 \)) to change the DNA binding specificity and affinity of the complex (either \( \alpha_2/MCM1 \) or \( \alpha_2/\alpha_1 \)). Additionally, the potentiating ability that \( \alpha_2 \) has on the DNA binding ability of \( \alpha_1 \), even when \( \alpha_2 \) can not bind DNA, serves as a potential model for some HD-
independent activities of FTZ.

ii.) extradenticle (exd)

exd mutations cause mild homeotic transformations in larvae (Jürgens et al., 1984) without grossly altering HOM-C gene expression patterns (Peifer and Wieschaus, 1990), suggesting that exd could encode a cofactor for HOM-C proteins (reviewed in Mann and Affolter, 1998; Mann and Chan, 1996). However, both HOM-C and exd genes have functions independent of each other, suggesting that EXD does not act solely as a HOM-C cofactor (Peifer and Wieschaus, 1990; Rauskolb et al., 1993). exd is ubiquitously expressed during early development, contains a homeobox (Rauskolb et al., 1993), and is homologous to the cancer causing PBX genes of humans and the C. elegans gene ceh-20 (Burglin and Ruvkun, 1992). Consistent with its hypothesized function, EXD can selectively raise the specificity of HOM-C proteins in vivo for their targets (Chan et al., 1994b; Chan et al., 1997; Ryoo and Mann, 1999). For example, decapentaplegic (dpp) expression in parasegment 7 of the visceral mesoderm requires both Ultrabithorax (UBX) and EXD, but not Antennapedia (ANTP). UBX and EXD binding sites are found juxtaposed in an enhancer fragment of dpp and they bind this site cooperatively in vitro. ANTP, however, which also recognizes the UBX site does not have its binding enhanced by EXD (Chan et al., 1994b).

Furthermore, EXD can cooperatively bind DNA with other HOM-C proteins in vitro, depending on the DNA sequence used (Chan et al., 1994b; van Dijk and Murre, 1994) (Pöpperl et al., 1995), and this cooperativity requires regions within the HD and a hexapeptide motif located N-terminal to the HD of the HOM-C protein (Chan and Mann,
1996; Chang et al., 1995a; Johnson et al., 1995; Passner et al., 1999; summarized in Mann and Chan, 1996). The EXD HD is 65% identical to the a1 HD and it has been suggested that the hexapeptide motif of HOM-C proteins is similar to the hydrophobic patch used by α2 to contact a1, since the hexapeptide motif also inserts into a pocket in the EXD HD surface (Chan and Mann, 1996; Passner et al., 1999). Also, sequences within the N-terminal arm of the HD which were suggested to be a source of specificity for different HOM-C proteins (see Pages 26-27) also appear to be partially involved in EXD interactions (Chan and Mann, 1996; Chan et al., 1997; Ryoo and Mann, 1999). A DNA site which binds to EXD/UBX or EXD/LABIAL (LAB) equally well can be made to specifically bind EXD/UBX or EXD/LAB when nucleotides in the minor groove, which are recognized by the N-terminal arm of the HD, are changed appropriately (Chan and Mann, 1996).

Further regulation of HOM-C proteins is achieved by temporally and spatially regulating the nuclear localization of EXD (Aspland and White, 1997; Mann and Abu-Shaar, 1996). Surprisingly, this localization is primarily dependent on the direct interaction between EXD and another HD-containing protein encoded by the homothorax (hth) gene (Rieckhof et al., 1997). The HTH HD is 43% identical to the EXD HD and is homologous to the murine protein MEIS1, suggesting that the HOM-C, EXD, and HTH components have been conserved during evolution. hth expression coincides with nuclear localization of EXD, although not all nuclear EXD requires hth expression. As EXD is only active when in the nucleus, the regulation of hth expression is another method by which HOM-C activity could be regulated (Henderson and Andrew, 2000; Ryoo and
Mann, 1999) since the target specificity of HOM-C proteins changes in the presence of EXD.

Interestingly, the presence of a HD motif in HTH may add another level of specificity to the EXD/HOM-C interaction (Ryoo et al., 1999). A tertiary complex containing HTH/EXD/LAB forms on an endogenous enhancer sequence from the lab gene and this complex is required for regulation of a shortened lab enhancer fragment in vivo. Mutations in any one of the three separate sites abolishes enhancer activity.

EXD has provided unique insight into how HOX genes obtain specificity in vivo. EXD can change the specificity and affinity of HOM-C DNA binding by specifically directing different HOM-C proteins to lower affinity sites not normally bound by HOM-C monomers in vitro. The importance of low affinity binding by HOM-C proteins was largely overlooked because the in vitro DNA-binding studies used for determining HOM-C binding sites selected for the highest affinity sites; it was thought that these would represent true HOM-C binding sites in vivo (see Draganescu et al., 1995 and references therein). These high affinity sites may not be relevant in vivo or, possibly, may not require EXD for specificity in vivo; several target genes are regulated by the same HOM-C genes in vivo, but to different degrees, and would not require EXD for specificity (Graba et al., 1997). The regulation of nuclear EXD by HTH allows for another level of control to be placed on HOM-C binding, as HOM-C proteins will potentially be targeted to different sites depending on which tissues contain nuclear EXD. Furthermore, the finding of HTH/EXD/HOM-C complexes provides additional specificity to HOM-C function.
An additional twist to the EXD story is the proposal that EXD only regulates the activity of HOX proteins bound to an enhancer. It has been suggested that HOM-C proteins act as repressors when bound as monomers to target sites, but that the presence of EXD turns the EXD/HOM-C complex into an activator complex (Li and McGinnis, 1999; Li et al., 1999; Pinsonneault et al., 1997). Further experimentation will be required to verify this hypothesis.

III. Maintenance of Cell Fate Specification

A) The Polycomb and trithorax Groups of Genes

Differentiation in higher organisms requires that cells maintain the activated or repressed transcriptional states of specific loci throughout development. In organisms ranging from flies to humans two conserved protein groups, the products of the Polycomb group (PcG) and trithorax group (trxG) of genes, play an important role (reviewed in Jacobs and van Lohuizen, 1999). Mutations in PcG genes in Drosophila cause posteriorly directed transformations in embryos and adults because of a failure to maintain the repressed transcriptional state of HOM-C genes (Simon, 1995). Conversely, mutations in trxG genes result in anteriorly directed transformations due to a failure to maintain the active transcriptional state of HOM-C genes (Kennison, 1995). However, since their initial characterization as regulators of the HOM-C, trxG and PcG genes have been shown to be involved in many other developmental processes, including specification of the CNS and PNS, dorsal ventral patterning, imaginal disc growth, and segmentation (see Simon, 1995 and references therein; Breen, 1999; Sinclair et al., 1992).
i.) The Polycomb Group

*Polycomb (Pc)*, the founding member of the PcG, was first characterized by its dominant extra sex combs phenotype in male flies (Lewis, 1978). This homeotic phenotype consists of the transformation of legs of the second thorax (T2) (which lack male sex combs), into legs of T1 identity (which possess male sex combs). Since the characterization of *Pc*, many genes have been identified which either mimic or enhance the extra sex combs phenotype, and hence have been classified as members of the PcG (Simon 1995). Approximately 15 PcG proteins have been identified and 30-40 are predicted to exist (Jürgens, 1985; Landecker et al., 1994).

Most PcG null/strong mutants are zygotically lethal and produce unhatched larvae with posteriorly transformed segmental identities (Struhl, 1981; Duncan, 1982; Jürgens, 1985; Breen and Duncan, 1986). In strong mutants these cuticles are abdominal 8 (A8) in appearance, suggesting ubiquitous expression of the most posterior HOM-C gene, *Abdominal-B*, throughout the embryo. Analysis of *Pc* (Beachy et al., 1985; Carroll et al., 1986; Celniker et al., 1989; Wedeen et al., 1986) and other PcG mutants (McKeon and Brock, 1991; Simon et al., 1992; Struhl and Akam, 1985) revealed that the initial expression of the HOM-C genes was normal but that later expression patterns were expanded into more anterior regions for most of the HOM-C genes studied. These results demonstrated that the PcG genes are not required during the initiation of HOM-C expression, but for maintaining the silenced state of HOM-C genes in those tissues where they have not been activated.

Although the cloning of PcG members has offered clues as to how these proteins...
maintain repressed states of transcription, there is still a wide gap in the understanding of how this occurs. Originally, it was suggested that PcG proteins create "heterochromatin" structures (Paro, 1990). This idea was based largely on the fact that PC contains a region of homology, called the chromodomain, with the Heterochromatin protein 1 (HP1) protein from *Drosophila* (Paro and Hogness, 1991). HP1 is encoded by the *Suppressor of variegation 205* gene and mutations in this gene act as suppressors of position-effect variegation (PEV) (Eissenberg et al., 1990; Eissenberg et al., 1992), a phenomenon in which euchromatic regions of DNA that are juxtaposed (by chromosome inversions etc.) to heterochromatin become randomly silenced to varied extents in different cells (reviewed in Weiler and Wakimoto 1995). This silencing is believed to occur by the variable spreading of the heterochromatic state into euchromatic regions. Because HP1 is localized to heterochromatic regions and is involved in maintaining heterochromatin, it was proposed that PC, because of its similarity to HP1, would be involved in packaging HOM-C genes into a heterochromatic state to maintain them in a silenced state.

Although this hypothesis is quite intriguing, the accumulating data do not support a simple model in which PEV and PcG repression occur by identical mechanisms. Screens for modifiers of PEV have not, for the most part, identified members of the PcG, suggesting that the heterochromatin associated with PEV does not require PcG members (Kennison, 1995; Sinclair et al., 1998a). Additionally, in some cases in which a PcG gene has been shown to modify PEV it cannot be ruled out that the effect is indirect, since many of the PcG proteins potentially regulate the transcription of PEV modifiers (see Stankunas et al., 1998). Furthermore, some of the PEV modification associated with PcG
genes is actually due to the presence of second site modifiers on the mutated chromosome (Rio, 1999; Sinclair et al., 1998a). However, these results do not rule out the possibility that PcG proteins act by packaging their DNA targets into heterochromatin (i.e. by expansion of a repressive nucleosome structure), but does suggest that this silencing is mechanistically different from PEV silencing. Different types of heterochromatin exist in *Drosophila* (Wallrath and Elgin, 1995) and it is possible that factors involved in PcG repression or PEV may interact with the nucleosome structure in unique ways to maintain the repressed state.

Besides a heterochromatin-like model, it has also been suggested that PcG proteins might silence genes through DNA looping (Pirrotta, 1995). PcG complexes bound at discrete sites along the enhancer might interact with each other to produce one large PcG complex that blocks activator binding sites. It has also been proposed that PcG complexes could sequester target genes to specialized compartments within the nucleus that are not transcriptionally active (Schlossherr et al., 1994; Strouboulis and Wolffe 1996). Understanding how the PcG group proteins silence genes will require more biochemical studies addressing the function of PcG complexes. To date, the only biochemical analysis of an intact PcG complex purified from *Drosophila* tissues suggests that at least one function of a PcG complex is to stabilize the formation of nucleosome structures (Shao et al., 1999), supportive of a “heterochromatin-like” model for PcG function. Nucleosome structures are postulated to be a primary target of transcriptional activators (see below) and nucleosome blockage of enhancer binding sites, the TATA box, etc., represents a plastic mechanism by which genes could be regulated.
Biochemical analyses (Franke et al., 1992; Horard et al., 2000; Jones et al., 1998; Kyba and Brock, 1998; Ng et al., 2000; Peterson et al., 1997; Shao et al., 1999; Strutt and Paro, 1997; Tie et al., 1998) and the detection of overlapping binding sites for PcG proteins on polytene chromosomes support the idea that multiple PcG complexes exist (DeCamillis et al., 1992; Franke et al., 1992; Lonie et al., 1994; Martin and Adler, 1993; Peterson et al., 1997; Rastelli et al., 1993; Sinclair et al., 1998b; Stankunas et al., 1998; Zink and Paro, 1989). As well, the presence of pleiotropic phenotypes, besides homeotic defects, in various PcG mutants implies that the different PcG complexes regulate distinct subsets of target genes. Currently, it is not known why different complexes are required.

One common component expected of PcG complexes, however, is a member(s) capable of binding DNA. PcG complexes associate with specific DNA sites on polytene chromosomes known as PcG response elements (PREs) (Chan et al., 1994a; Gindhart and Kaufman, 1995; Simon et al., 1993) and these elements are sufficient to induce repression in a PcG-dependent manner in vivo when attached to a reporter construct (Busturia and Bienz, 1993; Chan et al., 1994a; Christen and Bienz, 1994; Müller and Bienz, 1991; Simon et al., 1993; Simon et al., 1990; Zhang and Bienz, 1992). To date, the only PcG protein capable of sequence-specific DNA binding is the pleiohomeotic (pho) protein (Brown et al., 1998) and sites for PHO binding are found in many PRE sequences (Mihaly et al., 1998). Whether this protein is required for all PcG complexes is not yet known, but other DNA binding PcG members are predicted to exist since some PRE sequences identified by in vivo assays do not contain consensus PHO binding sites (Tillib et al., 1999).
Although progress is being made in determining which PcG complexes contain which PcG members, more biochemical analysis is required before it is understood how these genes function in maintaining repressed states of transcription. In addition, other aspects of PcG activity need to be addressed, including: how PcG complexes maintain silencing after mitosis (Buchenau et al., 1998) when it is expected that most complexes will be displaced from the DNA for replication, and how PcG complexes are initially recruited to their targets. Recruitment is an important consideration since, for the most part, PcG genes are expressed ubiquitously throughout the early stages of development but must act to silence different genes at different positions along the A-P axis. Although recruitment by repressor proteins is an attractive hypothesis, it has also been suggested that the transcriptional state alone (i.e. an “open” or “closed” enhancer/promoter conformation) could determine whether or not a PcG complex will form; no recruitment by repressor proteins would be required (Poux et al., 1996). However, this model cannot explain all the complexity associated with PcG function. For example, some PcG genes act in a tissue-specific manner (McKeon and Brock, 1991; Simon et al., 1992; Soto et al., 1995) and this activity will most likely require physical interactions with tissue-specific factors (Soto et al., 1995).

ii.) The trithorax Group

In contrast to PcG genes, trxG genes are required to maintain the activated state of target genes (reviewed in Kennison, 1995). trxG genes were identified in large part by screening for dominant suppressors of PcG mutant homeotic phenotypes (Kennison and Tamkun, 1988) and, like the PcG genes, are also required to regulate the expression of
non-HOM-C genes (Breen, 1999; Brizuela and Kennison, 1997; Elfring et al., 1998; Vazquez et al., 1999). Importantly, the identified trxG mutants cause homeotic transformations in embryos and adults on their own (Ingham and Whittle 1980; Kennison and Tamkun, 1988; Tamkun et al., 1992), indicating that they play a major role in maintaining gene function.

A breakthrough, in terms of understanding how PcG and trxG proteins might function, resulted from cloning brahma (brm), a gene identified in the Kennison and Tamkun screen (1988). Unlike many of the previously identified PcG and trxG genes, brm provided immediate insight because it was homologous to a known yeast protein with an identified function: the SWI2/SNF2 protein of the SWI/SNF complex (Tamkun et al., 1992). This complex, consisting of at least 11 proteins, is an ATP-dependent modifier of chromatin structure and is required for normal transcription of several yeast genes (reviewed in Kingston et al., 1996). As a general mechanism for transcriptional activation, it is postulated that chromatin remodeling by the SWI/SNF complex, in conjunction with the combined efforts of transcriptional activators and nucleosome-acetylating complexes, causes the dissociation or sliding of nucleosomes away from enhancer/promoter regions of a target gene, keeping it in an "open" configuration primed for transcription (reviewed in Workman and Kingston, 1998). It is important to stress that not all yeast genes require SWI/SNF or acetylation functions, and in some instances the SWI/SNF and acetylation functions may act independently of each other in gene regulation.

Evidence that acetylation "marks" genes as active in Drosophila has recently
come from study of the *Fab-7* DNA element from the BX-C (Cavalli and Paro, 1999). This element allows for PcG/trxG-dependent regulation of a reporter construct in vivo and conveys epigenetic inheritance of an active transcriptional state through meiosis (Cavalli and Paro, 1998). Activation of the reporter construct during embryogenesis, but not during larval stages, results in the hyperacetylation of H4 histones associated with the transgene. Interestingly, activation of the reporter transgene is inherited through meiosis only when the transgene is activated during embryogenesis; inheritance does not occur if activation takes place during larval stages. The authors postulate that the role of PcG/trxG proteins is to prevent erasure of such epigenetic marks.

