Order Number 9510670

Transcriptional regulation of the *Drosophila fushi tarazu* gene by its proximal enhancer

Han, Wei, Ph.D.

City University of New York, 1994

U·M·I
300 N. Zeeb Rd.
Ann Arbor, MI 48106

TRANSCRIPTIONAL REGULATION OF THE *DROSOPHILA* *fushi tarazu* GENE BY ITS PROXIMAL ENHANCER

by

WEI HAN

A dissertation submitted to the Graduate Faculty in Biomedical Sciences, in partial fulfillment of the requirements for the degree of Doctor of Philosophy,

The City University of New York

1994

This manuscript has been read and accepted for the Graduate Faculty in Biomedical Sciences in satisfaction of the dissertation requirement for the degree of Doctor of Philosophy.

July 5, 1994
Date

Robert A. Lazzarini, Ph.D.
Chair of Examining Committee

July 5, 1994
Date

Terry A. Krulwich, Ph.D.
Executive Officer

James Bieker, Ph.D.

Robert Krauss, Ph.D.

Leslie Pick, Ph.D.

Carl Wu, Ph.D.

Supervisory Committee

The City University of New York

ABSTRACT

TRANSCRIPTIONAL REGULATION OF THE *DROSOPHILA fushi tarazu* GENE BY ITS PROXIMAL ENHANCER

by

Wei Han

Advisor: Professor Leslie Pick

The expression of the *Drosophila* segmentation gene *fushi tarazu* (*ftz*) is controlled at the level of transcription. The *ftz* proximal enhancer, a 794-bp fragment located at ~3.4 kb upstream of the transcription start site, directs reporter gene expression in seven *ftz*-like stripes in transgenic embryos. We have taken a biochemical approach to unravel the molecular basis for the enhancer function. Nine footprint sites were identified using DNaseI footprinting and methylation interference experiments with staged *Drosophila* embryo nuclear extracts (NE). A 323-bp fragment containing all nine sites directed *lacZ* fusion gene expression in seven *ftz*-like stripes. Site-specific substitution mutations of sites 6, 8 and 9 abolished reporter gene expression. Twelve different sequence-specific DNA-binding complexes (fEBC1-12) interacted with eight of the nine sites. Twelve proteins were purified from NE with site 4 and 6 affinity chromatography. Reliable peptide sequences were obtained for P1 to P8. The cDNAs encoding P1 were isolated from an embryo cDNA library.

P2 is a known transcription factor, Adf1. Both *Escherichia coli* expressed P1 and Adf1 bound to multiple footprint sites specifically. Adf1 and FTZ-F1 are involved in the formation fEBC8 with footprint sites 6, 8, and 9, as suggested in gel retardation assays with antibodies. Both P1 and Adf1 have potential transcription activation function when tested in a yeast system. P5 is an another known transcription factor, TTK. It binds to five of the nine footprint sites. We propose that *ftz* expression is initiated and maintained by transcriptional activators identified in these studies, including FTZ-F1, P1, and Adf1, along with FTZ protein. These proteins interact with multiple sites in the *ftz* proximal enhancer. Potential repressors, including TTK, compete with the activators for binding to the same DNA sequences to actively repress *ftz* autoregulation and turn off *ftz* expression.

ACKNOWLEDGMENTS

I would like to thank all those who made my studies at Mount Sinai Graduate School possible, who helped me to make my studies productive, and who helped me enjoy my New York years. Of all these people, Leslie Pick, my thesis advisor, deserves special gratitude. I thank you, Leslie, for your continuous support, your advice, your patience, and your critical comments on the manuscript. My stay in your lab was a great learning experience. I also thank the members of my thesis committee, Robert Lazzarini and James Bieker for their advice and criticisms. I thank all the members of the Pick lab, past and present, for a stimulating and friendly atmosphere. My wife, Yan Yu, has been very supportive through all my years in the graduate school. Thank you, Yan.

I would like to dedicate this thesis to my parents.

CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1. INTRODUCTION

1. General introduction

The *Drosophila* embryo initially develops as a syncitium characterized by rapid synchronously dividing nuclei. After thirteen nuclear divisions, the totipotent nuclei located in the central portion of the fertilized egg have migrated to the cortical cytoplasm. By the end of the fourteenth cleavage cycle, the fate of each nucleus is determined depending upon its position in the periphery of the embryo (Simcox and Sams 1983; Technau 1987). The syncitial blastoderm shifts to cellular blastoderm when the determined nuclei are surrounded by the invagination of membranes forming an ~5000 cell monolayer. Gastrulation follows as soon as cellularization is complete. During this process, three germ layers (endoderm, mesoderm, ectoderm) form through coordinated cell movements. As gastrulation nears completion, the embryo is subdivided into segments or parasegments along the anterior-posterior axis, marked in the surface of the embryo by a series of indentations and bulges (Martinez-Arias and Lawrence 1985; Lawrence et al. 1987).

Nusslein-Volhard and co-workers set up a mutagenesis screen to look for genes regulating embryonic development. The approach was to eliminate the function of a gene required for embryonic development by mutation. The mutant phenotype is a specific abnormal pattern formed by the embryo. The deviation from normal development reflects the function of

the respective gene. The maternal coordinate genes and zygotic segmentation genes were identified. They were classified according to whether their mutant phenotypes depend on the genotypes of the mother or on that of the progeny (Nulsslein-Volhard and Wieschaus 1980; Nusslein-Volhard et al. 1984; Akam 1987; Johnston and Nusslein-Volhard 1992). The identification of the maternal coordinate genes and zygotic segmentation genes began to reveal the mechanisms underlying the elaboration of positional information in early *Drosophila* embryogenesis. The positional information refers to the location of an individual cell in an ellipsoidal monolayer of ~5000 cells in the cellular blastoderm stage. The maternal coordinate gene products lay down the initial positional information along the anterior-posterior and dorsal-ventral axes of the embryo. The dorsal-ventral axis is established by a single system of interacting gene products; among them, the product of *dorsal* (*dl*) gene is the primary determinant (Anderson 1987; Johnston and Nusslein-Volhard 1992). The anterior-posterior axis constitutes three discrete systems (Nusslein-Volhard et al. 1987). Each system is determined by a group of gene products. Briefly, *bicoid* and *hunchback* function in the anterior; *nanos*, *knirps* are responsible for posterior cues; *torso*, *huckebein*, and *tailess* specify the terminal regions of the embryo.

The three classes of segmentation genes were classified according to the extent of segment disruptions in mutant embryos (Nulsslein-Volhard and Wieschaus 1980; Nusslein-

Volhard et al. 1984; Akam 1987). The gap genes (e.g. *hunchback*) are expressed in broad domains that cover the primordia of several segments. These segments are missing in gap gene mutants. The pair-rule genes [e.g. *fushi tarazu, (ftz)*] define the primordia of alternate segments and the segment-polarity genes (e.g. *engrailed*) specify the primordia of domains within each segment.