The BRM homology to the SWI/SNF complex suggested a mechanism by which trxG proteins could counteract PcG activity (Collins et al., 1999; Crosby et al., 1999; Daubresse et al., 1999; Dingwall et al., 1995; Vazquez et al., 1999). However the purification of a BRM complex from *Drosophila* (Papoulas et al., 1998) identified only two known members of the trxG family (Crosby et al., 1999; Dingwall et al., 1995; Kal et al., 2000), suggesting that multiple independent trxG complexes likely exist. Consistent with this, two other complexes, each containing a separate trxG member, were identified in the Papoulas study (1998). Although the biochemical function of these additional two complexes is not yet known, several chromatin-remodeling complexes have been identified in yeast (Cairns et al., 1996; Peterson et al., 1994) and *Drosophila* (Ito et al., 1997; Tsukiyama et al., 1994; Varga-Weisz et al., 1997). The *Drosophila* complexes appear to have different activities in vivo (Kal et al., 2000), supporting a model in which several different trxG complexes are involved in maintaining the activated state of target
gene transcription.

Study of chromatin remodeling complexes has offered a mechanism to understand how an active state can be created, but has yet to address the interplay between the PcG and trxG complexes. In particular, how do the trxG complexes antagonize PcG activity to allow permanent activation of target genes? It is known that PcG/trxG function is required throughout development since, without the later function of some trxG components, PcG repression can be re-established (Brizuela and Kennison, 1997; Ingham, 1981; Ingham, 1985; Shearn et al., 1987). This fact points to a continuous role for some trxG members in directly counteracting PcG function.

A potential candidate for direct involvement in antagonizing PcG function is the founding member of the trxG, *trithorax (trx)*. *trx* was identified over 30 years ago as a regulator of HOM-C expression (Ingham and Whittle, 1980; Lewis, 1968) and was also identified in the Kennison and Tamkun screen (1988) for repressors of PcG mutants. Interestingly, TRX overlaps many PcG protein binding sites on polytene chromosomes (Chinwalla et al., 1995; Kuzin et al., 1994) although these sites, referred to as Trithorax response elements (TREs) (Chang et al., 1995b), appear to be distinct from PREs, at least in the *Ultrabithorax* regulatory regions (Tillib et al., 1999). Furthermore, TRX and PC were found to associate with their identified TREs and PREs in the HOM-C genes very early during development (Orlando et al., 1998), suggesting that these complexes are poised to react to the appropriate repressors and activators. These results suggest that TRX does not compete for PC binding but that TRX may be involved in direct protein-protein interactions with PcG complexes to counteract their effect.
How does TRX, then, function to maintain activated transcription? Although this is not yet known, an interesting finding is that TRX interacts with the *Drosophila* homologue of the SNF5 component of the yeast SWI/SNF complex (Rozenblatt-Rosen et al., 1998), suggesting that TRX, or a TRX complex (Rozovskaia et al., 1999), may be involved in recruiting or stabilizing a BRM-type complex at target promoters in reaction to the proper activators during early development.

B) The *Additional sex combs* Gene

*Additional sex combs* (*Asx*) was first identified by its larval head defects Nüsslein-Volhard et al., 1984) but was subsequently found to enhance *Pc* and other PcG mutants (Campbell et al., 1995; Jürgens, 1985). Like genes of the PcG, *Asx* loss-of-function (LOF) mutants cause posteriorly directed homeotic transformations in the cuticle of zygotic null larvae, although this phenotype is weaker than that seen in *Pc* mutants (Breen and Duncan, 1986). Heterozygous *Asx* mutant adults have dominant homeotic transformations, including occasional ectopic sex combs on T2 legs in males and posteriorly directed transformations of the abdominal tergites, while heterozygous adults and homozygous embryos exhibit occasional segmentation defects (Sinclair et al., 1992). Also, consistent with *Asx* acting as a member of the PcG, *Asx* mutants cause ectopic derepression of HOM-C genes (McKeon and Brock, 1991; Simon et al., 1992; Soto et al., 1995). Curiously, most *Asx* alleles act as gain-of-function (GOF) mutations for head phenotypes, since the head defects seen in these mutants are more severe (ie. more *Pc*-like) than null LOF alleles uncovered by larger deletions of the *Asx* region (Sinclair et al., 1992).
Several findings suggest that Asx also has trxG activity. First, a P-element induced allele of Asx, AsxP1, exhibits both anterior and posterior transformations (Sinclair et al., 1992). Second, several Asx alleles, including a deficiency, enhance thoracic and abdominal mutations in the trxG gene trithorax (trx) (Milne et al., 1999). Finally, Asx mutations were found to enhance position effect variegation (PEV), a result not expected for PcG members (Sinclair et al., 1998a). In light of these findings, it has been suggested that the weaker LOF Asx phenotype, when compared to other PcG genes, may result from the combined loss of trxG and PcG activities (Sinclair et al., 1992). This dual role may explain the complexity of the Asx GOF mutations in the head.

The possibility that trxG and PcG complexes have common components is supported by the discovery that Enhancer of zeste, originally identified as a PcG member, also has trxG activity (LaJeunesse and Shearn, 1996). Also, it was recently demonstrated that several genetic enhancers of both trxG and PcG genes were previously identified as PcG genes (Gildea et al., 2000). These results have led Gildea et al. (2000) to suggest that an additional group, ETP (Enhancers of trithorax and Polycomb mutations), should be included in the class of trxG and PcG genes.

The mechanism by which trxG/PcG proteins are directed to their targets and how they maintain a "memory" of the desired state is an active area of research (reviewed in Pirrotta, 1998). This question is especially relevant to Asx, as it has tissue-limited effects even though it is ubiquitously expressed in the early embryo (Sinclair et al., 1998). Asx mutants generally exhibit misexpression of HOM-C genes in the epidermis and visceral mesoderm, but not in the CNS (Soto et al., 1995). Interestingly, ASX is active in the
CNS, as a 14.5kb regulatory fragment from Ultrabithorax (*bxdI4*) attached to a reporter construct is responsive to Asx mutants in the CNS (Soto et al., 1995). Soto et al. (1995) have suggested that this type of tissue specificity could be achieved by one of two basic mechanisms. The first would require tissue-specific differences in ASX, which could include modifications of the Asx protein. This mechanism, as they point out, appears unlikely since ASX is capable of acting on the *bxdI4* element in the CNS. The second mechanism would require tissue-specific factors to interact with ASX at the target gene.

Currently, the role that ASX plays in either trxG or PcG complexes is unknown. Although ASX binding in polytene chromosomes overlaps sites for proteins of both the trxG and PcG (Sinclair et al., 1998b) it is difficult to determine the significance of this, in terms of a role for ASX in both complexes, since several trxG and PcG members also have overlapping binding sites in polytene chromosomes. Cloning of Asx has revealed limited homology to other members of the PcG, as well as two human ESTs (Sinclair et al., 1998b). The conserved domains include a C-terminal cysteine cluster, but the function of this domain is not yet known. However ASX acts, the requirement for Asx in both trxG and PcG functions, as well as the tissue-limited effects of ASX, suggest that physical interactions with other proteins are important.

**IV. Thesis Outline**

In this thesis, I discuss the characterization of the *tantalus (tan)* gene from *Drosophila melanogaster*. Results from the Krause laboratory and others (Fitzpatrick et al., 1992; Copeland et al., 1996; Hyduk and Percival-Smith, 1996) have demonstrated the importance of protein-protein interactions in Fushi tarazu (FTZ) function during early
segmentation in the embryo, and I have used the yeast two-hybrid screen to identify potential protein partners for FTZ. TAN was identified as a protein capable of interacting with the homeodomain of FTZ in the yeast two-hybrid and Far Western assays.

During the analysis of *tan*, it came to our attention that this gene had also been identified in the lab of Dr. Hugh Brock at the University of British Columbia as an interacting protein partner for Additional sex combs (ASX), a member of the Polycomb- and trithorax-groups (PcG and trxG respectively) of transcriptional maintenance factors. The observation in Dr. Brock’s lab that endogenous TAN protein binds to numerous sites on polytene chromosomes that overlap sites of ASX binding provided strong evidence that these two proteins were interacting in vivo and, in collaboration with Dr. Brock’s lab, precipitated an investigation into the role that TAN plays in development in conjunction with ASX.
**Drosophila Life Cycle**

- Fertilization
- **Embryo**
- **Adult**
  - **Male (♂)**
  - **Female (♀)**
  - 1 day

**Pupa**
- 3.5 - 4.5 days in pupal stage
- 1 day

**First instar larva**
- 1 day

**Second instar larva**
- 1 day

**Third instar larva**
- 1 day

**2.5 - 3.5 days**
Figure 1-2. Embryogenesis. (A) Stages of early embryogenesis are depicted with the focus on the migration of nuclei to the periphery of the egg and cellularization. Adapted from Hartwell, L.H. et al., “Genetics: From Genes to Genomes” Chapter 21. The McGraw Hill Companies, 2000. (B) The top two panels shows successive stages of germ band extension after gastrulation has occurred (3¼ hours – 5¼ hours AEL). The segments are clearly visible in the germ band retracted embryo in the bottom panel (10¼ hours AEL). VM, ventral midline; GB, germ band; T3, thoracic segment 3; A1, abdominal segment 1. Adapted from Hartenstein, V., “Atlas of Drosophila Development” M. Bate and A. Martinez-Arias, eds. Cold Spring Harbor Laboratory Press, 1993.
A

Diploid zygotic nucleus
30 min

Cortex

Mitotic cycle 7
1h 10 min

Multinucleate syncytium

Most nuclei migrate out to cortex

Mitotic cycle 10
1 h 30 min

Pole cells forming

Syncytial blastoderm

End of mitotic cycle 13
2 h 30 min

Primordial germ cells ("Pole cells")

Membranes in egg's cortex grow inward

2 h 30 min
2 h 45 min
3 h
3 h 15 min

Cellular blastoderm

B

T3

A1

VM

GB
Figure 1-3. A-P and D-V axes specification during oogenesis. The top panel shows a 16 cell cyst during early Gurken signaling. The nucleus (circle) and microtubule network (long strands) are visible. During stage 7 of oogenesis the microtubule reorganization signal required to specify the A-P axis is received in the oocyte, allowing the localization of bicoid RNA to occur (left side of figure). The rearranged microtubule network causes the nucleus to migrate to a margin location, specifying the D-V axis (right side of figure). The late Gurken signal prevents pipe expression in the future dorsal side, thereby limiting active Spätzle to ventral regions. Adapted from van Eeden, F and St. Johnston, D.. 1999.
A-P patterning

Stage 9

Anterior follicle cells

Posterior follicle cells

D-V axis patterning

Stage 8

Nuclear migration

Stage 10

Late Gurken signaling

bicoid mRNA

Rearranged microtubules

Stage 1-6

Early Gurken signal

Stage 7

Microtubule rearranging signal

Localised posterior determinants

Fertilized egg

bicoid mRNA

pipe mRNA

nanos mRNA

Active Spi"ettle
Figure 1-4. Hierarchy of genes involved in A-P patterning. The gene groups are shown on the left side (Maternal etc.). An example of a gene belonging to each group (bicoid etc.), with its corresponding expression pattern, is shown on the right side. A, anterior; P, posterior. Adapted from Wolpert, L., “Principles of Development” Chapter 5. Current Biology Ltd., 1998.
Gene Group | Gene Example
---|---
Maternal | bicoid
Gap | giant
Pair-rule | fushi tarazu
Segment polarity | engrailed
Homeotic | Ultrabithorax
Figure 1-5. Domains of gap gene expression. The embryo at top shows the location of the termini (T), anterior (A), posterior (P), and pole cells (Pc). The regions of expression for each gap gene are shown in solid lines, with the dotted lines representing later expression patterns for giant (gt) and hunchback (hb). For simplicity the anterior domains of huckebein (hkb) and tailless (ill) have not been shown. The domain of active caudal protein is also shown (CAD). Md, mandible; Mx, maxilla; La, labium; Proct. proctodeum; bcd, bicoid; Kr, Krüppel; kni, knirps. Adapted from Pankratz, M.J. and Jäckle, H., 1990.
Figure 1-6. D-V axis specification. Nuclear localization of Dorsal (represented by purple coloured gradient) activates the mesoderm specifying genes twist and snail. Snail represses rhomboid expression ventrally, allowing Dorsal to activate rhomboid in ventral lateral regions where it is involved in neuroectodermal specification. Dorsal also represses decapentaplegic expression in ventral regions, allowing the gradient of Decapentaplegic to specify amnioserosa and dorsal ectodermal fates.
Figure 1-7. Parasegments versus segments. Segmental divisions, as defined in the adult, consist of the anterior region of one parasegment and the posterior region of the next in the embryo. The posterior (p) part of each segment is defined by the engrailed expression domain in the early embryo; anterior, (a); head, (H); clypeolabrum, (C); thorax, (T); abdomen, (H). Adapted from Lawrence, P., “The Making of a Fly” Chapter 1. Blackwell Scientific Publications, 1992.
Figure 1-8. *fushi tarazu* expression. Seven stripes of expression are seen during the blastoderm stage of embryogenesis. Anterior is to the left and dorsal is up. Image courtesy of A. Nasiadka.
Figure 1-9. Expression of *even-skipped* stripe 2. Expression of stripe 2 occurs between the repressors Giant and Krüppel and the activators Bicoid and Hunchback. Adapted from Wolpert, L., “Principles of Development” Chapter 5. Current Biology Ltd., 1998.
Figure 1-10. Expression pattern of selected pair-rule and segment polarity genes. *fushi tarazu* (*ftz*) and *even-skipped* (*eve*) are expressed in the Even- and Odd-numbered parasegments (PS), respectively. At this time, each parasegment is four cells wide (represented by boxes in bottom of figure) The horizontal gradation of colour represents the narrowing of stripes during embryogenesis. The vertical gradation of colour for the *sloppy-paired 1&2* (*slp 1&2*) and *odd-skipped* (*odd*) genes represent the later appearance of secondary stripes for these genes. *prd; paired, opa; odd-paired, wg; wingless, en;engrailed*. Adapted from Copeland, J.W.R. “The role of protein-protein interactions and phosphorylation in the function of the homeodomain protein Fushi tarazu” Ph.D. thesis, 1997.
Figure 1-11. Specification in the larval epidermis. (A) Top panel shows the cuticle and a corresponding schematic diagram of a first instar larvae. The first and last thoracic (T) and abdominal (A) segments are marked. Each segment is composed of anterior rows of hairs which are unique within in the segment and between segments. The posterior of each segment is devoid of cuticle. Image courtesy of A. Nasiadka. (B-E) Role of the Wingless and Hedgehog signaling pathways in cell specification within the parasegment. (B) Diagram of genetic interactions. (C) The parasegmental border falls between the Wingless (wg) expressing cell (blue) and the Engrailed/Hedgehog (en/hh) expressing cell (red). Early negative signaling by Wingless and Hedgehog regulate the spatial domains of Serrate (ser) expression (green). (D) Wingless, Hedgehog and Serrate signaling regulate the position of the Veinlet (ve) expression domain. (E). Apparent correlations between gene expression and specific cuticular structures. Further experiments are needed to confirm the correlations in parenthesis. Adapted from Alexandre et al., 1999.
CHAPTER 2

Identification of Tantalus as a Fushi tarazu-Interacting Protein

I performed all experiments in this chapter. The bait constructs used in the yeast two-hybrid screen were prepared by Benoît St. Pierre and the yeast two-hybrid reagents were a gift from Roger Brent.
Abstract

The pair-rule gene *fushi tarazu* (*ftz*) is an essential gene in *Drosophila melanogaster* required to regulate many target genes. The *ftz* protein contains a homeodomain (HD) DNA binding domain, allowing FTZ to bind DNA sequence-specifically. However, the study of many HD-containing proteins has revealed that, on their own, they are poor discriminators of DNA target sequences. I have used a yeast-two hybrid screen to identify genes encoding potential cofactors for FTZ function, and describe here the identification of one such gene that we have subsequently called *tantalus* (*tan*). TAN interacts with the HD of FTZ in both a yeast two-hybrid screen and a Far Western analysis. Additionally, TAN binds DNA in vitro, and ectopic expression studies in vivo reveal that TAN is a nuclear protein.