The molecular cloning of the maternal genes and segmentation genes provided molecular probes. The probes were used to examine the expression patterns of these genes. In general, the expression pattern of a given gene is consistent with the pattern of defects in the mutant embryos for that gene. The availability of the probes also made it possible to study the relationships of genes by visualizing one gene expression pattern in various genetic backgrounds (Akam 1987; Nusslein-Volhard 1991). These epistatic studies revealed that the crude positional information supplied by the maternal genes is interpreted and progressively refined through the zygotic segmentation genes. The anterior-posterior maternal genes function to define gap gene expression patterns; the gap genes instruct the spatially restricted expression patterns of the pair-rule genes; specific combinations of pair-rule genes determine the expression of the segment-polarity genes. These studies suggest that those genes act in a hierarchical fashion to lay down the positional cues in early embryogenesis. However, genetic studies can never tell whether the relationship between two genes is direct or

indirect. Many steps may exist in between any two nearby levels of the hierarchy (examples given below).

2. The *fushi tarazu* (*ftz*) gene

a) Characterization of the *ftz* gene

One of the well studied segmentation genes is the pair-rule gene *fushi tarazu* (*ftz*). The Japanese name *fushi tarazu* meaning "not enough segments" describes the homozygous *ftz* mutant embryos that lack alternate body segments (Wakimoto and Kaufmann 1981). The posterior part of one segment and the anterior part of the next segment are missing, and the remaining segmental portions are fused (Nulsslein-Volhard and Wieschaus 1980). The deleted patterns correspond, in terms of parasegments (Martinez-Arias and Lawrence 1985), to the even numbered parasegments. The *ftz* gene was isolated by its homology to the homeobox of *Antennapedia* (*Antp*) (Kuroiwa et al. 1984). The *ftz* expression pattern has been studied by in situ hybridization to *ftz* mRNA, and by histochemical staining of *ftz* protein using antibodies (Hafen et al. 1984; Weir and Kornberg 1985; Carroll and Scott 1986; Lawrence et al. 1987; Krause et al. 1988). The *ftz* mRNA is first detected at low levels in the whole embryo and accumulates in seven evenly spaced bands corresponding to the even numbered parasegments. The seven *ftz* stripes persist throughout the cellular blastoderm, gastrulation and disappear before the germband is fully extended. The *ftz* gene is later expressed in a subset of cells in developing nervous system. Thus, the *ftz* gene is

expressed early in the primordia of the even-numbered parasegments which are missing in *ftz* mutant embryos (Hafen et al. 1984; Carroll and Scott 1985). Embryos also die when *ftz* is ectopically expressed (Struhl 1985). The correspondence of domains of expression and function strongly indicates that the correct spatial and temporal expression of *ftz* is necessary for normal development of embryos.

b) Identification of the cis-acting elements and trans-acting factors of the *ftz* gene

Molecular and biochemical approaches have begun to elucidate the molecular mechanisms regulating gene expression. Molecular cloning and sequencing of the homeotic genes (e.g. *Antp*) and some of the segmentation genes (e.g. *ftz*) revealed a common 180 bp sequence - the homeobox (McGinnis et al. 1984; Scott and Weiner 1984). The homeodomain encoded by the homeobox constitutes a DNA binding domain (Desplan and Theis 1985; Hoey and Levine 1988; Scott et al. 1989). That homeobox containing genes encode transcriptional regulators has been further suggested by co-transfection assays using *Drosophila* cell culture (Jaynes and O'Farrell 1988; Han and Levine 1989). In addition, many other segmentation genes (e.g. gap genes) contain zinc-finger DNA-binding motifs (Biggin and Tjian 1989); they also function as transcription factors. These observations along with genetic studies support the idea that the segmentation genes regulate each other at the level of transcription initiation

(Ingham 1988; Levine and Hoey 1988; Biggin and Tjian 1989; Murre et al. 1989; Scott et al. 1989). Like mammalian transcription regulation of gene expression, transcription initiation of *Drosophila* gene is also a primary control step in gene expression (Mitchell 1989). The regulation of differential gene expression depends on the gene structure itself: the cis-regulatory elements including the promoter and enhancers. Transcription factors, which are the trans-acting factors, interact specifically with the cis-regulatory elements. It is generally believed that the unique organization of cis-elements and the availability of trans-factors specify the gene expression pattern. The organization of the promoter and enhancer are quite similar, each of them composed of many individual elements which interact with one or more transcription factor(s) (Dynan 1989).

(I) The cis-acting elements of the *ftz* gene

To understand the transcriptional regulation of *ftz* expression, the cis-regulatory elements were examined by P-element transformation with recombinant promoters directing expression of a reporter gene, *E. coli lac-Z* (Hiromi et al. 1985; Hiromi and Gehring 1987). As shown in Figure 1, a 5' regulatory region of ~6 kb was shown to contain three cis-regulatory elements: A "zebra element" directed *lac-Z* fusion gene expression in seven stripes primarily in the mesoderm; a "neurogenic element" directed gene expression in neural precursor cells and an "upstream element", which will be

discussed below. All three cis-regulatory elements are necessary to rescue *ftz* mutant embryos. The *ftz* upstream element has been defined as an enhancer. It stimulates *ftz* expression from its own promoter (zebra element) in a distance and orientation independent fashion. It also can direct *lac-Z* fusion gene expression in seven *ftz*-like stripes via a heterologous promoter (Hiromi and Gehring 1987). A deletion analysis divided the enhancer into a distal enhancer, containing a mesodermally active element, and proximal enhancer, containing both mesodermally and ectodermally active elements (Pick et al. 1990). The *ftz* proximal enhancer contains two regulatory units: Prox A, a mesodermally restricted element, and Prox B that, when combined with Prox A, directs fusion gene expression in both mesodermal and ectodermal primordia (Pick et al. 1990).

(II) The trans-acting factors regulating *ftz* expression

a) Genetic studies. Analysis of *ftz* expression in embryos with various genetic backgrounds has revealed three classes of trans-acting regulators for the *ftz* gene. First, five maternal genes (*vasa, valois, stauffen, tudor* and *oskar*), required for normal abdominal segmentation, may be involved in the regulation of *ftz* expression. Each of the five mutants affected the *ftz* pattern differently, with the common features of aberrations in the region between the second and seventh stripes (Carroll et al. 1986). *bicoid* product has been suggested to be a repressor, which suppresses the *ftz*

expression in the anterior region of embryos (Frohnhofer and Nusslein-Volhard 1987; Vavra and Carroll 1989). Second, four gap genes (*Kruppel*, *knirps*, *hunchback*, *giant*) were shown to affect the *ftz* expression pattern (Carroll and Scott 1986). Finally, three of the eight pair-rule genes (*hairy*, *runt*, *even-skipped*) regulate *ftz* expression by refining and maintaining the stripe pattern (Ingham and Gergen 1988). *hairy* product suppresses *ftz* expression in the odd-numbered parasegments (spaces of the *ftz* stripes); *runt* and *even-skipped* products positively regulate part of the seven *ftz* stripes (Carroll and Scott 1986; Ingham and Gergen 1988; Carroll and Vavra 1989). It appears that the regulation of *ftz* expression by the gap and pair-rule genes are mediated by the *ftz* zebra element (Hiromi and Gehring 1987). This was examined by analysis of the expression of *ftz zebra* element-*lacZ* fusion genes in various mutant embryos of gap and pair-rule genes. The other pair-rule genes, except *ftz* itself (discussed below), and segment-polarity genes so far tested show no effects on the *ftz* expression pattern.