Introduction

As described in Chapter 1, the *fushi tarazu* (*ftz*) gene of *Drosophila melanogaster* is a member of the pair-rule class of segmentation genes (Nüsslein-Volhard and Wieschaus, 1980) and is expressed in an alternating seven-stripe pattern during early embryogenesis (Hafen et al., 1984). One of the major roles of *ftz* in the epidermis is to position the domains of expression of the segment polarity genes *engrailed* (*en*) and *wingless* (*wg*) (reviewed in Ingham and Martinez-Arias, 1992). These segment polarity genes act to define the boundaries of the parasegments, which act as fundamental patterning units in the anterior-posterior axis of the embryo. Patterning of these parasegments occurs through the parasegment-specific expression of genes found in the homeotic complexes (HOM-C). The expression of HOM-C genes defines the second
major role of ftz, as FTZ increases the level of expression of several HOM-C genes while also defining their domains of expression (Ingham et al., 1988; Ish-Horowitz et al., 1989; Müller and Bienz, 1992).

FTZ contains a homeodomain DNA binding motif (Kuroiwa et al., 1984; Laughon and Scott, 1984; McGinnis et al., 1984a; McGinnis et al., 1984b; Weiner et al., 1984) and this domain is required for FTZ to activate transcription in vitro (Fitzpatrick and Ingles, 1989; Han et al., 1989; Jaynes and O'Farrell, 1988; Ohkuma et al., 1990; Winslow et al., 1989). FTZ directly regulates its own expression (Schier and Gehring, 1992; Schier and Gehring, 1993), as well as that of en (Desplan et al., 1988; DiNardo and O'Farrell, 1987; Howard and Ingham, 1986; Nasiadka and Krause, 1999), wg (Copeland et al., 1996; Ingham et al., 1988; Ish-Horowitz et al., 1989; Nasiadka et al., 2000; Nasiadka and Krause, 1999) and the HOM-C genes (Ingham et al., 1988; Ish-Horowitz et al., 1989; Müller and Bienz, 1992). It was expected that these activities would require direct binding of FTZ, through its HD, to the regulatory regions of these genes. However, results from our lab and others have demonstrated that the HD of FTZ is not essential for it to carry out many of its functions in the epidermis (Copeland et al., 1996; Fitzpatrick et al., 1992; Hyduk and Percival-Smith, 1996). On the other hand, the HD is required for FTZ to rescue null ftz embryos in vivo, in the context of a genomic rescue construct (Furukubo-Tokunaga et al., 1992). Together, these results have lead to the suggestion that the HD is only required at low levels of FTZ, when an autoregulatory loop is required to increase ftz expression levels. Later, however, when the levels of FTZ are increased, the HD appears to be dispensable for FTZ function.
These findings strongly suggested that FTZ would require additional cofactors to regulate its target genes, and research in our lab and others has identified two such cofactors (Copeland et al., 1996; Guichet et al., 1997; Yu et al., 1997). Testing in our lab for protein-protein interactions between FTZ and other pair-rule proteins revealed a strong in vitro interaction with the paired protein, and this interaction appears to be required for the proper regulation of wg expression in vivo (Copeland et al., 1996). Additionally, FTZ interacts with the orphan nuclear receptor encoded by α-Ftz-F1. α-Ftz-F1 is a requisite cofactor for FTZ regulation of both en and wg, but not ftz itself (Guichet et al., 1997; Yu et al., 1997). FTZ and α-FTZ-F1 bind to lower affinity FTZ sites cooperatively (Yu et al., 1997) and both FTZ and α-FTZ-F1 sites are required for the expression of an en reporter construct in vivo (Florence et al., 1997). These results support the idea that FTZ requires cofactors to function. To extend these results, I have used the yeast two-hybrid screen in a less biased attempt to identify novel protein partners for FTZ.
Results and Discussion

Screening for Fushi tarazu interacting proteins

Original attempts at screening a 0-12hr cDNA library from Drosophila expressed in E. coli with radioactively labeled FTZ were unsuccessful in identifying protein-protein partners for FTZ (data not shown) and the yeast two-hybrid system was employed as an alternative approach (Figure 2-1). Because full-length ftz protein activates transcription in yeast cells (Fitzpatrick et al., 1992) and would therefore interfere with the screening approach (see Figure 2-1), several “Bait” constructs were created which contained different portions of the FTZ protein (Figure 2-2). These Baits were selected based on their hypothesized relative importance for FTZ function. In particular, Bait #2 consisted of a region of strong homology to the FTZ homologue from Drosophila hydei (Figure 2-2). Bait *3, on the other hand, contained 17 amino acids upstream of the highly conserved FTZ HD and continued until amino acid 49 of the 60 amino acid HD. The entire HD was not used because it was feared that the DNA binding activity of the construct might interfere with the two-hybrid assay. Of the four baits tested, only Baits #2 and #3 were suitable for the screen, as they did not activate transcription on their own when introduced into yeast.

The results of the screen are summarized in Table 2-1. Of the 9 clones identified by Bait #2 and sequenced, clones for rRNAs (2 clones), cytochrome c (1 clone) and several novel sequences (6 clones) were identified. Unfortunately, most of the novel clones had only short coding regions in frame with the acidic activation domain, making it difficult to assess their significance. In some cases, longer coding regions were found...
out-of-frame with the activation domain, making these clones unlikely candidates for FTZ interactors. One explanation for the isolation of these clones was the appearance of a weak coiled-coil protein interaction motif (Hodges, 1992) in both the Bait *2 construct and several of these out-of-frame clones. Sequencing of the 3 positive clones identified using Bait *3 revealed 2 rRNAs and one novel clone (see below) whose coding region was in frame with the acidic activation domain.

The isolation of several rRNAs using both baits is a common finding during two-hybrid screens, most likely due to the abundance of such species in the total amount of RNA in tissue samples. Unfortunately, clones for the strong FTZ interactors Paired (PRD) (Copeland et al., 1996) and α-FTZ-F1 (Guichet et al., 1997; Yu et al., 1997) were not identified in the screen. This result at first seems surprising since recent results from in vitro studies have demonstrated the importance of the N-terminal regions of FTZ, specifically amino acids 100-150, for cofactor binding to both PRD (Copeland et al., 1996) and α-FTZ-F1 (Guichet et al., 1997, Schwartz et al., 2000), and these residues were included in the Bait *2 construct. However, there are several possible explanations as to why the PRD and α-FTZ-F1 interactors were not identified in the screen.

First, although amino acids 100-150 of FTZ are required for the interactions with PRD and α-FTZ-F1, other regions of FTZ contribute to these interactions (Copeland, 1997; Schwartz et al., 2000). The absence of these regions in the Bait *2 construct may lower the FTZ/PRD and FTZ/α-FTZ-F1 interactions to levels below detection by the two-hybrid assay. As an extension of this idea, it is also possible that the isolated 100-150
residues of FTZ are not folded properly in the bait construct to participate in a protein-protein interaction. Second, yeast cells expressing the PRD and α-FTZ-F1 fusion proteins may not be viable or may be under-represented in the library, making their detection difficult. Finally, errors in the experimental method must also be considered.

*Tantalus* interacts with the HD of *Fushi tarazu*

Clone 33-13, hereafter referred to as *tantalus (tan)*, was identified using Bait #3. To verify the yeast two-hybrid results and to confirm that the interaction between TAN and FTZ is direct, the 1kb tan cDNA recovered in the screen was subcloned into a T7-expression vector and used in a Far Western assay with different FTZ deletion constructs. Bait #3 was comprised of 17 amino acids upstream of the FTZ HD and continued until amino acid 49 of the 60 amino acid HD, suggesting that TAN was making contacts with at least part of the HD. To test this, several constructs containing different portions of the ftz polypeptide were created for use in a Far Western assay. First, full-length FTZ, N- and C-terminal portions of FTZ (the HD is found in the C-terminal construct), GST and GST fused to the HD of FTZ (4 amino acids upstream and 10 amino acids downstream of the HD) were expressed in bacteria and run on SDS-PAGE (Figure 2-3A). To verify the position and identity of each construct on the blot, a Western using a polyclonal antibody against the FTZ peptide was performed (Figure 2-3B). Protein expressed from the tan cDNA was then radioactively labeled with 35S-methionine using a rabbit reticulocyte lysate and used to probe a second blot containing the same FTZ constructs as those detected on the Western. TAN bound to the full-length, C-terminal region, and isolated HD of FTZ but did not bind to the N-terminal portion of FTZ or to GST alone (Figure 2-
These results confirm the yeast two-hybrid interaction and suggest that the region of FTZ sufficient for an interaction with TAN includes the 4 amino acids preceding the HD and the first 49 amino acids of the of the HD.

This region of FTZ incorporates helices 1 and 2 of the HD motif, and studies of several HOM-C proteins (Gibson et al., 1990; Kuziora and McGinnis, 1989; Mann and Hogness, 1990), as well as other HD containing proteins (Chan and Mann, 1996; Mak and Johnson, 1993; Passner et al., 1999), have stressed the importance of residues within the N-terminus of these various HDs for target specificity. In particular, these residues are important for making protein-protein contacts with cofactors, such as Extradenticle (EXD) and OCA-B (Chan and Mann, 1996; Lai et al., 1992; Mak and Johnson, 1993; Passner et al., 1999; Pomerantz et al., 1992). For example, the HOM-C proteins in Drosophila have increased DNA binding specificity when complexed with the HD-containing EXD protein (reviewed in Mann and Chan, 1996), while the mammalian OCT-1 HD-containing protein has an altered DNA binding specificity, and becomes a more potent transcriptional activator, when complexed with OCA-B or VP16 (reviewed in Wegner et al., 1993). However, what role TAN may play in FTZ function will require further experimentation.

The identification of tan using the HD of FTZ may at first seem paradoxical, considering that this screen was initiated based on the finding that many activities of FTZ are HD-independent. However, it is important to remember that the analyses of FTZ function were, for the most part, based on results in the epidermis (Fitzpatrick et al., 1992; Copeland et al., 1996; Hyduk and Percival-Smith, 1996). ftz is also expressed in
other tissues, like the CNS, and the role of the HD in these tissues has not been addressed. Additionally, the HD is required for the auto-activating ability of FTZ (Schier and Gehring, 1992) and it is possible that cofactors are required for this activation.

*Sequence and genomic location of tantalus*

Sequencing of the 1kb *tan* cDNA recovered from the yeast two-hybrid screen revealed an open reading frame of 269 amino acids. Northern analysis (see below) suggested that the *tan* transcript was approximately 1.6kb in length, so a 0-12hr cDNA library was screened to identify overlapping cDNA clones. One cDNA of approximately 1kb was identified which extended the *tan* sequence in the 5' direction. Combining the new cDNA sequence with the old produced a total sequence of approximately 1.6kb (Figure 2-4A,B), in close agreement with the size predicted by Northern analysis.

Approximately 7kb of genomic DNA flanking the cDNA sequence was obtained by screening a genomic library (Figure 2-4C) and 4kb surrounding the *tan* locus was sequenced. Figure 2-5A shows the sequence of the 1.6kb *tan* cDNA and the *tan* genomic regions. A canonical TATA box sequence was not found in upstream of the recovered cDNA sequence. Comparison of the cDNA and genomic sequences reveals two introns, one located upstream of the putative ATG translational start site and the second within the coding region. This ATG conforms well to *Drosophila* consensus translation start sites (Cavener, 1987) and the predicted open reading frame would encode a protein with a mass of 33kDa.

*TAN* lacks homology to any proteins currently identified in GenBank; this is not unusual since only 50% of *Drosophila* proteins display sequence similarity to mammalian
and worm proteins, with even fewer showing homology to yeast proteins (28%) (Rubin et al., 2000). TAN does, however, contain several regions of interest (Figure 2-5B). Two large basic regions consisting of 32% (21/66) and 41% (19/46) basic amino acids are located at residues 90-154 and 188-235. Despite these large patches of basic residues, the overall pI predicted for TAN is 6.4. The protein has two canonical nuclear localization signals at residues 90-106 (RRSSTFGARAGVARRRM) and 217-220 (KRRR). The protein also contains two PEST regions (Proline, Glutamate, Serine and Threonine), which are believed to act as protein degradation signals. These motifs are located at positions 74-90 and 154-188, and have “PESTfind” scores of 9.5 and 11.7 respectively (a PESTfind score of +5 or greater is considered significant; Rechsteiner and Rogers, 1996). Numerous canonical sites for protein kinase C, cAMP-dependent kinase and casein kinase II phosphorylation are also present (data not shown).

To determine if the tan locus had been identified genetically, in situ hybridizations to polytene chromosomes were performed (Figure 2-6) and a P1 genomic library was screened (data not shown). Both methods localized tan to position 65A on the polytene map, consistent with the recent sequencing of the Drosophila genome (Adams, 2000). However, tan does not appear to co-map with any candidate FTZ-interacting genes.

The tantalus expression pattern

Before pursuing additional biochemical studies of the TAN/FTZ interaction, it was first important to determine whether these two proteins had the potential to interact in vivo, i.e., do their temporal and spatial domains of expression overlap? As ftz is expressed during the 2-4, 6-8 and 10-12hr time periods during development (Kuroiwa et
1984; Weiner et al., 1984) tan should also be expressed at this time if the two proteins are to interact in vivo. A developmental Northern blot shows high levels of tan mRNA in early embryos, decreased expression in 12-24 hr embryos and high levels of expression in third instar larvae, pupae, and females (Figure 2-7).

This early temporal expression pattern for tan overlaps that of FTZ. To determine whether the spatial expression domains of the two genes also overlap, in situ hybridizations were performed (Figure 2-8). Transcripts were first looked for in ovaries since the Northern blot results indicated the presence of tan expression in early embryos before zygotic transcription takes place. Indeed, tan mRNA is first detected in stage 10 ovaries in the nurse cells surrounding the oocyte (Figure 2-8A). These cells will eventually deposit their mRNA and protein contents into the oocyte. tan is also expressed ubiquitously in pre-cellularized and cellularizing embryos (Figure 2-8C,D). Signal is not detected in control embryos hybridized with a sense probe (Figure 2-8B). Transcript levels begin to decrease after cellular blastoderm but increase again during germ band extension. At this time, localized expression is seen in what appears to be the somatic and visceral mesodermal layers (Figure 2-8E). tan is also expressed ubiquitously in imaginal discs of third instar larvae, with higher levels of expression seen in the morphogenetic furrow of the eye-antennal disc (Figure 2-8F). These results demonstrate that the early expression pattern of tan partially overlaps that of ftz. However, the mesodermal and disc expression of tan suggest that at least some TAN functions must be independent of FTZ, since FTZ is not expressed in these tissues.
Tantalus is a nuclear DNA binding protein

As FTZ acts in the nucleus, the subcellular localization of TAN was determined by using an epitope tagged version of the protein. A DNA sequence encoding the p53 epitope tag (Dalby and Glover, 1993) was inserted in the tan cDNA at amino acid 19, and the cDNA was cloned into a T7 inducible vector to verify that the tag was functional (Figure 2-9A). The p53-tan cDNA was then cloned into the pCaSpeR-hs vector (Thummel and Pirrotta, 1992), which uses the hsp70 promoter to permit transgene expression in vivo upon heat shock (HS). Transgenic lines were produced using P-element mediated germ line transformation (Spradling, 1986). Third instar transgenic or wild-type (WT) larvae were exposed to HS and the location of p53-TAN was analyzed in dissected salivary glands using a monoclonal anti-p53 antibody (Figure 2-9B). No signal was detected in the salivary glands of WT controls (Figure 2-9C) but a strong nuclear signal was seen in both the salivary glands and attached fat bodies of the transgenic line (Figure 2-9B). At higher magnifications TAN appeared to be associated with the chromatin (Figure 2-9D).