The limitation of the genetic studies make it impossible to determine if any of these regulatory effects is direct. For example, in *knirps* mutant embryos, the seven *ftz* stripes appears as four strips, the first two and seventh looks normal, an abnormal broad band was detected that extends across the area where the third through sixth stripes normally form (Carroll and Scott 1986). This observation does not reveal the molecule basis the effects on the *ftz* pattern.

The *knirps* may regulate the *ftz* expression by turning on or off the other pair-rule genes or unknown genes. Thus, *ftz* may not be the direct target of *knirps* at all. The genetic studies only reveal the hierarchical relationship between two regulatory genes, that is, *ftz* is downstream of *knirps*. Is *knirps* the immediate upstream regulator of the *ftz* gene? Even with the identification of the *ftz* zebra element as the potential target for knirps in genetic studies (Hiromi and Gehring 1987), it would still have to be shown that the *knirps* product binds to the zebra element biochemically and activates the transcription of the *ftz* gene in a DNA-binding site dependent fashion. Therefore, a combination of both genetic and biochemical approaches are necessary to unravel the molecular basis of transcriptional regulation of *ftz* and other developmental genes in early embryogenesis.

b) Biochemical studies. The FTZ protein itself was the first suggested trans-regulator for which both genetic and biochemical evidence was provided. Genetic studies showed that the expression of *ftz* enhancer-*lacZ* fusion genes depends on the wild type *ftz* gene (Hiromi and Gehring 1987; Pick et al. 1990). Protein-DNA binding assays revealed that the homeodomain containing FTZ protein binds to DNA specifically (Gehring 1987; Scott et al. 1989). It can function as a transcription factor regulating gene expression in yeast and in *Drosophila* cultured cells (Jaynes and O'Farrell 1988; Fitzpatrick and Ingles 1989; Han and Levine 1989). The *ftz*

distal and proximal enhancers contain multiple binding sites for FTZ protein *in vitro* (Pick et al. 1990). These binding sites are thought to mediate direct FTZ homeodomain-DNA interaction *in vivo,* as suggested by experiments of Schier and Gehring (Schier and Gehring 1992). These results indicate that FTZ protein mediates autoregulation of *ftz* expression by interacting with the *ftz* enhancer.

tramtrack protein (TTK) was the second potential *ftz* regulator identified. Its cDNA was cloned by screening a *Drosophila* expression library with several oligonucleotides from the *ftz* distal enhancer (Harrison and Travers 1990), *ftz* zebra element (Brown et al. 1991) and *even-skipped* (*eve*) promoter (Read et al. 1990). Two forms of TTK resulting from differential splicing have been identified. One form is 69 kDa and the other is 88 kDa (Read and Manley 1992). The zinc-finger DNA binding motifs of the two forms are different. The 69 kDa TTK is thought to act as a transcriptional repressor since overexpression of the TTK protein severely suppressed *ftz* expression in transgenic *Drosophila* embryos (Read 1992; Brown and Wu 1993). Null mutants for *ttk* result in embryonic lethality, before dorsal closure (Xiong and Montell 1993). However, no misexpression of FTZ protein was detected in the null *ttk* mutants. The 69 kDa TTK appears to regulate other segmentation genes, such as *even skipped, odd skipped, hairy* and *runt* (Read 1992; Brown and Wu 1993). Both forms of TTK protein were shown to be involved in cell fate determination in the *Drosophila* compound eye (Xiong and Montell 1993).

The third potential *ftz* regulator is FTZ-F1. FTZ-F1 was purified by DNA affinity chromatography using a sequence from the *ftz* zebra element (Ueda et al. 1990). The deduced amino acid sequence revealed that it is a member of the hormone receptor superfamily (Lavorgna et al. 1991). Two DNA binding sites for FTZ-F1 were found in the ftz zebra element. Mutations of those two sites decreased expression of a reporter gene in transformed embryos. Therefore, FTZ-F1 could be a direct transcriptional activator of ftz expression. FTZ-F1β was cloned using low stringency hybridization with the DNA-binding domain of FTZ-F1 as probe. It is a member of the nuclear receptor family with homology to FTZ-F1 in its DNA-binding domain (Ohno and Petkovich 1992). It bound to the *ftz* zebra element in a similar manner as did FTZ-F1 (Ohno and Petkovich 1992), and it also interacts with the *Adh* distal enhancer (Ayer and Benyajati 1992).

3. Specific aims for the study of the *ftz* proximal enhancer

While genetic screens have led to the identification of a large number of developmentally regulated genes, the identification of potential *ftz* regulators by molecular and biochemical approaches, such as FTZ-F1, FTZ-F1β and TTK, revealed that some regulatory genes were missed by genetic screens. Rather, a combination of genetic, molecular biology and biochemistry approaches allows for the elucidation of mechanisms by which the *ftz* gene is regulated. These genetic and biochemical approaches are necessary to identify all the

regulatory genes and to understand the biochemical mechanisms involved in segmentation and embryonic development.

The *ftz* gene has been shown to contain three cis-acting regulatory elements: the zebra element, the neurogenic element, and the upstream element (Fig. 1) (Hiromi et al. 1985). The upstream elements contains two independent enhancers: the distal enhancer and the proximal enhancer (Pick et al. 1990). The zebra element and distal enhancer each direct reporter gene expression in the mesoderm only; the proximal enhancer leads to reporter gene expression in both ectoderm and mesoderm, which most closely resembles the expression pattern of the endogenous *ftz* gene. Striped expression directed by each of the enhancers is dependent upon the presence of wild-type *ftz* product in the embryo (Hiromi and Gehring 1987; Pick et al. 1990). Therefore, the FTZ protein is the only identified trans-acting regulator which might interact directly with the *ftz* proximal enhancer mediating autoregulation. Since the endogenous *ftz* is transcribed in both germlayers, it was proposed that each of the enhancers requires at least one other factor, in addition to the FTZ protein, to direct reporter gene expression in different germlayers (Pick et al. 1990). Given that the proximal enhancer appears to play key roles in the activation, autoregulation (maintenance) and repression of the *ftz* expression (see Discussion), it is more likely that multiple trans-acting regulators interact with the *ftz*

proximal enhancer, which determine the expression pattern of the *ftz* gene.

To understand the mechanisms by which the *ftz* proximal enhancer directs the *ftz* expression pattern, a biochemical approach was taken to identify transcription factors that interact directly with the *ftz* proximal enhancer. The results of these studies are presented in this thesis.

CHAPTER 2. MATERIALS AND METHODS

1. Preparation of *ftz* proximal enhancer fragments

The seven proximal enhancer fragments shown in Figure 1 were generated from plasmid XL-7 containing an XbaI fragment from positions 1502 to 2574 of the *ftz* proximal enhancer (Pick, unpublished) (Harrison and Travers 1988). The XbaI fragment was released from XL-7, gel-purified and cut with HincII. The digested fragments were inserted into the EcoRV site of a Bluescript KS+ vector to generate subclone C191 and into the EcoRV and XbaI sites of the vector to generate subclones C300 and C500. C300 was cut with EcoRV and XbaI; the XbaI site of the subclone was filled in with Klenow fragment and the two blunt ends were ligated generating subclone C129. The released EcoRV-XbaI fragment from C300 was gel purified and inserted into the EcoRV-XbaI site of a Bluescript KS+ vector to generate subclone C211. F1 and F2 were released from C211 by digestion with EcoRV, XbaI and BstEII. F3 was released from C129 by digestion with XbaI and Hind III. F4 was released from C191 by digestion with EcoRV and Hind III. F5, F6 and F7 were released from C500 by digestion with EcoRV, XbaI, BglII and AvaII; F5 is a EcoRV-AvaII fragment, F6 is a AvaII-BglII fragment, F7 is a BglII-XbaI fragment (Fig. 1). All of the fragments contain at least one 5'-protruding end. All procedures were carried out using standard techniques (Maniatis et al. 1982).

2. Gel retardation assays

The procedure was carried out essentially as described by Carthew et al. (1985). Fragments with 5'-protruding ends were labeled with α-^{32}P-dATP and Klenow fragment. 2-6 µg of the nuclear extract was incubated with 10 fmole of labeled fragment in a total reaction volume of 25 µl containing 0.1 M KCl, 25 mM Hepes, pH7.6, 0.5 mM DTT, 10% glycerol and 4-6 µg poly(dI-dC) (Pharmacia). Samples were kept on ice for 1 hr and analyzed by electrophoresis through 4% native polyacrylamide gels using 0.5 X TBE as running buffer. The gels were dried after running at 12 mA for 1 hr in the unlabeled room. The dried gels were exposed overnight for autoradiography at -70°C with an intensify screen.

The oligonucleotide gel retardation assay was done essentially as described for the DNA fragment gel retardation, except that 0.1 µg random single stranded oligonucleotide was included as non-specific competitor. In the oligonucleotide competition reactions, the unlabeled oligonucleotide competitors were added to the binding reactions which were initiated by addition of protein fraction.

3. DNaseI footprinting assay

The protocol used is modified from Ueda, et al. (Ueda et al. 1990). Subclones C211, C129 and C191 containing DNA fragments F2, F3 and F4 were digested with restriction enzymes generating 5'-protruding ends that were labeled with

α-^{32}P-dATP and Klenow fragment. One end labeled fragment was released from the vectors by digestion of the opposite ends of the fragments. The labeled fragments were gel-purified. Ten reactions were carried out as described in the gel retardation assay with labeled DNA fragments. After 1 hr incubation, 7 µl of 1 U/ml DNaseI (Boehringer), 10 µl solution containing 125 mM $MgCl_2$ and 62.5 mM $CaCl_2$ was added. The digestion proceeded for 1 min on ice, and was stopped by addition of 8 µl of 0.5 M EDTA. The samples were analyzed by electrophoresis through a 4% native polyacrylamide gel as described above. After the free and bound DNA on the gel were localized by autoradiography, they were recovered from the gel slices by electroelution. The DNA was purified by phenol/chloroform extraction followed by passage through Elutip-d minicolumns (Schleicher and Schuell Co.), and ethanol precipitation. Equal amounts of radioactivity from the bound and free DNA samples were electrophoresed through an 8.3 M urea, 8% polyacrylamide gel with guanine (G) cleavage reaction as marker (Maxam and Gilbert 1977). The gel was autoradiographed after drying.

4. <u>Methylation interference analysis</u>

The methylation interference assay was done essentially as described by Hendrickson and Schleif (1985). The DNA fragments F2, F3 and F4 were end-labeled as above. They were partially methylated with dimethylsulfate (Maxam and Gilbert 1980). The methylated fragments as probes were incubated with

the nuclear extracts as described in Method 2. After separation of free from bound probes on native gels, each of them was electroeluted from the gel slices. The DNA was purified as above and cleaved at the methylated guanine sites with piperidine treatment (Maxam and Gilbert 1980). The cleavage products were electrophoresed through an 8.3 M urea, 8% polyacrylamide gel. The gels were autoradiographed after they were dried.

5. Purification of potential *ftz* enhancer binding proteins

a) Preparation of DNA-affinity columns

The DNA-affinity columns were prepared as described by Kadonaga and Tjian (1986). Four complementary oligonucleotides (O2, O3, O4, O6,) were synthesized with the four footprint sites (site 2, 3, 4 and 6 in Fig. 7) in the center flanked by HindIII recognition sequences. 440 µg of each oligonucleotide strand was suspended in 100 µl TE (10 mM Tris-HCl, pH7.6, 1 mM EDTA). They were annealed by sequential incubation at 88°C for 2 min; 65°C for 10 min; 37°C for 10 min and 25°C for 5 min. The 5'-ends of the oligonucleotides were phosphorylated and partially labeled with polynucleotide kinase (150 U, Biolab) in the presence of 2.4 mM ATP and 5 µCi γ-^{32}P-ATP. The phosphorylated oligonucleotides were concatamerized using 2000 U of T4 DNA ligase (400 U/µl, Biolab) at 15°C overnight. The concatamerized oligonucleotides were coupled to CNBr-sepharose 4B beads (Pharmacia) (Briggs et al. 1986). 10 ml of swelled beads was

mixed with 4 ml partially labeled DNA in 10 mM potassium phosphate buffer, pH 8.0 (coupling buffer). The mixture was rotated at room temperature overnight, followed by sequential washing with 12 ml of coupling buffer, two times 100 ml dH_2O and 100 ml of 1 M ethanolamine-HCl, pH 8.0. The remaining CNBr active sites were blocked with 4 ml of 1 M ethanolamine-HCl, pH8.0, by rotating for 4 hr at room temperature. The DNA coupled beads were washed sequentially with 100 ml of 10 mM and 1 M potassium phosphate buffer, 100 ml of 1 M HCl and 100 ml of dH_2O. The final DNA affinity beads were stored in column storage buffer (0.3 M NaCl, 10 mM Tris-HCl, pH7.5, 1 mM EDTA, 0.02% NaN_3) at 4°C. The coupling efficiency was monitored by measuring the radioactivity before and after coupling. It was about 60%. The DNA affinity columns were used repeatedly. After each purification, they were regenerated by washing sequentially with 10 ml of 1 M NaCl HEG and 40 ml of 0.1 M HEG.

b) Purification procedures

The purification procedures were developed by pretest on a small scale. 100 mg of 0-12 hr nuclear extract prepared from the frozen embryos (for detail, see Result 7) was loaded into a 10 ml Heparin-sepharose column. The 0.4 and 1 M NaCl heparin fractions were precipitated with 40% ammonium sulfate. The pellets were further purified using O2-affinity chromatography for three-rounds. The purification was monitored by gel retardation assays with O2. The detail large

scale purification with O4 and O6 affinity chromatography are described in the Result 7.