Careful analysis of the TAN sequence revealed a small stretch of residues with homology to the third recognition helix of the homeodomain (HD) DNA binding motif (Figure 2-5C). This homology may be relevant as the zeste protein also contains a somewhat diverged HD sequence required for DNA binding in vitro (Chen et al., 1992). This raises the possibility that the association of TAN with specific sites on polytene chromosomes may depend in part on its own ability to bind DNA. To test this possibility, 35S-labeled TAN was produced using a reticulocyte lysate and incubated with salmon
sperm DNA that had previously been attached to nitrocellulose filters (Figure 2-10). The Fushi tarazu (FTZ) protein, with and without its DNA binding HD (FTZΔHD), and Luciferase (LUC), which does not bind DNA, were used as controls. Of the four polypeptides, TAN yielded the strongest signal. Full-length FTZ also bound to the blotted DNA, and as expected, FTZΔHD and LUC both failed to bind.

The lack of homology between tan and any other genes in the databases makes it difficult to predict how TAN may affect FTZ function. However, the finding that TAN binds DNA and is found in the nucleus is consistent with a role for TAN in assisting/stabilizing the DNA binding of FTZ. It will be important to determine whether TAN binds DNA sequence specifically, and how a TAN/FTZ interaction might affect this binding.

Summary

The tan protein was identified in a yeast-two hybrid screen and Far Western analysis as a potential partner for the pair-rule protein FTZ. A 53 amino acid region of FTZ which includes helices 1 and 2 of the HD is sufficient for the interaction with TAN. However, additional studies will be required to determine how specific the interaction is for the FTZ HD. TAN expression partially overlaps the temporal and spatial window of FTZ, consistent with a potential in vivo interaction between the two proteins. Although TAN does not show strong homology to any sequences in GenBank, our analysis supports a nuclear role for TAN since it binds DNA strongly and a tagged version of the protein is localized to the nucleus. Although the evidence presented here is consistent with an in vivo interaction between FTZ and TAN, proof that such an interaction is relevant in vivo
is still required.
Materials and methods

The yeast two-hybrid screen

The two-hybrid screen has been described in detail (Golemis, E.A., et al., 1997). The RFLY1 library was used in the screen and represents cDNAs from 0-12hr of embryogenesis (Golemis, E.A., et al., 1997). The FTZ Bait constructs were made by PCR amplification using the following primers:

Bait *1 5’CGGAATTCATGGCCACCACAAACAGC’3 and 5’CGGGATCCTCAAAAGCTTCTGCTCCTGATTGTGTA’3;
Bait *2 5’CGGAATTCGCCGCCCCAACAGGCACC’3 and 5’CGGGATCCTCAAAAGCTTGGGAGCCTTCTTACCTG’3;
Bait *3 5’CGGAATTCGCCGCTTTCAATGGTGC’3 and 5’CGGGATCCTCAAAAGCTTGGGAACCAGATCTTGATCTG’3;
Bait *4 5’CGGAATTCAGCTGGACAGCTCCCG’3 and 5’CGGGATCCTCAAAAGCTTGGGTAATCATGAGCAG’3

Far Western assay

The Far Western analysis was performed as previously described (Guichet et al., 1997) using 35S-labeled TAN expressed using a reticulocyte lysate (Promega TNT). The described FTZ constructs expressed on the gel/blot (Figure 2-3) were cloned into pET19B (Novagen). The HD of FTZ (4 amino acids upstream and 10 amino acids downstream of the HD) was cloned as a GST fusion into pGEX2T (gift of C. Desplan). Proteins were expressed in E. coli and run on SDS-PAGE using standard techniques (Sambrook, et al., 1989). The polyclonal anti-FTZ antibody has been described (Krause and Gehring,
Southern blots

DNA preparation was modified from Ballinger et al. (1989). Batches of twenty flies were collected in 1.5ml eppendorf tubes (flies were stored at -20°C until required). To each tube, 200µl of TENS buffer was added (TENS is 100mM Tris-HCl pH 7.6, 100mM EDTA, 100mM NaCl and 0.5% SDS). The flies were homogenized with a pestle and incubated at 65°C for 30min to denature protein. 30µl of 8M Potassium Acetate was then added (for a final concentration of 1.2M) to precipitate protein. The tube was spun at 14,000xg in a table top centrifuge after which the supernatant was removed and the spin repeated. An equal volume of 100% EtOH was added to the supernatant and the tube was mixed at room temperature and allowed to incubate for 5min to precipitate the DNA. The tube was again spun at 14,000xg after which the pellet was washed in 70% EtOH and allowed to dry. The high molecular weight genomic DNA was resuspended in 200µl TE overnight at 4°C. The DNA was ethanol precipitated 2X and resuspended in 40µl TE at 37°C. The DNA was digested with the appropriate enzymes, blotted onto positively charged nylon membranes (Boehringer Mannheim) and UV cross-linked. Blotting was performed as described (Sambrook et al., 1989) with the following modifications. DIG labeled DNA probes were made by PCR amplification. Labeling efficiency of the probe was verified by comparison to a control DIG labeled DNA sample. Probes were used at a concentration of 50ng/ml in 5X SSC, 1% (w/v) blocking reagent (Boehringer Mannheim), 0.1% (w/v) N-laurylsarcosine, and 0.02% (w/v) SDS. Incubations were at 65°C overnight.
and, unless noted, washes were 2X 5min at room temperature with 2X SSC, 0.1% (w/v) SDS followed by 2X 15min at 65°C with 0.1X SSC, 0.1% (w/v) SDS. Signal was detected using CDP-Star according to the manufacturer's instructions (Boehringer Mannheim).

Northern blots

Northern blots were performed as described for agarose/formaldehyde gels (Sambrook et al., 1989). RNA was isolated by direct phenol extraction (Andres and Thummel, 1994) and 10μg of total RNA loaded onto the gel. After blotting and UV cross-linking the blot was stained with Methylene Blue (Ausubel et al., 1997). Probe preparation and signal detection were performed as per the Southern blotting procedure.

Cloning tantalus

Genomic (Tamkun et al., 1992) and cDNA (Poole et al., 1985) libraries were screened using standard procedures (Sambrook et al., 1989). To obtain a full-length cDNA tan construct, two overlapping cDNAs were fused together using a common Bst EII site located at +582 (site positions refer to the genomic map in Figure 2-5) and subcloned into pBluescript II (Stratagene). The cDNA was subcloned into the pET19b vector (Novagen) for T7 expression. The p53 tag was inserted at the Bst EII site (before the first PEST domain) of the tan cDNA. The inserted sequence is N-RSRAFRHSVVR (new sequence between dashes).

In situ hybridization

Hybridization of embryos and imaginal discs with RNA probes was performed as described (Hughes and Krause, 1999) except that secondary antibodies were HRP-
conjugated and the ABC system of Pierce was used for detection. In situ hybridization to polytene chromosomes was carried out as follows. Salivary glands were dissected from larvae grown at 18°C. Glands were dissected in 1X PBS and placed in 45% acetic acid and allowed to fix for 3-5min. Glands were squashed and then placed in liquid nitrogen for 30sec followed by 3X 10min washes in 95% EtOH. Glands were often stored in 95% EtOH before use. Slides were rehydrated by successive washes of 2min each in 95% EtOH, 70% EtOH, 50% EtOH, 30% EtOH, 0.1X SSC and 2X SSC. Slides were then incubated in 2X SSC at 70°C - 80°C for 30min followed by denaturation in 0.1N NaOH for 90sec at room temperature. Slides were washed for 30sec in 2X SSC before another dehydration step of 2min each in 30% EtOH, 50% EtOH, 70% EtOH and 95% EtOH. Slides were air dried for 5min and incubated with the appropriate probe (made as per the Southern blotting procedure) in a moist chamber overnight at 37°C. Slides were washed for 2X 15min in 2X SSC followed by a 2min wash in 1X PBS and a 5min wash in PBT. Secondary antibody was added and slides were incubated for 1hr at 37°C. Slides were then washed in 1X PBS briefly followed by 2X 15min washes in 2X SSC and developed using standard NBT/BCIP procedures (Sigma). Chromosomes were then stained using Giemsa (Andrew and Scott, 1994).

**DNA binding of Tantalus**

5μg aliquots of sonicated salmon sperm DNA were vacuum blotted to nitrocellulose (nitrocellulose gave less background than other membranes) in a 30μl mixture of DNA and 2X SSC followed by washes in 500μl 2X SSC. The blot was
blocked with 1X Binding Buffer2 (1X BB2: 2% milk powder, 20mM Tris pH 7.6, 100mM NaCl, 0.25mM EDTA, 0.25mM DTT, 0.1% Tween-20 and 10% glycerol) for 2hrs on ice. Radioactively labeled proteins were synthesized according to the protocol provided by the manufacturer (Promega TNT T7 kit) and incubated with the blots in 1X BB2 for 25min on ice. Blots were washed with 1X BB2 with several changes over 15min, then dried and exposed to film.

*Preparation of larval tissues for antibody staining*

*tan* and *p53-tan* cDNAs were cloned into the pCaSpeR-hs (Thummel and Pirrotta, 1992) or pUAST (Brand and Perrimon, 1993) vectors for in vivo expression. Depending on constructs used, transgenic or wild-type third instar larvae were heat shocked for 2hrs and allowed to recover for 30min. Larvae were dissected in cold 1X PBS plus 0.3% (v/v) Triton X-100 and then fixed in 4% paraformaldehyde in 1X PBS for 20min. Tissues were dehydrated using several changes of methanol followed by rehydration in PBST (1X PBS + 0.3% Triton X-100) and blocking in PBST + 0.5% BSA for 2.5hrs. Tissues were incubated overnight at 4°C in a 1/50 dilution of a monoclonal anti-p53 antibody (Santa Cruz Biotechnology, Inc.) in 1X PBS + 0.5% BSA. Tissues were washed quickly several times followed by 3X 30min washes in PBST + 0.5% BSA. The secondary antibody (HRP) was added at a dilution of 1/300 for 45min after which time samples were washed in PBT for 1hr. The staining was developed using the ABC system of Pierce. After staining, tissues were washed in PBS and a glycerol solution (50:50 with 1X PBS) before dissection and mounting.
Table 2-1. Results of the yeast two-hybrid screen

<table>
<thead>
<tr>
<th></th>
<th>Bait #2</th>
<th>Bait #3</th>
</tr>
</thead>
<tbody>
<tr>
<td>Colonies screened</td>
<td>$1 \times 10^6$</td>
<td>$1 \times 10^6$</td>
</tr>
<tr>
<td>Colonies selected</td>
<td>300-400</td>
<td>800-900</td>
</tr>
<tr>
<td>Gal dependent colonies</td>
<td>35</td>
<td>100 classified into 4 categories*</td>
</tr>
<tr>
<td>Positive clones after retransformation</td>
<td>9</td>
<td>3</td>
</tr>
</tbody>
</table>

*only 8 were retransformed
Figure 2-1. Schematic diagram of the yeast two-hybrid screen. The LexA DNA binding domain (yellow) is fused to the Bait (green) and binds DNA through the LexA operator. The acidic activation domain B42 (purple) is fused to a library of proteins (black) representing cDNAs from the 0-12 hr time period of Drosophila embryogenesis. An interaction between the Bait and a library protein brings the B42 activation domain to the DNA, resulting in activation of the reporter genes LEU 2 and lacZ.
Figure 2-2. Bait constructs used in yeast two-hybrid screen. Comparison of the ftz gene from *Drosophila melanogaster* (upper sequence) and *Drosophila hydei* (lower sequence). The highly conserved N-terminal region (yellow) and homeodomain sequences (green) are indicated. Identities are indicated by lines and similarities are indicated by dots. Regions comprising each of the Bait constructs are shown.
Figure 2-3. Far Western assay. (A) FTZ constructs, as described adjacent to the figure were expressed in *E. coli* and separated by SDS-PAGE and stained with Coomasie blue. (B) Expression of each construct is verified by a Western blot using a polyclonal FTZ antibody. Fewer epitopes for the antibody are found in the HD, resulting in a weaker signal in the GST-FTZHD lane. (C) Labeled TAN binds to full-length FTZ, a C-terminal portion of FTZ containing the HD and the isolated HD of FTZ fused to GST in a Far Western assay. Interacting FTZ constructs are denoted by an asterisk.
Figure 2-4. Cloning *tan* from cDNA and genomic libraries. (A-B) Schematic diagrams of the *tan* cDNA clones. Coding sequences are in black, UTR sequences are in gray and the ATG and stop sites for translation are indicated. (A) The clone identified in the yeast two-hybrid screen. (B) Clones identified from the cDNA library screen. A full-length *tan* construct was created by pasting the two constructs together using the Bst EII site. (C) Genomic region surrounding the *tan* locus (black with introns in gray) based on clones identified from screening a genomic library.
Figure 2-5. Genomic and amino acid sequence of tan. (A) Sequence of the tan locus with introns (two boxed sequences) and putative polyadenylation signal (AATAAA) indicated. Arrows labeled 1, 2, and 3 denote the 5' ends of the cDNAs identified by, respectively, my screen and two ESTs from the Berkeley EST project (Flybase). (B) Schematic diagram of TAN indicating PEST (a PEST score of +5 or greater is considered significant; Rechsteiner and Rogers, 1996), Basic, and Acidic (AC) domains. (C) Alignment of residues 190-201 of TAN with the third helix of the HD from several different classes of HD containing proteins (see Laughon, 1991). Identities are dark gray while similarities are light gray.
Figure 2-6. Polytene in situ hybridizations. (A) In situ hybridization using a *tan* probe detects one signal on polytene chromosomes (arrowhead). The ends of the five large chromosome arms are indicated by arrows. (B) *tan* maps to 65A.
**Figure 2-7. Northern blot analysis.** Developmental Northern analysis of tan (top panel) detects an approximately 1.6kb signal. The lower panel shows the same blot stained with Methylene blue as a loading control. Larval, (L); Pupal, (P); Female, (F); and Male, (M).
Figure 2-8. Expression pattern of tan. (A) In situ hybridization to oocytes shows tan expression in Stage 10 oocytes. (B) Hybridization to a cleavage stage embryo with a tan sense probe. (C-F) In situ hybridizations using an anti-sense tan probe; (C) cleavage stage; (D) cellular blastoderm; (E) germ band extension; specific expression is seen in the somatic and visceral mesoderm (arrow) in addition to the ubiquitous transcript. For all embryonic stages anterior is to the left and dorsal up. (F) tan expression in an eye-antennal disc: higher expression is seen in the morphogenetic furrow (arrowhead).
Figure 2-9. An epitope tagged version of TAN localizes to the nucleus. (A) Left panel is in vitro transcribed and $^{35}$S-labeled TAN run on SDS-PAGE. Right panel is a Western blot of in vitro transcribed and translated p53-TAN (unlabeled) detected by anti-p53 antibodies. (B) Ectopic p53-TAN produced by heat shock (see text) is localized to the nucleus in salivary glands (arrow) and associated fat bodies (arrowhead) of a third instar larvae. (C) No signal is detected in wild-type larvae exposed to heat shock. (D) Higher magnification of the nuclei from (B) reveals TAN staining associated with the chromosomes (arrow).
Figure 2-10. TAN binds DNA. Top panel shows in vitro transcribed and radioactively labeled FTZ (Lane 1), FTZΔHD (Lane 2), TAN (Lane 3) and LUC (Lane 4) proteins analyzed by SDS-PAGE. Bottom panel shows the ability of each protein to bind to salmon sperm DNA blotted to a filter. TAN and FTZ bind strongly in this assay while FTZΔHD and LUC do not bind. The faster migrating band in Lane 3 most likely results from alternate translational start sites.
CHAPTER 3

Tantalus Interacts Physically and Genetically with the Polycomb- and trithorax-Group Member Additional sex combs

In this chapter, Dr. Hugh Brock’s laboratory performed the original yeast two-hybrid assay identifying Tantalus as an Additional sex combs-interacting protein (Michael Kyba) and mapped Tantalus binding sites on polytene chromosomes (Hugh Brock) (Figure 3-3 and Table 3-1). TAN antibodies were made by myself and the Brock lab. Shelley Lumba assisted with the characterization of the TAN antibodies. Experiments described in Figures 3-2B and 3-4A,B were performed by Jocelyn Moore. Gilbert Dos Santos and Fiona McCloskey assisted in the experiments described in Figure 3-4C-E. I performed the experiments described in the remaining Figures and Tables.
Abstract

The Drosophila trithorax- and Polycomb-groups of proteins maintain activated and repressed transcriptional states at specific target gene loci. The Additional sex combs (Asx) gene is of particular interest as it acts as a member of both protein complexes and its effects on target genes are tissue-limited. A novel protein, Tantalus (TAN), was identified in a yeast two-hybrid screen for ASX-interacting proteins that might confer tissue-specific activities. Although TAN contains consensus nuclear localization sites and binds DNA in vitro, the protein localizes to both the cytoplasm and nucleus. In salivary glands, TAN is nuclear and associates with 66 euchromatic sites on the polytene chromosomes, more than half of which overlap with ASX. Unlike the majority of trxG and PcG proteins, however, TAN does not appear to regulate homeobox complex genes. Rather, tan mutant phenotypes are specifically limited to sensory lineage defects and one of these phenotypes, shared by Asx, is genetically enhanced by Asx. Taken together, the data suggest that TAN may act as a tissue-specific cofactor for ASX, and that its activity may be controlled by subcellular trafficking.