6. Preparation of the purified proteins for peptide micro-sequencing

Purified proteins from the DNA-affinity chromatography were precipitated with 4-fold excess of acetone (pre-cooled to -20°C) for 18 hrs at -80°C. The proteins were separated by gel electrophoresis through a 10% SDS polyacrylamide gel (Laemmli 1970). The protein bands were revealed by staining the gel with 0.1% Coomassie Brilliant Blue R250 in water for 10 min, followed by destaining in water for 15 min. The stained bands were excised and loaded onto a 15% SDS polyacrylamide gel. 0.1 µg of V8 protease in 20 µl of 10 mM Tris-HCl, pH 8.0, was loaded on top of the gel slice. The gel was run for 1 hr at 50 V, and then at 150 V until the Bromophenol Blue reached the bottom of the gel. The digested peptides in the gel were transferred to a PVDF membrane (Millipore). The electoblotting was carried out in 1 L of transfer buffer (10 mM TrisBase, 60 mM glycine, 0.03% SDS, 30% methanol) for 2 hrs at 150 mA. The PVDF membrane was stained for 30 sec in 50 ml of staining buffer (0.1% Coomassie Brilliant Blue R250, 50% methanol, 10% acetic acid), followed by destaining with two changes of 50 ml of destaining buffer (50% methanol, 12% acetic acid) until background was gone (about 10 min). The stained PVDF filters

were sent to the Protein Core Facility at Mount Sinai for sequencing.

7. Renaturation of the purified proteins

The renaturation experiment was done essentially as described previously (Briggs et al. 1986; England et al. 1990). 0.2 ml of eluate from the second-round of DNA-affinity chromatography was precipitated with 0.8 ml unlabeled acetone. The proteins were separated by 10% SDS-PAGE. Prestained protein molecular weight markers were run on the gels. The gels were cut into 4 regions as indicated in Figures 16A-19A. Each gel slice was crushed with a Teflon pestle in an Eppendorf tube containing 0.25 ml elution buffer (0.1% SDS, 50 mM Tris-HCl, pH7.6, 0.1 mM EDTA, 5 mM DTT, 0.1 mg/ml BSA, 0.15 M NaCl), and incubated for one hour at room temperature (R/T). After the supernatant was recovered by a brief spin for 1 min in a Eppendorf tube centrifuge, the gel was washed with 0.1 ml elution buffer. 0.3 ml recovered supernatant was precipitated with 1.2 ml acetone. The pellet was resuspended in 10 µl of 6 M guanidine-HCl made in dilution buffer (25 mM Hepes, pH7.6, 0.1 M NaCl, 0.5 mM DTT, 10% glycerol, 0.1 mg/ml BSA) and incubated for 30 minutes at R/T. 0.5 ml dilution buffer was added and incubated for 50 minutes at R/T. The renatured protein was concentrated in a Centricon-30 filter (Amacon) from 0.5 ml to 50 µl. Finally, 50 µl renatured protein was brought to 100 µl by adding 50 µl

of dilution buffer. 8 µl was used in one reaction for the gel retardation assay.

8. Site-specific mutagenesis of the *ftz* proximal enhancer

Site-specific mutagenesis was done using the overlap extension method (Steffan et al. 1989). The principle is shown schematically in Figure 11. The polymerase chain reactions were carried out according to the standard protocol with the Taq DNA polymerase (Promega) (Innis and Gelfand 1990). Oligonucleotides a and b were used in PCR reaction 1 (Rx1); oligonucleotides d and c were used as primers in reaction 2 (Rx2). M indicates introduced mutant sites. Products from Rx1 and Rx2 were gel-purified and used in PCR reaction 3 (Rx3) with oligonucleotide a and d as primers. Rx3 was an overlap extension step generating a full size DNA template carrying the mutations as designed in oligonucleotide c and b. The DNA template used was the *ftz* proximal enhancer fragment covering positions 2168-2574 (DraI - XbaI fragment) subcloned in Bluescript KS+ vectors. The primers a and d indicated in Figure 11 correspond to the polylinkers of the KS vector with XbaI recognition sites at the 5'-end. They were synthesized according to the following sequences: The primer a is 5'-TCTAGAAACAGCTATGACCATG-3'(upper strand); primer b is 5'-TCTAGAGTAAAACGACGGCCAGT-3' (lower strand). Nine pairs of mutant oligonucleotides were synthesized according to the mutant sites shown in Figure 12. Each pair of mutant oligonucleotides was used as primers in

PCR reactions 1 and 2 (b and c primers), as shown in Figure 11. The products from PCR reaction 3 were cut with XbaI, and subcloned back into the KS vectors. These were then used as the templates for the next round of site-specific mutagenesis to generate multiple site-specific mutations. In this way, multiple fEBC binding sites were mutated (Fig. 13). The sequences of the mutant DNA constructs in the Bluescript KS+ vectors were confirmed by double stranded DNA sequencing using sequencing grade Taq DNA polymerase (Promega) (Sanger et al. 1977; Innis et al. 1988). Generally, one out of five clones for each mutant construct contained the right sequence.

9. Germ-line transformation and immunohistochemical staining of embryos

P-element transformation vectors containing the fusion genes were coinjected with helper plasmid p25.1WC into the embryos of $rosy^{506}$ strain according to Rubin and Spradling (Rubin and Spradling 1982), by Dr. L. Pick. Following transformation into the germ line of *Drosophila*, multiple independent lines were established for each fusion gene. The expression of the fusion genes was monitored by immunohistochemical staining (Gutjahr et al. 1993). Briefly, dechorionated 0-12 hr embryos were fixed with formaldehyde. The fixed embryos were incubated with anti-β-galactosidase antibodies (Cappell). After washing, the embryos were incubated with anti-rabbit IgG biotinylated antibodies. ABC

reagents (Vector Lab) were added, providing avidin-mediated biotinylated peroxidase coupling to the biotinylated antibodies. Staining was carried out in the presence of diaminobenzidine (DAB), $NiCl_2$ and H_2O_2. The embryos were photographed using Nomarski optics on a Zeiss Axiophot microscope with TMAX-100 film after staining.

10. Isolation of *P1* cDNA clones

To isolate the *P1* cDNA, low stringency hybridization of P1 degenerate oligonucleotides to an embryo cDNA library was performed. Degenerate oligonucleotides were synthesized according to the six amino acid sequence of P1 based upon the genetic code. The sequence of the oligonucleotide was 5'-ATGCGC(T)ACC(T)ATC(T)AAC(T)T(A)C(G)C (G)-3'. It was labeled with γ-^{32}P-ATP in a T4 polynucleotide kinase reaction. The probe was used to screen a *Drosophila* embryo cDNA library prepared from mRNA isolated from 2 to 14 hr old embryos (Stratagene). 300,000 plaques were screened. It is estimated that there are about 5000 essential proteins in *Drosophila* (Lewin 1980). Prehybridization buffer (PHB) was 6 X SSC, 20 mM sodium phosphate buffer, pH 6.4, 0.1% SDS, 0.1 mg/ml sonicated salmon sperm DNA, 1 X Denhardt's solution. Prehybridization was carried out at 37°C for 5 hr in a MAXI oven (Labnet). Hybridization solution was PHB without Denhardt's solution containing 100 ng of the labeled probe. Hybridization was carried out at 37°C for 18 hr. Filters were washed with 6 X SSC/0.1% SDS for 15 min at 37°C, 40°C and

45°C sequentially. Three positive clones were obtained. They were converted from the original λZAP vectors to plasmid Bluescript SK vectors by the excision procedure described by the manufacturer (Stratagene). Briefly, the isolated phage clones were mixed with Exassist helper phage and XL-1 blue cells at 37°C for 2 hr, then 70°C for 20 min. The supernatant containing the excised plasmid was used to transform SOLR cells. Individual plasmid clones were obtained by plating transformed cells on ampicillin-containing LB plates. The plasmid DNA was purified by the boiling lysis method. The DNA sequences of the three clones were determined by "oligonucleotide walking" using Sequenase Version 2.0 (United States Biochemical). Both strands of the three clones were sequenced. The oligonucleotides used as sequencing primers are schematically shown in Figure 20.

11. Southern and Northern Blot of the *P1* gene

A 26 bp oligonucleotide, P1-R2 (3'-GGTTGGTCACGCACT TCTAGGGATTC-5', lower strand of *P1* cDNA clone B at position of 290-306 bp, Fig. 21) used as a primer for sequencing of the *P1* cDNA was labeled with γ-^{32}P-ATP in a polynucleotide kinase reaction. This probe was used in both Southern and Northern blot for the *P1* gene.

Genomic DNA was isolated from *Drosophila* embryos. 0.5 g embryos were homogenized in 5 ml of lysis buffer (0.1 M Tris-HCl, pH9.0, 0.1 mM EDTA, 1% SDS). The genomic DNA was extracted from the homogenate with 1 M potassium acetate,

followed by precipitation with one half volume of isopropanol. The pellet was suspended in 0.4 ml TE and was extracted with an equal volume phenol:chloroform (1:1), and precipitated with 100% ethanol. 20 μg of genomic DNA was digested with different restriction enzymes including SalI, BamHI, EcoRI, HincII, HindIII, and EcoRV. The digested DNA was separated on an argarose gel and blotted to a nitrocellulose membrane by capillary transfer in a 20 X SSC solution for 18 hr. Prehybridization in PHB (see Method for library screening) was carried out at 37°C for 2 hr. Hybridization was done in the PHB with 10 ng labeled probes at 48°C for 18 hr. The filter was washed with 6 X SSC/0.1% SDS for 30 min at 48°C, 60°C, and 68°C sequentially. The filter was autoradiographed at -70°C with an intensifying screen for 2 days.

The mRNA from 0-12 hr *Drosophila* embryos was provided by Yan Yu. 2 μg of mRNA was applied to a 1% formaldehyde argarose gel. The mRNA was blotted to a nitrocellulose membrane by capillary transfer in 20 X SSC for 18 hr. Prehybridization and hybridization were carried out as for the Southern blot. Three washes were carried out with 6 X SSC/0.1% SDS for 30 min at 25°C, 55°C, and 60°C individually. The filter was autoradiographed at room temperature for 4 hr.

12. Determination of the *P1* transcript distribution in embryos

In situ hybridization of digoxigenin (DGG) labeled *P1* to 0-16 hr *Drosophila* embryos was carried out using a method modified from Tautz and Pfeifle (1989). DGG-labeled *P1* RNA probes were synthesized in *in vitro* transcription reactions with T7 (for anti-sense RNA probe), and T3 (for sense RNA probe) RNA polymerase using *P1* cDNA in the Bluescript SK vector as template. The DNA templates were cut with XhoI for the T3 RNA polymerase reaction; with EcoRI for the T7 RNA polymerase reaction. The reactions were carried out using an mRNA DGG-labeling kit (Boehringer).

Dechorionated embryos were fixed with 4% formaldehyde, and treated with proteinase K at a concentration of 50 mg/ml for 3 min. Prehybridization was performed at 55°C for 4 hr in a solution containing 50% formamide, 5 X SSC, 100 µg/ml sonicated salmon sperm DNA, 50 µg/ml heparin, 0.1% Tween 20. Hybridization was at 55°C for 18 hr in the prehybridization solution with 20 ng probe. The embryos were washed at 60°C for 20 min with, first, the prehybridization solution, then 0.5 X prehybridization solution. The embryos were incubated with anti-DGG antibody coupled to alkaline phosphatase (Boehringer) for an hour. The enzyme activity was detected by adding its substrates, NBT and X-phosphate (Boehringer). The embryos were photographed using Nomarski optics on a Zeiss Axiophot microscope with TMAX-100 film after staining.

13. *In vitro* transcription and translation of the *P1* cDNA

The *P1* cDNA was translated *in vitro* using a reticulocyte lysate system (Promega) according to instructions from the manufacturer. The *in vitro* transcription and translation was carried out in a 50 µl reaction containing 25 µl TNT rabbit reticulocyte lysate, 2 µl T3 RNA polymerase, 1 µl of 1 mM amino acid mixture minus methionine, 4 µl of 10 mCi/ml ^{35}S-methionine (1000 Ci/mmol), 1 µl of RNasin ribonuclease inhibitor and 1 µg of P1 cDNA (clone A) in the Bluescript-SK plasmids. For the control experiment, the template was the Bluescript-SK plasmid only. The reaction was carried out at 30°C for 90 min. 5 µl of the reaction mixed with 2 µl of 3x PAGE sample buffer (6 mM EDTA, Tris-HCl, pH6.8, 6% SDS, 20% glycerol, 1% (W/V) Bromophenol Blue) was loaded onto a 10% SDS polyacrylamide gel. The gel was exposed for 18 hr for autoradiography after being dried.