Introduction

PcG and trxG genes are required to maintain states of repressed (PcG) or activated (trxG) transcription during Drosophila development, and mutations in these genes are most commonly associated with homeotic transformations in the adult fly (Simon, 1995; Kennison, 1995). Mutations in PcG genes, such as Polycomb or polyhomeotic, cause ectopic HOM-C activation, while mutations in trxG genes, such as trithorax, result in a loss of expression of HOM-C genes. In both cases, however, HOM-C gene expression is
initiated normally but the maintenance stage of HOM-C expression is compromised. Both groups appear to act as large protein complexes (reviewed in Pirrotta, 1998), and recent results support the idea that each group is involved in the maintenance of an “open” or “closed” configuration of nucleosomes around target genes (Kal et al., 2000; Shao et al., 1999).

Asx was classified as a member of the PcG because Asx mutations enhance the defects seen in other PcG mutants and cause PcG-like transformations (Breen and Duncan, 1986; Sinclair et al., 1992). As with other PcG members, the homeotic transformations seen in Asx mutants are associated with an inability to maintain repression of HOM-C genes (McKeon and Brock, 1991; Simon et al., 1992; Soto et al., 1995). However, one Asx allele, created by a P-element insertion within the Asx locus (AsxP1), causes opposite - anteriorly directed - homeotic transformations in the adults of homozygous flies (Sinclair et al., 1992), suggesting an inability to maintain the active state of HOM-C gene expression. This result, and the finding that several Asx alleles including AsxP1 can genetically enhance the homeotic phenotypes associated with the trxG member trithorax, suggested that Asx may be a member of both the PcG and trxG. ASX is not the first protein suggested to have both PcG and trxG activities: the Enhancer of zeste (E(z)) gene was originally identified as a PcG gene, but has since been shown to have trxG activity (LaJeunesse and Shearn, 1996). More recently, alleles of several genes that had previously been characterized as PcG genes were shown to be genetic enhancers of both trxG and PcG loci (Gildea et al., 2000). These results prompted the suggestion of an additional gene class, enhancers of trithorax and Polycomb mutations (ETP).
Unfortunately, there is little biochemical data available for ASX, E(Z) or proteins of the proposed ETP group, making it difficult to assess what their respective roles might be in PcG/trxG function. However, ASX associates with a limited number of sites on polytene chromosomes, suggesting that it plays a specific role in gene regulation.

Another interesting attribute of Asx is that its effects are tissue-limited. Asx mutations cause ectopic derepression of HOM-C genes in the epidermis and visceral mesoderm, but not in the central nervous system (CNS) where other trxG and PcG genes play important roles (Soto et al., 1995). Experiments with reporter constructs have demonstrated that ASX is active in the CNS however, suggesting that it is functional in this tissue. One interpretation of these results is that cofactors are required to facilitate tissue- and promoter-specific functions of particular PcG complex components.

A tissue-specific cofactor(s) for ASX activity would be expected to exhibit tissue-specific defects when mutated and to interact genetically with Asx mutations. ASX cofactors should also interact physically with ASX, either directly or indirectly. TAN and ASX interact directly in yeast two-hybrid and GST pull-down experiments, and co-localize at a subset of sites on polytene chromosomes from third instar larvae. Asx and tan mutations also interact genetically. Interestingly, TAN is unlike other trxG/PcG complex components in that it exhibits differential subcellular localization. It is also unusual in that it appears to play no role in HOM-C gene regulation. Rather, TAN appears to be a novel cofactor of ASX that specifically facilitates a role in sensory organ development.
Results

Identifying novel Additional sex combs cofactors

The Asx protein contains 1668 amino acids (Sinclair et al., 1998b). The C-terminal 500 amino acids constitute the main homology region with respect to mammalian counterparts (E. O'Dor and H.W.B., unpublished results), and includes a cysteine cluster with 25/28 conserved residues. As most conserved PcG protein domains are required for protein-protein interactions (Peterson et al., 1997; Gunster et al., 1997; Kyba and Brock, 1998; Hashimoto et al., 1998), this region (residues 1139-1668) was used to make a yeast two-hybrid bait construct. A 0-12 hr embryonic cDNA library was screened, and out of approximately $10^5$ clones, 11 unique interacting constructs were recovered (M.K., J.M. and H.W.B., in preparation). One of these clones was found to encode the tantalus gene. The tan clone did not interact with control constructs, nor did it interact with 23 of 24 two-hybrid baits constructed from 11 PcG or PcG-related proteins.

To verify the two-hybrid interaction between ASX and TAN, a GST pull-down assay was used. Residues 1139-1668 of ASX were used to make a GST-ASX fusion protein. GST and GST-ASX were purified and bound to glutathione beads. The two resins were then incubated with $^{35}$S-labeled TAN. TAN was not precipitated by GST, but was efficiently precipitated by the GST-ASX fusion protein (Figure 3-1). These results suggest that ASX and TAN are likely to interact directly.

TAN and ASX binding sites overlap on polytene chromosomes

Staining of polytene chromosomes with antibodies against PcG and trxG members reveals limited but overlapping euchromatic binding for many members, and is a
hallmark of PcG/trxG proteins (see Chinwalla et al., 1995; Rastelli et al., 1993). To determine whether TAN is capable of binding polytene chromosomes in vivo, polyclonal antibodies were raised against GST fusions of both full-length and peptide portions of TAN. As a control, the p53 tagged version of TAN was employed (see page 72). tan and p53-tan cDNAs were then placed in appropriate vectors for induction by heat shock (HS) or the GAL4/UAS system, and transgenic lines were produced.

Figure 3-2 shows a Western blot from wild-type (WT), HS-tan and HS-p53-tan third instar larvae which have been exposed to HS. A signal that runs at approximately 55kDa is detected by the TAN antibody, and this signal is dramatically increased when extracts were made from HS-tan or HS-p53-tan fly lines (Figure 3-2A). Monoclonal antibodies specific for the p53 epitope tag also recognized this inducible protein (Figure. 3-2A). A protein of similar size was detected by affinity purified serum in nuclear extracts of cultured Drosophila Kc cells (Figure 3-2B), and by in vitro transcription and translation of the tan cDNA (Figure 3-1). Although 55kDa is significantly larger than the predicted TAN molecular weight of 33kDa, many transcription factors are known to run aberrantly on SDS-PAGE gels (Krause and Gehring, 1988; Martin and Adler, 1993; Yamamoto et al., 1997).

The affinity purified anti-TAN antibodies were then used to stain third instar polytene chromosomes to determine whether endogenous TAN binds discrete loci. Approximately 66 discrete euchromatic sites recognized by TAN were reliably detected and mapped (Figure 3-3 and Table 3-1). Over half of these sites (35/66) overlap previously mapped ASX binding sites (Sinclair et al., 1998b), suggesting that a subset of
ASX-regulated loci may be co-regulated by TAN. Double staining with TAN and ASX antibodies confirmed that a number of these sites overlap (H.W.B unpublished observations). TAN also binds to a number of Polycomb/polyhomeotic binding sites not recognized by ASX, indicating that TAN may have functions that are independent of ASX. Somewhat surprisingly, two of the PcG/trxG sites that do not appear to interact with anti-TAN antibodies are those of the ANT-C (84A,B) and BX-C (89E) homeotic gene clusters (Figure 3-3B). These results suggest that TAN function is likely to be unrelated to homeobox gene activities.

*Cellular distribution of Tantalus*

Affinity purified TAN antibodies were also used to follow endogenous TAN. Although patterns of expression similar to those detected by in situ hybridization were detected (Figure 3-4A,B), the ubiquitous nature of TAN expression made it difficult to definitively determine whether endogenous TAN was localized to the nucleus or cytoplasm. However, while following the expression of TAN and p53-TAN from the HS and UAS transgenic lines, it was observed that TAN was not found exclusively localized to the nucleus. The proteins detected by anti-TAN and anti-p53 antibodies showed similar subcellular distributions and, unexpectedly, this localization varied in different tissues. For example, when expressed in imaginal discs using a patched-GAL4 driver, TAN is clearly enriched in the cytoplasm (Figure 3-4C-E). However, when expressed in salivary glands or associated fat bodies, the protein is nuclear (see Figure 2-9). Further studies will be required to catalogue protein distributions in different tissues and to determine how these distributions affect activity.
Creating a tantalus null allele

As TAN and ASX bind to a subset of shared sites on polytene chromosomes, I was interested in obtaining mutations in tan to test for a genetic interaction between the two genes. tan is located at 65A on chromosome 3. Searches of Flybase and Southern analysis of P-elements within the 65A region (data not shown) were unsuccessful in identifying a mutation within the tan locus. Therefore, a P-element mutagenesis screen was undertaken. Deák and coworkers (1997) have created an extensive library of P-element inserts on the third chromosome and several of these lines were obtained to begin the screen. Three different P-elements in the 65A region were mobilized to generate 270 independent lines for screening (see Materials and methods). Using PCR, one line containing a P-element inserted near the start site of tan transcription (hereafter referred to as tan') was identified (Figure 3-5A).

A tan null allele was created by mobilizing the P-element and using PCR to screen for imprecise excisions that disrupted tan, but not an upstream gene identified from the genomic clones. After screening over 1000 lines, one deletion was identified which removed approximately 1.4kb of tan sequence, leaving all sequences upstream of the P-element insert site intact (Figure 3-5B-E). Sequencing through the deleted region confirmed the removal of 1.4kb of DNA. This allele is designated as tan^2 and appears to be a null allele (see below) as the deletion removes all of the 5' UTR and most of the coding region, extending to residue 266.

During the Southern analysis to confirm the P-element inserts and excisions, several background bands were detected which were originally attributed to probes
detecting homologues of the serine protease gene located upstream to tan (Figure 3-5B). However, the use of a specific tan probe still resulted in cross-reaction with the strongest background band under highly stringent conditions (Figure 3-5D), suggesting a homologue to tan may exist. The annotated Drosophila genome does not reveal such a homologue to date but numerous gaps still exist in the sequence (Adams et al., 2000; Flybase).

**tantalus mutant phenotypes**

Despite the fact that tan is widely expressed during much of development, tan' flies, which have the majority of the gene deleted, are homozygous viable and fertile (as are their progeny). Examination of both tan' and tan' homozygous adults (unless stated otherwise, tan' and tan' mutants described from hereon are homozygous), however, reveals a number of morphological defects. These include a rough-eye phenotype, the loss or duplication of sensory bristles, and shortened veins in the wings. A common feature of each of these tissues is that they comprise or contain enervated sensory organs. These defects are discussed further below.

Consistent with the expression of tan in the eye-antennal disc (see Figure 2-8F), tan' mutants have a rough eye appearance due to ommatidia defects and the deletion or duplication of ommatidial bristles (Figure 3-6A-D). Wild-type ommatidia have a uniform size and hexagonal appearance with bristles equally spaced at three of the six corners (Figure 3-6C). In contrast, many tan' mutants have ommatidia that are smaller, and occasionally the ommatidia are deleted altogether or fused with neighbors (Figure 3-6B,D). Bristles tend to be shifted adjacent to one another or are missing altogether.
In wild-type flies the medial part of the adult head contains three light sensitive organs called ocelli with numerous sensory bristles spaced around (macrochaetae) and between (microchaetae) them (Figure 3-6E). There are approximately eight microchaetae, referred to as interocellar bristles, between the ocelli and the number of these bristles is greatly reduced in \( \text{tan}^1 \) and \( \text{tan}^2 \) mutants (Figure 3-6F and Table 3-2). This phenotype was rescued by two transgenic copies of genomic \( \text{tan} \) sequence (referred to as P[\( \text{tan} \)]) which included 1.4kb of upstream and 3kb of downstream sequence (Figure 3-6G). Strikingly, 79% of \( \text{tan}^2 \) mutants have five or fewer bristles compared with 2% of the rescued flies, with no rescued flies having fewer than five bristles. The mutants averaged 4.8 bristles per fly versus 7.5 bristles per rescued fly, similar to wild-type, demonstrating that the mutants are deficient in the ability to properly specify these microchaetae.

\( \text{tan} \) mutants also have other bristle and sensory organ defects. These include macrochaetae duplications, predominantly of scutellar bristles (Figure 3-6H and Table 3-2), while males frequently have ectopic hairs in the A6 sternite, a segment normally devoid of hairs (Figure 3-6I and Table 3-2). The penetrance of this phenotype was increased when flies were reared at 30°C (Table 3-2). The bristle pattern of the abdominal tergite also appears to be modified variably, with many of the mutants showing smaller, irregularly spaced bristles and ectopic bristles (data not shown). Finally, \( \text{tan} \) mutants also frequently have a shortened fifth wing vein (Figure 3-6J and Table 3-2).

These phenotypes are also fully or partially rescued by two copies of P[\( \text{tan} \)] (Table 3-2 and data not shown), verifying that the phenotypes are not a consequence of second site mutations. To further confirm \( \text{tan} \) mutations as the cause of the sensory defects
described above, \textit{tan} mutants were crossed to a deletion that uncovers the 65A region (Table 3-2 and data not shown). Hemizygous flies display phenotypes with similar frequencies as those of homozygous \textit{tan}^1 and \textit{tan}^2 mutants. This suggests that \textit{tan}^1 is a strong hypomorph and that \textit{tan}^2 is a null allele. These results are also consistent with TAN being the sole source of the observed defects.

\textit{Developmental defects caused by ectopic tantalus expression}

\textit{tan} and \textit{p53-tan} were expressed under the control of several different promoters to determine whether ectopic expression or over-expression might induce additional developmental defects that shed further light on the function of the protein. As a first approach, the genomic \textit{tan} rescue construct was introduced into wild-type flies to double the gene copy number to 4. These flies were viable and had subtle effects on sensory tissues. For example, of the two lines generated, one (Line A) showed deletions of the macrochaetae referred to as post-vertical bristles, located just posterior to the ocelli (Figure 3-7A). The second line (Line B) had a more variable effect on these bristles, including reductions in size (Figure 3-7B), occasional deletions with sockets remaining (Figure 3-7C), and rare deletions where the sockets were also missing, similar to Line A. This graded effect suggests that \textit{tan} may function at different stages of bristle specification and that the level of \textit{tan} activity is important for these functions.

\textit{HS-tan} and \textit{HS-p53tan} lines reared at 30°C to induce low levels of \textit{tan} expression also had frequent deletions of the post-vertical bristles (up to 40% of the flies had one or both bristles missing). Interestingly, all HS lines also had ectopic bristles (usually three or four) in the A6 sternite of males (Figure 3-7E). This phenotype was found frequently and,
surprisingly, was more severe than the \( tan^2 \) mutant phenotype. These phenotypes were not seen in control crosses (Table 3-2).

In order to increase the levels of \( tan \) expression, and to express the protein ectopically, the UAS-\( tan \) and UAS-\( p53-tan \) constructs described earlier were expressed under the control of various GAL4 drivers. When \( tan \) was expressed under the control of a patched-GAL4 driver, adults were frequently missing bristles in the ocellar and scutellum regions, and occasionally all bristles from the ocellar region were missing (Figure 3-7D). Ectopic UAS-\( tan \) expression under the control of a hairy-GAL4 driver was larval lethal. However, when grown at 18°C, escapers were recovered from one line, and these occasionally displayed hypertrophy of the wing veins. Similar wing vein defects were seen in one (Line A) of the two lines with 4 copies of endogenous \( tan \) (Figure 3-7F). These wing vein effects are reciprocal to those observed in \( tan \) mutants.