14. Identification and purification of Adf-1 and P1 expressed in *E. coli* by using a pET-3a system

The *Adf-1* cDNA in a plasmid expression vector pET3a (Adf1/pET3a) was obtained from Dr. Robert Tjian's laboratory (England et al. 1992). The Adf1/pET3a plasmid was transformed into a host *E. coli* strain BL21(DE3)plysS (Novagen). 5 ml of overnight culture of the transformed cells was added to 500 ml LB to continue growth for 2 hr. The expression of Adf-1 was induced with 0.2 mM isopropyl thiogalactopyranoside (IPTG) at 37°C for 3 hr. The cells were collected by centrifugation, and resuspended in 20 ml lysis buffer [20 mM

Hepes, pH7.6, 0.1 M NaCl, 1 mM DTT, 0.1 mM phenylmethylsulfonylfluoride (PMSF)]. They were disrupted by sonicating 6 times for 30 sec repeatedly with a 5 mm probe at 3.5 output. Between sonications the cell suspension was cooled on ice. Proteins (180 mg) in the supernatant were loaded onto 1 ml of an O6-sepharose DNA-affinity column equilibrated with 0.1 M NaCl lysis buffer (for preparation of the column, see Method 5). The column was washed with 10 ml of lysis buffer, and eluted with 8 ml of 0.8 M NaCl lysis buffer. 8 ml of the eluate was concentrated and desalted to 1 ml (~0.3 mg protein) using a Centriprep-10 concentrator (Amico). Protein concentration was measured by the method of Bradford (1976) using the Bio-Rad protein assay solution (Bio-Rad Laboratory).

P1 cDNA was subcloned into the pET-3a vector to express the protein in *E.coli*. Two primers were synthesized in order to introduce two NdeI sites into the *P1* cDNA. The 5'-end primer sequence is 5'-GGAATTCCATATGCGTACCATTA-3'; the 3'-end primer sequence is 5'-GGAATTCCATATGGTCTAAGACTCTA-3'. The two primers were used in a polymerase chain reaction (PCR) with the *P1* cDNA (clone A) as template. The amplified DNA fragments were digested with NdeI, and then the NdeI fragments were gel-purified. They were inserted into the NdeI site of the pET-3a vector using T4 DNA ligase. The forward orientation P1/pET-3a construct was checked by restriction enzyme digestion and double stranded DNA sequencing with Sequenase Version 2.0 (United States Biochemical). The

P1/pET-3a was transformed into the *E. coli* strain BL21(DE3)plysS (Novagen). The protein production in *E. coli*, purification of P1 using an O6 column were essentially as described above for that of the Adf1/pET-3a.

The purified proteins were analyzed by 12% SDS-PAGE. The PAGE sample buffer was the same as above in Method 13; the running buffer contained 25 mM TrisBase, 193 mM glycine, 0.1% SDS. The gel was prepared using a minigel system - Mighty small slab gel electrophoresis Unit SE280 (Hoefer Scientific Instructions). The gel was stained for 10 min in 50 ml staining buffer (0.1% Coomassie Brilliant Blue R-250, 40% methanol, 10% acetic acid), and destained for 30 min with several changes of 50 ml of destaining buffer (40% methanol, 10% acetic acid).

For Adf1, the recombinant protein enriched by O6 DNA-affinity chromatography was further confirmed by a Western blot of the purified proteins with anti-Adf-1 antibody (obtained from Dr. Robert Tjian's laboratory). The purified proteins were separated using SDS-PAGE, followed by electroblotting onto a PVDF membrane (Millipore). The protein blotting was carried out using 1 L of transfer buffer (700 ml of SDS-PAGE running buffer and 300 ml of methanol), electroblotting for 2 hr at 150 mA at room temperature in a small transfer electrophoresis Unit (Hoefer Scientific Instructions). The PVDF membrane was blocked with 3% bovine serum albumin (BSA) in TTBS (0.1% Tween 20, 0.1 M Tris-HCl, pH 7.5, 0.9% NaCl) for 18 hr, followed by incubating for 1 hr

with anti-Adf-1 antibody at a 1:5000 dilution in a buffer containing 7 ml of TTBS, 2 ml of 2% BSA, 1 ml of fetal calf serum. The membrane was washed with TTBS, and incubated for 15 min with the second antibody (biotinylated anti-rabbit IgG, Vector Laboratories) at 1:5000 dilution in the same buffer. After washing, the membrane was transferred to 20 ml of ABC reagent (Vector Laboratories) plus 2 ml 15% BSA for 30 min. The protein bands were revealed by detecting peroxidase activity with its substrates diaminobenzidine plus nickel chloride (DAB-Ni) in the presence of H_2O_2 as described by the manufacturer (Vector Laboratories).

15. Purification of the recombinant P1 expressed as a GST fusion protein in *E. coli* using a pGEM expression system

To express P1 as a GST fusion protein, the *P1* cDNA was cloned into an *E. coli* expression plasmid - pGEM-3X (Pharmacia). The *P1* cDNA fragment was released by cutting the *P1* clone A (Fig. 20) in the Bluescript-SK vector with EcoRI and XhoI. The fragment was purified using a GENECLEAN kit (BIO-101). pGEX-3X was cut with EcoRI and ends were dephosphorylated with calf intestine phosphatase. After ligation of the EcoRI sites of the fragment and vector, the unligated ends were filled in using DNA polymerase large fragment (Klenow). The resulting blunt-ends were ligated with T4 DNA ligase. The construct was transformed into host strain HB101. The construct was verified by restriction enzyme digestion and partial DNA sequencing.

100 ml overnight culture of the transformed host cells was added to 400 ml LB and grown for 1 hr at 37°C. Production of the fusion protein was induced with 0.2 mM IPTG for 3 hr at 30°C. Disrupting cells with sonication was done as in Method 14. A glutothione-Sepharose 4B column (Pharmacia) was first used to try to purify the GST/P1 fusion protein according to the instructions from the manufacturer. Proteins in the supernatant of the sonicated cells were loaded onto the column. After washing, the proteins were eluted with 20 mM reduced glutothione. The fusion protein was also isolated with the O6 affinity column as described in Method 14. The Western blot of the purified proteins with anti-GST antibody was carried out as in Method 14. The result of this experiment is not shown.