**Genetic interaction between** Additional sex combs and tantalus

The results above support a limited role for TAN during development since both \( tan \) over- and under-expression result in specific defects in sensory lineages. To test whether ASX and TAN act together to control sensory organ specification, \( Asx \) and \( tan \) mutants were combined and \( tan \) bristle phenotypes scored. The highly penetrant and well characterized interocellar bristle phenotype was focused on since \( tan \) mutants have an easily quantifiable effect on the number of these bristles (see Figure 3-6G). To ensure a stringent assessment, only flies with 4 or fewer interocellar bristles were scored.

To determine if \( Asx \) plays a role in bristle specification, several \( Asx \) gain-of-function (GOF) alleles (Sinclair et al., 1992) were crossed to \( tan \) mutants to test for a
specific genetic interaction. The Asx\textsuperscript{t5} and Asx\textsuperscript{y} alleles both strongly rescued the tan interocellar bristle phenotype (Table 3-3). Because these alleles were made in the same background, another GOF allele, Asx', was also tested. This allele also rescued the tan phenotype (Table 3-3).

Additionally, a loss-of-function (LOF) Asx allele, Df(2R)trix, was crossed to tan mutants. In contrast to the GOF alleles, this allele enhanced the penetrance of the tan bristle phenotype when compared to sibling controls (Table 3-3). Note, however, that the penetrance of the tan phenotype on its own is reduced in the Df(2R)trix/CyO genetic background (only 2-5\% of CyO;tan/tan flies versus 37\% of tan\textsuperscript{y} homozygous flies exhibit the phenotype). Although the basis of this suppression is unknown, this suppression effect has been observed previously with this Asx allele (see Table II in Milne et al., 1999).

To confirm this genetic enhancement, the genomic tan rescue line was crossed to Df(2R)trix/CyO;tan\textsuperscript{y}/tan\textsuperscript{y} flies. Introduction of one copy of P[tan] completely rescues the tan bristle phenotype in CyO progeny and, as expected if Asx is involved in interocellar bristle specification, only partially rescues the bristle phenotype in Df(2R)trix siblings (Table 3-3). To verify that the Df(2R)trix allele affects bristle specification, this allele was out-crossed to control flies. Seven percent (n=167) of the heterozygous Df(2R)trix progeny exhibited the bristle phenotype, while all control siblings were wild-type (n=88), demonstrating that the Df(2R)trix allele affects interocellar bristle specification.

Asx mutants produce homeotic defects including ectopic sex combs on the T2 legs of males. These transformations were not enhanced by the presence of tan, nor did
deletions of tan alter the expression patterns of the homeotic genes Antennapedia, Ultrabithorax or Sex combs reduced (data not shown), consistent with the inability to detect binding of TAN at the sites of the ANT-C and BX-C on polytene chromosomes. Notch is a genetic modifier of tantalus

The effects of tan on bristle and wing vein differentiation resemble the effects generated by certain alleles of Notch (N) (see Lindsley and Zimm, 1992) while expression of tan in the wing margin using the margin-specific C96-GAL4 driver caused severe notching of the wings (Figure 3-8), also reminiscent of certain alleles of N. As N functions within a field to specify individual cell fates (reviewed in Artavanis-Tsakonas et al., 1999) it is possible that N signalling could be involved in limiting TAN function, a potential requirement since tan is ubiquitously expressed but appears to be active in only a subset of cells. To test whether N and tan interact genetically, loss-of-function N alleles were crossed with tan2 flies to generate trans-heterozygotes (Table 3-4). As shown earlier, tan2 heterozygotes exhibit no detectable defects in interocellar bristle formation. However, in the presence of N, deletions of the interocellar bristles were observed with very high frequency (75% of N+/+;tan2/+ flies displayed the interocellar bristle phenotype). This phenotype was reverted by crossing in the genomic tan rescue construct (Table 3-4).
Discussion

_Tantalus, a new Additional sex combs cofactor_

A major objective of this study was to identify ASX-interacting proteins that might help explain the tissue-specific activities of ASX (Soto et al., 1995). Such tissue-specificity could result from many different mechanisms, including the tissue-specific expression of different PcG/trxG genes or through the actions of tissue-specific cofactors. Since Asx does not show tissue-specific patterns of expression, it was speculated that tissue-specific cofactors might exist (Soto et al., 1995; Sinclair et al., 1998b).

TAN was identified as an ASX-interacting protein using a yeast two-hybrid screen, and the interaction was confirmed by a GST pull-down assay. TAN shows no extended regions of homology to other proteins currently listed in the databases but does feature several general properties typical of transcription factors. These include stretches of basic residues, consensus nuclear localization motifs and a short motif found in other DNA binding proteins. Consistent with these properties, TAN binds DNA in vitro, is enriched in the nucleus in a subset of tissues and associates with 66 sites on polytene chromosomes.

TAN also displays properties that are consistent with it being an important cofactor for ASX. It associates with ASX in vitro, co-localizes with ASX at 35 of 66 TAN chromosomal binding sites and interacts genetically with Asx in vivo. The ability of TAN to bind DNA in vitro suggests a possible role in recruiting or anchoring ASX-containing complexes to specific sites on DNA.
Additional sex combs and Tantalus control sensory organ development

Most members of the PcG and trxG protein complexes characterized thus far associate with both the ANT-C and BX-C gene clusters and cause homeotic transformations when mutated. However, TAN appears to have no role in the regulation of these genes. Rather, TAN appears to play a specific role in the differentiation of sensory organs. The data show that ASX is also required in at least a subset of these tissues, and when tan and Asx mutations are combined, these sensory organ defects are specifically enhanced or suppressed. Taken together, these results suggest that TAN acts as a tissue-specific ASX cofactor in sensory organ differentiation. This study only examined the role of tan and Asx in ocellar bristle development where, possibly because of small field in which bristle specification occurs, bristle fate is extremely sensitive to tan over- or under-expression. Further work will be required to determine if ASX acts together with TAN to regulate all other tan-dependent processes identified. This need not be the case, as ASX only colocalizes with 35/66 of the mapped TAN polytene chromosome binding sites.

Although the majority of PcG and trxG components identified to date were isolated via genetic screens for the suppression or enhancement of homeotic transformations, loss of sensory bristle phenotypes has been observed when other trxG/PcG genes are misexpressed. Examples include LOF mutations in the trxG genes brahma (Elfring et al., 1998), absent small/homeotic disks2 (Adamson and Shearn, 1996), leg arista wing complex (Zorin et al., 1999) and Asx (this study). Loss of bristle phenotypes are also observed after ectopic expression of the trxG gene osa (Collins et al.,
1999) and the PcG genes *Posterior sex combs* and *Suppressor 2 of Zeste* (Sharp et al., 1994). Additional studies will be required to determine which bristle genes TAN and ASX are acting through, and whether this regulation is positive or negative in nature.

The defects in sensory organ tissues caused by *tan* misexpression are remarkably similar to those caused by mutations in the *Notch* (*N*) gene (see Lindsley and Zimm, 1992). In addition, expression of *tan* in the wing margin causes severe notching of the wings and *N* mutations specifically and strongly enhance the *tan* interocellar bristle phenotype. *N* is required at two different stages of sensory organ development, during specification of the sensory organ precursor (SOP) cells and during the subsequent specification of the SOP's daughter cells (reviewed in Artavanis-Tsakonas et al., 1999). Similarly, the results presented here demonstrate that TAN can affect successive stages of bristle development; over- or under-expression of TAN can cause bristle and socket loss.

One additional observation regarding the effect of *tan* on sensory bristles needs comment. The presence of ectopic bristles in the A6 sternite of *tan* mutant male flies is characteristic of a loss of *Abdominal-B* (*Abd-B*) expression, and is a phenotype exhibited by mutations in *trx* and *trithorax-like* of the *trxG* (Farkas et al., 1994; Ingham, 1981). However, ectopic expression of *tan* using the HS promoter resulted in an even more severe anterior homeotic transformation of male sternites. Considering this finding, and the other evidence presented here for a sensory-specific role for TAN, it seems likely that the ectopic hairs in the A6 sternites of males are a result of TAN activity in bristle specification and not an effect on *Abd-B* expression. Such an interpretation is consistent
with the inability to detect TAN binding at the site of the BX-C (where Abd-B is located) in polytene chromosomes.

**Tantalus subcellular localization**

An interesting feature of TAN that has not previously been noted for other trxG or PcG complex components is that its subcellular localization varies in a tissue-specific manner. In embryos, protein expressed by the GAL4 system is primarily cytoplasmic, while in third instar larvae it is cytoplasmically enriched in some tissues and localized to the nucleus in others. Although the mechanism and importance of this localization has yet to be addressed, it suggests a novel means of functional regulation.

To date, the number of transcription factors known to cycle between cytoplasm and nucleus are relatively few. Other well-characterized examples include proteins such as Dorsal, Arm/β-catenin, STAT proteins, MAD proteins and components of the Notch signaling pathway, including N itself (Struhl and Adachi, 1998). The last example is particularly notable, given the genetic interaction observed between N and tan.

Control of TAN nuclear localization may explain the restriction of mutant phenotypes to sensory lineages, despite fairly ubiquitous patterns of gene expression. Retention in the cytoplasm may be a general way of relegating TAN activity to a subset of TAN-expressing cells. On the other hand, the cytoplasmically localized protein may also serve a function that has yet to be elaborated.

**Other roles of Tantalus**

Another explanation for the relatively specific nature of tan mutant phenotypes, despite the prolonged and widespread expression of the gene, is that some of its functions
are obscured by redundantly acting gene products. Several additional observations lend weight to this argument. First, *tan* bristle phenotypes are highly variable in terms of penetrance and severity. Second, more severe phenotypes can be induced by ectopic- or over-expression. Third, slightly less than half of the TAN polytene chromosome binding sites do not appear to colocalize with ASX. The mapped TAN binding sites also show heterogeneity with respect to other PcG/trxG protein binding sites. This heterogeneity of protein complexes at different loci suggests the likelihood of different functions and outputs for each protein complex. Revealing the full extent of these TAN complex activities will likely require the elimination of redundantly acting factors.
Materials and methods

GST pull-downs

Amino acids 139-1668 of ASX were subcloned into pGEX4T1. GST beads were bound to GST and GST-ASX at a concentration of 2μg/μl for GST and 400ng/μl for GST-ASX. Beads were resuspended in an equal volume of 1X Binding Buffer1 (1X BB1: 25mM HEPES pH 7.5, 5mM KCl, 1mM EDTA, 0.25mM DTT, 0.05% Tween-20, 1mg/mL BSA, 10% glycerol). The Promega TNT T7 Rabbit Reticulocyte Lysate System was used to synthesize full length TAN (25μl reaction) which was passed over a Sephadex G-25 column and diluted to 100μl with 1X BB1. The probe was incubated at 4°C for 30min to block and then 50μl added to each of the GST and GST-ASX beads (total volume is 100μl). Reactions were incubated at 4°C with mixing for 2hrs and then washed 2X with 1X BB1 and 2X with Wash Buffer (20mM HEPES pH 7.5, 0.15M NaCl. 10% glycerol). Bound probe was eluted with 15μl of 1X SDS buffer and analyzed by SDS-PAGE.

Antibodies

Either full-length TAN or residues 87-164 were fused to GST and purified for injection to rabbits. Polyclonal antibodies were purified as described (Sinclair et al., 1998b) with both antibodies giving similar results. Kc extracts were made as described previously (Kyba and Brock 1998). Polytene chromosomes were stained as described (Sinclair et al. 1998b) with a 1/75 dilution of affinity purified anti-TAN antibody. Serum depleted by passage over columns containing GST-TAN did not react with polytene
chromosomes, demonstrating the specificity of the antibody. Double-staining of polytene chromosomes was undertaken with the sheep anti-ASX antibody described previously (Sinclair et al. 1998b) and the rabbit anti-TAN antibody described here. Binding was detected with appropriate secondary antibodies labeled with Alexa Fluor 488 (Molecular Probes) and Cy3 (Jackson Labs). To immunostain imaginal discs, larvae were dissected and fixed as described (Hughes and Krause, 1999). Fixed tissues were dehydrated using several changes of methanol followed by washes in PBST (1X PBS + 0.3% Triton X-100). Tissues were blocked in PBST + 0.5% BSA for 2.5hrs. Tissues were incubated overnight at 4°C in appropriate antibodies (1/50 dilution of a monoclonal anti-p53 antibody (Santa Cruz Biotechnology, Inc.) or 1/500 dilution of anti-TAN against full length TAN) in 1X PBS + 0.5% BSA. Tissues were washed quickly several times followed by 3X 30min washes in PBST + 0.5% BSA. Tissues were incubated in secondary antibodies for 45min (1/300 anti-mouse HRP (Bio-Rad) or 1/300 anti-rabbit Alexa Fluor 488 (Molecular Probes)) followed by washes in PBT for 1hr. For fluorescent staining, tissues were incubated in propidium iodide to stain nuclei. The HRP staining was developed using the ABC system of Pierce. After staining HRP tissues were washed in PBS and a glycerol solution (50:50 with 1X PBS), while fluorescent tissues were washed in 70% glycerol/2% DABCO, before dissection and mounting.

P-element screen

The screen was performed based on the protocol of Hamilton and Zinn (1994). Lines 0545/01, 0666/10 and 1203/07 from Deák and coworkers (1997) were used for the P-element mutagenesis screen. Groups of 20 flies containing hopped P-elements were
analyzed by PCR using the P-element inverted repeat sequence as a primer and several primers from the tan genomic sequence. PCR was performed according to standard procedures (Ausubel et al., 1997). The following primers were used:

P-element inverted repeat 5'CGACGGGACCACCTTATGTATTTTCATCATG'3

13-1 5'CAGCGATTGCATCAGTG'3

13-2 5'GATTGTCTCCAGATTTGATG'3

G-1 5'GGCTGAACCCAAGTACACTA'3

G-2 5'CCGAACGACGTGGATATGTG'3

One line was obtained from 0545/01 that contained an insert in tan (see Figure 3-5). The 0545/01 line contains a lethal mutation located outside of the 65A region which was recombined off the chromosome to make homozygously viable tan'/tan' flies. A similar procedure was used to screen for deletions of tan after hopping the P-element from the gene. Southern blots were performed as described in Chapter 2 except high stringency washes were carried out at 68°C.

Genomic rescue

The rescue construct is approximately 7kb in size consisting of 1477bp upstream of tan (a Bam H1 site) and 3kb downstream of tan (a Sal I site) cloned into pW8 (Klemenz et al., 1987). Flies were made homozygous for both the rescue construct and tan2 deletion (P[tan]/P[tan];tan2/tan2).

Drosophila strains, crosses, and analysis

AsxJ, AsxQ, AsxT, and Df(2R)trix have been described (Sinclair et al., 1992) and the wild-type strain used is Oregon-R. N8 and N55el alleles of N are described in Lindsley
and Zimm (1992). patched-GAL4, hairy-GAL4 and C96-GAL4 are described in Flybase (http://flybase.bio.indiana.edu:82). Crosses for genetic analysis were 15-20 males and females mated for 3-4 days. Parents were either transferred to new bottles or dumped and progeny allowed to hatch until Day 17 for counting (parents are introduced to bottles at Day 0). For the genomic rescue of tan\(^2\) mutant bristles (Figure 3-6C) 200 flies were randomly selected at Day 17. Tissue samples were prepared as described (Sinclair et al., 1992).
Table 3-1. Polytene chromosome binding sites for TAN, with ASX and PC/PH sites indicated.

<table>
<thead>
<tr>
<th>X</th>
<th>3L</th>
</tr>
</thead>
<tbody>
<tr>
<td>1D</td>
<td>ASX</td>
</tr>
<tr>
<td>1D</td>
<td>61F, ASX, PC/PH</td>
</tr>
<tr>
<td>3D</td>
<td>ASX</td>
</tr>
<tr>
<td>3D</td>
<td>62F, ASX, PC/PH</td>
</tr>
<tr>
<td>6A</td>
<td>ASX</td>
</tr>
<tr>
<td>6A</td>
<td>63A, ASX</td>
</tr>
<tr>
<td>8A</td>
<td>ASX, PC/PH</td>
</tr>
<tr>
<td>8A</td>
<td>64A, PC/PH</td>
</tr>
<tr>
<td>9CD</td>
<td>ASX</td>
</tr>
<tr>
<td>9CD</td>
<td>64C, ASX</td>
</tr>
<tr>
<td>10F</td>
<td>ASX</td>
</tr>
<tr>
<td>10F</td>
<td>65B</td>
</tr>
<tr>
<td>13E</td>
<td>ASX, PC/PH</td>
</tr>
<tr>
<td>13E</td>
<td>66A, ASX</td>
</tr>
<tr>
<td>19D</td>
<td>ASX, PC/PH</td>
</tr>
<tr>
<td>19D</td>
<td>67D, ASX, PC/PH</td>
</tr>
<tr>
<td>19D</td>
<td>68C</td>
</tr>
<tr>
<td>2L</td>
<td>ASX</td>
</tr>
<tr>
<td>2L</td>
<td>70A, ASX, PC/PH</td>
</tr>
<tr>
<td>2L</td>
<td>70D, ASX, PC/PH</td>
</tr>
<tr>
<td>21A</td>
<td>ASX, PC/PH</td>
</tr>
<tr>
<td>21A</td>
<td>71EF</td>
</tr>
<tr>
<td>24A</td>
<td>ASX, PC/PH</td>
</tr>
<tr>
<td>24A</td>
<td>74EF</td>
</tr>
<tr>
<td>25EF</td>
<td>ASX, PC/PH</td>
</tr>
<tr>
<td>25EF</td>
<td>75F</td>
</tr>
<tr>
<td>27B</td>
<td>ASX</td>
</tr>
<tr>
<td>27B</td>
<td>80A</td>
</tr>
<tr>
<td>29E</td>
<td>PC/PH</td>
</tr>
<tr>
<td>29E</td>
<td>82C</td>
</tr>
<tr>
<td>30A</td>
<td>ASX</td>
</tr>
<tr>
<td>30A</td>
<td>3R</td>
</tr>
<tr>
<td>33C</td>
<td>PC/PH</td>
</tr>
<tr>
<td>34A</td>
<td>84D, ASX, PC/PH</td>
</tr>
<tr>
<td>34D</td>
<td>84F, ASX, PC/PH</td>
</tr>
<tr>
<td>35A</td>
<td>ASX</td>
</tr>
<tr>
<td>38D</td>
<td>85D, ASX</td>
</tr>
<tr>
<td>38D</td>
<td>86C, ASX</td>
</tr>
<tr>
<td>38D</td>
<td>87EF, PC/PH</td>
</tr>
<tr>
<td>38D</td>
<td>88C</td>
</tr>
<tr>
<td>39EF</td>
<td>ASX, PC/PH</td>
</tr>
<tr>
<td>39EF</td>
<td>89A</td>
</tr>
<tr>
<td>41C</td>
<td>ASX, PC/PH</td>
</tr>
<tr>
<td>41C</td>
<td>89B, ASX, PC/PH</td>
</tr>
<tr>
<td>42B</td>
<td>91D</td>
</tr>
<tr>
<td>42D</td>
<td>92C</td>
</tr>
<tr>
<td>44C</td>
<td>ASX</td>
</tr>
<tr>
<td>44C</td>
<td>93E, ASX, PC/PH</td>
</tr>
<tr>
<td>46A</td>
<td>ASX</td>
</tr>
<tr>
<td>46A</td>
<td>94D, PC/PH</td>
</tr>
<tr>
<td>47C</td>
<td>ASX</td>
</tr>
<tr>
<td>47C</td>
<td>95A</td>
</tr>
<tr>
<td>48A</td>
<td>ASX, PC/PH</td>
</tr>
<tr>
<td>48A</td>
<td>96A, ASX</td>
</tr>
<tr>
<td>50A</td>
<td>ASX</td>
</tr>
<tr>
<td>50A</td>
<td>97E</td>
</tr>
<tr>
<td>52E</td>
<td>ASX</td>
</tr>
<tr>
<td>52E</td>
<td>98C, ASX, PC/PH</td>
</tr>
<tr>
<td>58CD</td>
<td>PC/PH</td>
</tr>
<tr>
<td>59F</td>
<td>ASX, PC/PH</td>
</tr>
</tbody>
</table>
Table 3.2. Analysis of tan mutant and rescued phenotypes.

<table>
<thead>
<tr>
<th></th>
<th>A6 Sterne</th>
<th>Interocellar Bristles</th>
<th>Ectopic Scutellars</th>
<th>Wing Vein</th>
</tr>
</thead>
<tbody>
<tr>
<td>tan\textsuperscript{2}/tan\textsuperscript{2}</td>
<td>24.7 % (361)</td>
<td>36.7 % (387)</td>
<td>8.3 % (387)</td>
<td>85.1 % (174)</td>
</tr>
<tr>
<td>P[tan]/P[tan]; tan\textsuperscript{2}/tan\textsuperscript{2}</td>
<td>5.5 % (274)</td>
<td>0 % (357)</td>
<td>0.84 % (357)</td>
<td>30.8 % (120)</td>
</tr>
<tr>
<td>tan\textsuperscript{1}/tan\textsuperscript{1}</td>
<td>19.6 % (363)</td>
<td>29.3 % (314)</td>
<td>5.1 % (314)</td>
<td>ND</td>
</tr>
<tr>
<td>P[tan]/P[tan]; tan\textsuperscript{1}/tan\textsuperscript{1}</td>
<td>0.9 % (221)</td>
<td>0.67 % (450)</td>
<td>0 % (450)</td>
<td>ND</td>
</tr>
<tr>
<td>CH4/tan\textsuperscript{2} (30°C)</td>
<td>36.0 % (344)</td>
<td>40.4 % (383)</td>
<td>8.6 % (713)</td>
<td>ND</td>
</tr>
<tr>
<td>TM3. Sb/tan\textsuperscript{2} (30°C)</td>
<td>4.5 % (308)</td>
<td>3.4 % (357)</td>
<td>3.8 % (677)</td>
<td>ND</td>
</tr>
<tr>
<td>CH4/tan\textsuperscript{1} (30°C)</td>
<td>29.0 % (345)</td>
<td>33.3 % (342)</td>
<td>11.2 % (712)</td>
<td>ND</td>
</tr>
<tr>
<td>TM3. Sb/tan\textsuperscript{1} (30°C)</td>
<td>4.6 % (345)</td>
<td>1.3 % (377)</td>
<td>2.7 % (737)</td>
<td>ND</td>
</tr>
<tr>
<td>tan\textsuperscript{2}/tan\textsuperscript{2} (30°C)</td>
<td>57.8 % (128)</td>
<td>ND</td>
<td>ND</td>
<td>ND</td>
</tr>
<tr>
<td>P[tan]/P[tan]; tan\textsuperscript{2}/tan\textsuperscript{2} (30°C)</td>
<td>14.9 % (181)</td>
<td>ND</td>
<td>ND</td>
<td>ND</td>
</tr>
<tr>
<td>tan\textsuperscript{1}/tan\textsuperscript{1} (30°C)</td>
<td>38.5 % (200)</td>
<td>ND</td>
<td>ND</td>
<td>ND</td>
</tr>
<tr>
<td>P[tan]/P[tan]; tan\textsuperscript{1}/tan\textsuperscript{1} (30°C)</td>
<td>3.7 % (191)</td>
<td>ND</td>
<td>ND</td>
<td>ND</td>
</tr>
<tr>
<td>HS-tan 4.1 (30°C)</td>
<td>76.9 % (108)</td>
<td>ND</td>
<td>ND</td>
<td>ND</td>
</tr>
<tr>
<td>HS-p53-tan 37.3 (30°C)</td>
<td>69.3 % (88)</td>
<td>ND</td>
<td>ND</td>
<td>ND</td>
</tr>
<tr>
<td>HS-p53-tan 29.4 (30°C)</td>
<td>97.1 % (69)</td>
<td>ND</td>
<td>ND</td>
<td>ND</td>
</tr>
</tbody>
</table>

The % of flies displaying the indicated phenotypes and the number of flies analyzed (in brackets) are shown for each cross. P[tan] is the genomic rescue construct. TM3. Sb is a balancer chromosome and CH4 is a deletion of 64E-6SB1.2. For the CH4 containing cross, reciprocal matings were performed and, as the results did not vary, the data were pooled. All crosses except those noted were at room temperature. HS-tan 4.1, HS-p53-tan 37.3, and HS-p53-tan 29.4 are independent lines with the tan gene (or a p53 epitope tagged version of tan) under the control of a HS promoter.

\textsuperscript{a} refers to flies with at least 1 ectopic hair in the A6 sternite of males
\textsuperscript{b} refers to flies with 4 or fewer interocellar bristles
\textsuperscript{c} refers to flies with at least 1 ectopic scutellar bristle
\textsuperscript{d} refers to flies with a shortened fifth vein

116
Table 3-3. *tan* and *Ax* interact genetically in interocell bristle specification.

<table>
<thead>
<tr>
<th>Mutant allele</th>
<th>CyO/+: <em>tan</em>/tan</th>
<th><em>Ax</em>+/+: <em>tan</em>/tan</th>
<th>CyO/+, P[<em>tan</em>]/: <em>tan</em>/tan</th>
<th><em>Ax</em>+/+, P[<em>tan</em>]/: <em>tan</em>/tan</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>Ax</em>(1^3)</td>
<td>26.0% (292)</td>
<td>4.4% (475)</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td><em>Ax</em>(1^5)</td>
<td>25.9% (255)</td>
<td>7.0% (440)</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td><em>Ax</em>(1^7)</td>
<td>9.3% (400)</td>
<td>6.0% (465)</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td><em>Ax</em>(1^9)</td>
<td>18.8% (256)</td>
<td>6.3% (286)</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Df(2R)(rix)</td>
<td>1.7% (301)</td>
<td>15.0% (381)</td>
<td>0% (376)</td>
<td>1.5% (464)</td>
</tr>
<tr>
<td>Df(2R)(rix^*)</td>
<td>4.6% (241)</td>
<td>16.9% (313)</td>
<td>-</td>
<td>-</td>
</tr>
</tbody>
</table>

Crossoes were *Ax*+/CyO: *tan*/tan\(^1\)/tan\(^2\) females to *tan*/tan\(^2\) males, except * which were *Ax*/CyO: *tan*/tan\(^1\)/tan\(^2\) females to *tan*/tan\(^1\) males. *Ax*+/+: *tan*/tan and CyO/+: *tan*/tan siblings were then compared (columns 2 and 3). Df(2R)\(rix\)/CyO: *tan*/tan\(^2\) females were also crossed to P[*tan*]/P[*tan*]: *tan*/tan\(^2\) males to rescue the interocell bristle phenotype associated with the *tan* mutant (columns 4 and 5). Values indicate the percentage of flies with four or fewer interocell bristles, with the number of flies counted in brackets. All crossoes were repeated twice, except for those crossoes involving the *tan*\(^1\) allele, and the data pooled. Results were verified for significance by \(\chi^2\) analysis (\(\alpha = 0.05\)).
Table 3-4. Genetic interaction between *tan* and N.

<table>
<thead>
<tr>
<th>Mutant allele</th>
<th>NI+/tan&lt;sup&gt;+/+&lt;/sup&gt;</th>
<th>NI+: P[tan]&lt;sup&gt;+&lt;/sup&gt;/tan&lt;sup&gt;+/+&lt;/sup&gt;</th>
</tr>
</thead>
<tbody>
<tr>
<td>N&lt;sup&gt;a&lt;/sup&gt;</td>
<td>74.7% (75)</td>
<td>16.3% (104)</td>
</tr>
<tr>
<td>N&lt;sup&gt;2421&lt;/sup&gt;</td>
<td>26.7% (90)</td>
<td>4.1% (73)</td>
</tr>
</tbody>
</table>

The percentage of flies with 4 or fewer interocellar bristles is shown, with the number of flies counted in brackets. The loss of function allele N<sup>a</sup> and hypomorphic allele N<sup>2421</sup> were crossed to either *tan*<sup>2</sup> homozygotes or P[tan]:*tan*<sup>2</sup> homozygotes. Both N alleles show a strong genetic interaction when N and *tan*<sup>2</sup> are present heterozygously (column 2). *tan*<sup>2</sup> heterozygotes do not display the interocellar bristle phenotype when heterozygous (see Tables 3-2 and 3-3). This interaction is strongly rescued by the addition of the P[tan] construct (column 3).
Figure 3-1. GST pull-down experiment. Labeled TAN was incubated with GST or GST-ASX and the unbound (5% of sample) and bound (100% of sample) fractions were analyzed by SDS-PAGE. Unbound (Un) (Lanes 1 and 3) and Bound (B) (Lanes 2 and 4) fractions of TAN from GST (Lanes 1 and 2) and GST-ASX (Lanes 3 and 4) samples are shown. The two faster migrating bands most likely result from alternate translational start sites. The interaction of the fastest migrating band with ASX appeared to be weaker than full-length protein (compare right two lanes), and may uncover part of an ASX interaction domain in TAN.
Figure 3-2. Western blots. (A) Third instar larval extracts from HS-treated samples incubated with anti-TAN antibodies (first panel). A band of approximately 55kDa (arrow) is detected in all three lanes and is more abundant in HS-tan and HS-p53-tan lanes (WT: wild-type). The p53 tag adds 10 amino acids to the protein, resulting in a slower migrating band. The blot was overexposed to demonstrate equal loading (compare cross reacting bands indicated by arrowhead). A 55kDa protein is also detected by monoclonal p53 antibodies in HS-p53-tan samples, but not in HS-WT samples (second panel). (B) Affinity purified anti-TAN antibodies detect a 55kDa protein in Kc cell nuclear (Nu) extracts; cytoplasmic (Cy) fraction.
Figure 3-3. TAN binding to polytene chromosomes. (A) Affinity purified anti-TAN antibodies detect endogenous TAN at euchromatic sites on polytene chromosomes from salivary glands of third instar larvae. TAN binding sites are stained brown and the chromosome is counterstained in blue. Although numerous sites mapped show strong TAN binding (arrowheads), no staining is detected at the site of the BX-C (89E) (B). Similar results were obtained for the ANT-C.
**Figure 3-4. In vivo detection of TAN.** (A,B) Embryos from stage 16-17 stained with affinity purified TAN antibodies. (A) No signal is detected in a homozygous tan<sup>2</sup> mutant. (B) Staining is seen in the somatic and visceral (arrow) mesoderm, consistent with tan expression (see Figure 2-8E). (C-E) Confocal images of ectopic TAN localization. (C) Wing imaginal disc nuclei stained with propidium iodide (red). (D) TAN expression from a UAS-tan construct driven by patched-GAL4 (green). (E) Merged image reveals staining predominantly in the cytoplasm of the cells (compare arrow in panels). patched is expressed in a strong stripe along the A-P axis of the disc and no TAN staining is detected in non-patched expressing cells.
Figure 3-5. P-element mutagenesis of tan. (A) Schematic drawing of tan genomic region and surrounding Eco R1 (E) sites. Directed arrows indicate direction of transcription from an endopeptidase gene (gray box) located upstream of tan (white box). The 12kb P-element that inserted into the 5' region of tan is shown above (not to scale). The direction of transcription from the lacZ reporter gene within the P-element is indicated by the directed arrow. G-1, G-2, 13-1, and 13-2 show primers for PCR analysis and P1 and P2 indicate probes used for Southern analysis. The box below the P1/P2 probes indicates the extent of the deletion in tan\(^2\) (see below). (B-D) Genomic Southern blots. (B) The starting P-element line 0545-01/+ (Lane 1) and the tan\(^1\)/+ line (Lane 2) were digested with Eco R1 and probed with P1 and P2 probes. One ~8kb E\(^1\)-E\(^3\) band (Lane 1) is shifted upwards in the tan\(^1\) chromosome (Lane 2) due to the E\(^5\) site within the P-element. The 3kb E\(^2\)-E\(^3\) band is not affected by the insert but a new ~5kb band (star on right side) is created by E\(^5\) and the E\(^5\) site within the P-element. The E\(^2\)-E\(^5\) band contains only 200bp of tan sequence, resulting in a weak signal. The lower arrow on the right side shows one background band seen under high stringency washes using only the P1 probe (see text). (C) Southern blot with lacZ probe. Eco R1 digested DNA from 0545-01/+ (Lane 1), tan\(^1\)/+ (Lane 2) and tan\(^2\)/tan\(^2\) (Lane 3) flies probed with a lacZ probe reveal a new band in the tan\(^1\)/+ line (the original P-element is still present). This band is completely overlapping with the ~5kb E\(^2\)-E\(^3\) band detected in Lane 2 of panel B, as expected. tan\(^2\)/tan\(^2\) flies have lost the P-element inserted at the tan locus and the P-element has not reinserted in the genome. (D) Blot of 0545-01/+ (Lane 1) and tan\(^2\)/tan\(^2\) (Lane 2) using only the P1 probe. The E\(^1\)-E\(^3\) band detected in Lane 1 is weak because of the limited overlap between P1 and this fragment (~200bp). The deletion in tan\(^2\)/tan\(^2\) flies (Lane 2) created by the jump has deleted the E\(^2\) site leaving a weak ~8.3kb E\(^1\)-E\(^3\) band. The arrow denotes the background band also detected in panel B. (E) PCR analysis confirms the deletion of 1.4kb of tan DNA. The G-1/G-2 primers amplified a 2kb fragment from 0545-01/+ genomic DNA (Lane 1) but a 600bp fragment in tan\(^2\)/tan\(^2\) (Lane 3), while the 13-1/13-2 primers, located within the tan coding sequence, amplified a 200bp fragment in 0545-01/+ (Lane 2) but no product in tan\(^2\)/tan\(^2\) (Lane 4). The 600bp band (Lane 3) was purified from the gel and sequenced.
Figure 3-6. Adult defects in tan\(^2\) homozygotes. (A-D) Scanning electron micrographs with corresponding high magnification views. (A,C) Wild-type eye. (B,D) tan\(^2\) mutant eye. Notice the rough appearance due to disruption of ommatidial spacing and the misplaced, duplicated or missing bristles. Both groups of flies were reared at 30°C. (E) Wild-type dorsal head with one of the three ocelli indicated by an asterisk and interocellar bristles indicated by arrow (anterior is up). The larger ocellar bristles (top) and post- vertical bristles (bottom) are also visible. (F) A tan\(^2\) homozygous mutant with only three interocellar bristles (arrow). (G) The number of interocellar bristles was counted in 200 tan\(^2\)/tan\(^2\) and 200 P[tan]/P[tan];tan\(^2\)/tan\(^2\) rescued flies and graphed. tan\(^2\) mutants average approximately half the number of interocellar bristles. (H) Scutellum with a duplicated machrochaetae (arrow). (I) Male sternite with an ectopic bristle in the 6\(^{th}\) abdominal (A6) segment (arrow). (J) tan\(^2\) mutant showing a fifth vein which does not reach the margin (arrow).
Figure 3-7. Ectopic and over expression of *tan* disrupts sensory lineages. (A) Post-vertical bristles (arrows) are frequently deleted in Line A flies (4 copies of *tan*) or flies from a HS-*tan* line grown at 30°C. (B,C) Line B flies (4 copies of *tan*). Mutants showing either smaller post-vertical bristles (B) (compare left and right bristles indicated by arrows), or deletions of the bristle but not the socket (C). (D) A UAS-*p53-tan* line driven by *patched*--GAL4 has all bristles in the ocellar region missing, while the ocelli and cuticle are unaffected. (E) Ectopic expression of *tan* using a HS promoter resulted in the appearance of usually 3 or more hairs in male A6 sternites (denoted by asterisks). (F) A Line A fly (4 copies of *tan*) with ectopic vein material (indicated by arrows).
Figure 3-8. Ectopic expression of *tan* in the wing margin. (A) Wild-type wing. (B) UAS-*tan* expression driven by a margin specific driver (C96-GAL4) results in severe notching of the wing margins.
CHAPTER 4

Summary and Future Directions
Summary

In this thesis, I have described the characterization of the novel gene *tantalus* and outlined experiments which have begun to address its role during *Drosophila* development. Although the current understanding of TAN’s in vivo function is limited, enough details exist from my analysis of the *tantalus* gene to put forth a simple model amenable to further experimentation (Figure 4-1). Because ectopic TAN produced by the GAL4/UAS and heat shock systems is found in both the cytoplasm and nucleus, it is reasonable to assume that endogenous TAN will also be found localized to both compartments. Based on the genetic interaction between *tan* and *Notch* (*N*), and the similarity of *N* and *tan* phenotypes, it is possible that translocation of TAN to the nucleus could occur through *N* signaling, or possibly through another signal transduction pathway. One possibility is that translocation by these signaling pathways could require phosphorylation of TAN, since TAN contains many canonical phosphorylation sites. A role for signal transduction has recently been proposed for the *trxG* member *trithorax*, based on extensive phenotypic analysis of *trx* mutants (Breen, 1999) and the finding that cell maturation and differentiation are promoted by dephosphorylation of HRX, the human homologue of *trithorax* (Cui et al., 1998; De Vivo et al., 1998).

Once translocated to the nucleus, the data support a model in which TAN associates with the DNA, possibly through direct binding, and interacts with ASX. It is currently impossible though to predict which event might occur first. The specific binding of TAN and ASX to euchromatic sites on polytene chromosomes suggests that these factors play a direct role in target gene regulation. However, the dual role of ASX in both
PcG and trxG activities does not allow one to address whether TAN is acting in a PcG or trxG fashion.

**Future Directions: Testing the model**

The model proposed offers several testable predictions, which I discuss below. Determining whether TAN cycles between the nucleus and cytoplasm simply requires close examination of endogenous TAN expression, but could be supplemented by studies using the GAL4/UAS system. When *tan* is over-expressed in the presumptive notum using the *patched*-GAL4 driver, all machrochaetae are deleted. The p53-tagged version of TAN would allow one to easily follow protein distribution and determine when, and if, TAN becomes localized to the nucleus during bristle development. Additionally, the A101 enhancer trap line, which marks sensory organ precursor cells and their descendants (Bellen et al., 1989; Huang et al., 1991), would be useful in following TAN localization during bristle development. It would be quite exciting if TAN nuclear localization was coupled to sensory lineage specification.

Evidence presented in this thesis suggests that the N pathway may be involved in TAN function: interocellar bristle loss is dramatically enhanced in *tan/N* transheterozygous mutants when compared to *tan* heterozygous mutants alone (see Table 3-4). The presence of temperature-sensitive and inducible dominant forms of N could be used to determine whether the N pathway affects TAN localization and/or activity in SOP cells and their descendants.

Several groups have reported the dissociation of PcG or trxG members from polytene chromosomes in the absence of other members (Carrington and Jones, 1996;
Kuzin et al., 1994; Rastelli et al., 1993). For example, *Enhancer of zeste* mutations cause a decrease in the binding of both the *Suppressor2 of zeste* and *Posterior sex combs* proteins to polytene chromosomes (Rastelli et al., 1993). The fact that homozygous *tan* mutants are viable would make it simple to test whether *tan* loss-of-function mutations affect ASX binding to polytene chromosomes. If the DNA binding ability of TAN is used to stabilize/recruit an ASX complex then one might expect to see a decrease in ASX binding to at least some of the sites bound by both proteins on polytene chromosomes.

As mentioned, PcG/trxG members are found as protein complexes, and experiments such as the Far Western assay or co-immunoprecipitations could be used to address whether TAN interacts with other members of the PcG/trxG. Indeed, an interaction between TAN and Polycomb (PC) has been observed in the yeast two-hybrid system and by GST pull-down assays (Kyba and Brock, unpublished observations). This finding is rather interesting considering TAN and PC also bind to a large number of overlapping sites on polytene chromosomes. Several of these overlapping sites are not bound by ASX, suggesting that TAN could have functions that specifically require an interaction with Polycomb.

The early identification of HOM-C genes as targets of the PcG/trxG has been a boon for understanding many questions surrounding the function of PcG/trxG genes, and identification of TAN targets would greatly aid the understanding of its function. The bristle phenotypes associated with over- and under-expression of *tan*, combined with the mapped TAN binding sites on polytene chromosomes, allow one to postulate potential TAN target genes. By identifying potential bristle-regulating genes that map to the 66
sites of TAN binding on polytene chromosomes, the expression of these genes could be followed in *tan* mutants to determine if TAN regulates their expression. Unfortunately, the low penetrance/expressivity of the associated *tan* phenotypes may hinder the success of this approach.

One manner by which to overcome this hindrance is to combine a dominant modifier of the *tan* bristle phenotype, like *Notch* (see Table 3-4), with the *tan*<sup>2</sup> mutant to increase the penetrance of the phenotype. Another approach would be to study the phenotypes associated with *tan* over-expression. The penetrance of the loss of bristle phenotype in the notum approaches 100% when *tan* expression is driven by *patched*-GAL4. However, a potential drawback to this approach is that over-expression of *tan* may have slightly different developmental consequences than under-expression which, although bristle specific, would not necessarily reflect the role of endogenous TAN.

Finally, it is necessary to look more broadly at the role of TAN in sensory lineage specification, and such experiments could easily be combined with attempts to identify potential target genes. There are numerous bristle-specific genes and markers (see Kavalier et al., 1999) that may not be directly regulated by *tan*, but would still serve as useful molecular markers for SOP cells and their differentiated progeny. These markers could be used to determine the stage at which bristle differentiation is disrupted in *tan* mutants.

*Redundancy in Tantalus function*

Strong evidence for a tissue-specific role for TAN has been presented here. However, the incomplete penetrance of *tan* phenotypes is consistent with some
redundancy in *tan* function. Additionally, results of the Southern analysis leave open the possibility that a homologue to *tan* exists. One main goal of future studies will be to determine if redundancy plays a role in the penetrance/expressivity and tissue-specific activities observed for TAN.

The recent sequencing of the *Drosophila* genome (Adams et al., 2000; Rubin et al., 2000) may provide some perspective for this question. Approximately 13,600 *Drosophila* genes are predicted from the sequence analysis, representing ~ 8,000 distinct gene families. The remaining ~ 5,500 genes, or roughly 40%, represent duplicated genes. It is important to note, however, that the 5,500 genes do not represent separate gene families; for example, there are approximately 100 homeobox containing genes included within the 5,500 genes. The large number of duplicated genes may partly explain the fact that less than one-third of *Drosophila* genes give rise to obvious phenotypes when mutated (Ashburner and et al., 1999; Miklos and Rubin, 1996). Therefore, it would appear that *tan*, with its subtle defects, most likely represents the norm. Once the remaining gaps in the *Drosophila* sequence are filled, it should become clear as to whether or not a second gene with homology to *tan* exists.

Screens for dominant suppressors or enhancers of *tan* phenotypes are one way to search for redundant factors, as well as genes that may assist in TAN function. Such a procedure, based on results in Table 3-4 with the *N* gene, should be capable of identifying potential modifiers of *tan* function. Also, other identified PcG/trxG members should be analyzed to determine which of them, like *Asx*, can modify *tan* mutant phenotypes. As TAN binding to polytene chromosomes overlaps the binding sites of many PcG/trxG
members, it seems likely that at least some of these genes will also be involved in sensory lineage specification.

*Homeodomain binding of Tantalus*

Is there relevance to the finding that TAN interacts with the HD of FTZ? Although this interaction may be significant in vivo, it is also possible that this interaction is not specific to the HD of FTZ. The *Drosophila* genome project has identified ~100 HD-containing proteins (Rubin et al., 2000) and it will be important to address the specificity of the TAN/HD interaction. There are several HD-containing proteins, like D-PAX2 and Prospero, which are involved in the proper differentiation of sensory bristles (Kavaler et al., 1999; Reddy and Rodrigues, 1999) and it is possible that TAN could interact with these, or with similar proteins, during bristle differentiation.

In particular, study of *D-Pax2* has uncovered some interesting similarities to *tan* function (Kavaler et al., 1999). N signaling is required during development of sensory organ precursor (SOP) cells, as well as in the daughter cells of the SOPs. Within these daughter cells N signaling regulates the expression of *D-Pax2*. *D-Pax2* mutant alleles cause shortened/stunted bristles as well as empty sockets (Kavaler et al., 1999), and over-expression of D-PAX2 leads to the formation of ectopic bristle structures. Interestingly, *D-Pax2* mutants also have a rough eye appearance (Fu and Noll, 1997). The similarities of the *D-Pax2* phenotypes to the phenotypes seen when *tan* is over- or under-expressed make D-PAX2 an ideal candidate to test for a potential protein-protein interaction with TAN.
Conclusions

This project has illustrated some of the problems that will be encountered during the deciphering of the *Drosophila* genome. As mentioned, roughly two-thirds of *Drosophila* genes may not have easily identifiable phenotypes (Ashburner et al., 1999; Miklos and Rubin, 1996) and many of these may not possess sequence homology to currently known genes (Rubin et al., 2000). Both facts make study of new genes in *Drosophila* a slow and arduous process, and suggest that future studies will have to exert great effort to find detectable and highly penetrant phenotypes. As demonstrated here, over- and ectopic expression studies combined with genetic modifiers of subtle phenotypes (like the effect of *N* on *tan*) may be excellent ways to address these difficulties.
Figure 4-1. **Model of TAN function.** Cytoplasmic TAN translocates to the nucleus upon signaling by Notch or other, unidentified (?), pathways. This translocation step could involve phosphorylation (P). Once translocated to the nucleus TAN binds DNA and this step could require other cofactors (CoF). TAN then becomes associated with an ASX complex composed of either PcG or trxG members to maintain transcriptional states of target genes.
REFERENCES


Bejsovec, A. and Martinez-Arias, A. (1991) Roles of *wingless* in patterning the larval


899-913.


Chan, S.K., Jaffe, L., Capovilla, M., Botas, J. and Mann, R.S. (1994b) The DNA binding specificity of Ultrabithorax is modulated by cooperative interactions with extradenticle, another homeoprotein. *Cell*, 78, 603-615.


Chan, S.-K. and Mann, R.S. (1996) A structural model for a homeotic protein-


Deák, P. et al. (1997) P-element insertion alleles of essential genes on the third
chromosome of *Drosophila melanogaster*; correlation of physical and cytogenetic maps in chromosomal region 86E-87F. *Genetics, 147* 1697-1722.


Dingwall, A.K., Beek, S.J., McCallum, C.M., Tamkun, J.W., Kalpana, G.V., Goff, S.P.


homeotic locus Antennapedia in Drosophila. EMBO J., 2, 2027-2036.


bHLH activators delimit threshold responses to the *dorsal* gradient morphogen.

*Cell,* **72**, 741-752.


Karr, T.L. and Kornberg, T.B. (1989) *fushi tarzu* protein expression in the cellular
blastoderm of *Drosophila* detected using a novel imaging technique.

*Development*, 105, 95-103.


Kuzin, B., Tillib, S., Sedkov, Y., Mizrokhi, L. and Mazo, A. (1994) The *Drosophila trithorax* gene encodes a chromosomal protein and directly regulates the region-


*even-skipped*: A second pair-rule gene of *Drosophila* containing a homeo box. 

*Cell*, 47, 721-734.


protein α2 is required for function with a second homeo domain protein. *Genes 

Dev.*, 7, 1862-1870.


Dev.*, 8, 423-429.

Mann, R.S. and Chan, S.-K. (1996) Extra specificity from *extradenticle*: the partnership 

between HOX and PBX/EXD homeodomain proteins. *Trends Genet.*. 12, 258- 

262.

Mann, R.S. and Hogness, D.S. (1990) Functional dissection of Ultrabithorax proteins in 


the definition and maintenance of cell states in the *Drosophila* embryo.
Development, 103, 157-170.


Mlodzik, M., Gibson, G. and Gehring, W.J. (1990) Effects of ectopic expression of


Orlando, V., Jane, E.P., Chinwalla, V., Harte, P.J. and Paro, R. (1998) Binding of
trithorax and Polycomb proteins to the bithorax complex: dynamic changes during early Drosophila embryogenesis. EMBO J., 17, 5141-5150.


Rusch, J. and Levine, M. (1994) Regulation of the *dorsal* morphogen by the *Toll* and


Schüpbach, T. (1987) Germ line and soma cooperate during oogenesis to establish the


Sinclair, D.A., Milne, T.A., Hodgson, J.W., Shellard, J., Salinas, C.A., Kyba, M.,


Tie, F., Furuyama, T. and Harte, P.J. (1998) The Drosophila Polycomb Group proteins ESC and E(Z) bind directly to each other and co-localize at multiple chromosomal