16. Transcriptional activation function of Adf-1 assayed in a yeast growth system

To express Adf-1 in yeast *Saccharomyces cereviciae*, the *Adf1* cDNA in a plasmid pBSADFa (obtained from Dr. R. Tjian's laboratory) was subcloned into a yeast expression vector pADNS, which contains an ADH (alcohol dehydrogenase gene) promoter and terminator and a Leu 2 selective marker (obtained from Dr. J. Gerst). Genes expressed with this vector have to provide their own translation start codon. The *Adf-1* cDNA fragment was released from the vector pBSADFa by cutting with BglII and NotI, and the fragment was purified with a GENECLEAN kit (Bio-101). pADNS was cut with NotI, and

ends were dephosphorylated using calf intestine phosphatase. Two NotI sites of the fragment and vector were first ligated. The BglII site of the fragment and the other NotI site of the vector were then filled in using DNA polymerase large fragment (Klenow). The construct Adf1/pADNS was transformed into *E. coli* strain HB101, and the correct orientation was confirmed by restriction enzyme digestion.

Adf1/pADNS was transformed into two yeast host strains. The yeast strains were constructed by Yan Yu in our laboratory. One yeast host strain contains a reporter *His3* gene. The transcription of the *His3* gene is controlled by a 323 bp *ftz* proximal enhancer fragment (323-HIS3 reporter, see Fig. 29). Another host strain carries the His3 reporter, but under the control of the NP6 sequence (NP6-HIS3), which contains a six copies of an *engrailed* protein binding site (Fig. 29). The transformants were selected using a selective plate without leucine. As a control, the pADNS was also transformed into the host strain carrying 323-HIS3 reporter. The *ftz* gene, cloned into Ycp50, a modified yeast expression vector containing the ADH (*alcohol dehydrogenase* gene) promoter and terminator, was also constructed by Yan Yu. This *ftz* expression vector was transformed into the host strain with the NP6-HIS3 reporter. Growth of the host strains transformed with Adf1/pADNS, pADNS, *ftz*/Ycp50 were tested on selective plates with or without 3-aminotriazole (3-AT).

17. <u>Transcriptional activation function of P1 assayed in a yeast growth system</u>

The *P1* cDNA (clone A, Fig. 20) was cloned into the pADNS vector (Method 16). The P1 cDNA fragment was released from the Bluescript SK vector by digestion with NotI and XhoI. The pADNS yeast expression vector was cut with NotI, and ends were dephosphorylated with calf intestine phosphatase. After ligation of the NotI sites of the fragment and vector, the other NotI and XhoI sites were filled in using DNA polymerase large fragment (Klenow). They were ligated again to generate the P1/pADNS construct. The correct orientation of the construct was confirmed by restriction enzyme digestion. The construct was transformed into two yeast host strains, one with the 323-HIS3 reporter, another with the NP6-HIS3 reporter (see Method 16). Growth of the host strains transformed with P1/pADNS and pADNS were tested on selective plates with or without 3-aminotriazole (3-AT).

CHAPTER 3. RESULTS

1. Identification of DNA-binding activities with DNA fragments from the *ftz* proximal enhancer

The *ftz* upstream element was previously identified as a enhancer for the *ftz* gene expression in seven stripes (Hiromi et al. 1985; Hiromi and Gehring 1987). It was further divided into distal and proximal enhancers by deletion analysis followed by P-element mediated transformation (Pick et al. 1990). The proximal enhancer directs *lacZ* fusion gene expression in seven *ftz*-like stripes both in the mesoderm and ectoderm via a heterologous *hsp70* minimal promoter (Pick et al. 1990).

A biochemical approach was carried out to identify trans-regulators interacting with this well-defined ~1 kb proximal enhancer. The ~1 kb fragment was digested with restriction endonucleases and subcloned. Seven smaller fragments (F1-F7) were generated (Method 1 and Fig. 1). They were ^{32}P-labeled and incubated with 0-12 hr *Drosophila* embryo nuclear extract (NE, Result 7). The DNA-protein complexes were separated from the free probe by electrophoresis through 4% native argarose gels in gel retardation assays (Method 2). The DNA binding conditions for each fragment were optimized, by titrating the concentrations of NE, probe, poly(dI-dC), $MgCl_2$, and KCl. The presence of $MgCl_2$ in the reaction buffer was found to enhance the degradation of the probe when a high concentration of NE was used. Therefore $MgCl_2$ was omitted

from the reaction buffer. The concentrations of NE and poly(dI-dC) were varied from 2 to 6 μg. The salt concentration also was varied between 0.1 M to 0.3 M NaCl without changing the efficiency of binding for any of the complexes.

Of the seven DNA fragments tested in gel retardation assays, only three of them (F2, F3 and F4) were found to interact with proteins in the 0-12 hr NE. These three protein-DNA complexes are shown in Figure 2: complex I with F4, Panel A; complex II with F3, Panel B; complex III with F2, Panel C. To determine whether or not the three complexes (I, II, III) represented sequence-specific protein-DNA interactions, unlabeled probes were used as specific competitors in the binding reactions, and a 230 bp AvaII plasmid fragment from Bluescript KS+ was used as a non-specific competitor. The competitors were added to the reactions at a 40-fold molar excess over the probes. As shown in Figure 2, complex I was competed by F4 (Panel A, lane 4), not by the AvaII fragment (lane 9); complex II was competed by F3 (Panel B, lane 3), not by the AvaII fragment (lane 9); complex III was competed by F2 (Panel C, lane 2), not by the AvaII fragment (lane 8). It is clear that all the protein-DNA complexes are competed by the specific but not by the non-specific competitors. Therefore, the three protein-DNA complexes are the result of proteins in the *Drosophila* nuclear extract interacting specifically with the *ftz* proximal enhancer.

Since the probes F5, F6 and F7 did not interact with proteins in the NE, although various conditions were tested, we wondered if these fragments could compete with the three complexes identified above. As shown in Figure 2, F7 competed for complex I, but not complexes II and III; F5 and F6 did not compete for any of the three complexes (Fig. 2A, B, C, lanes F5, F6, F7). Generally, F5, F6 and F7 in the distal half of the proximal enhancer neither interact with proteins in the NE nor compete for the three complexes. Thus, the ~1 kb fragment of the *ftz* proximal enhancer was narrowed down to ~500 bp (F2, F3, F4), which is the proximal half of the enhancer. This short DNA fragment interacts specifically with DNA-binding proteins present in nuclear extracts prepared from 0-12 hr *Drosophila* embryos.

Since the three complexes migrated at nearly the same position in gel retardation assays, we wondered if they were the same protein(s) binding to the different proximal enhancer fragments repeatedly. As shown in Figure 2, fragment cross-competition showed that the DNA-protein complexes I, II, and III were each competed by F2, F3 and F4 at a 40-fold molar excess over the probes (Fig. 2A, B, C, lanes F2, F3, F4). To test if F2, F3, and F4 compete for the three complexes equally well, we titrated the amount of competitors added to the binding reactions. An increasing amount of each competitor (F3, F2, F4) (Fig. 2D, lanes 2 to 9) was added to the gel retardation reactions with probe F3. As shown in Figure 2D, F3 is the strongest competitor for complex II. F2

and F4 are the strongest competitors for complexes III and I, respectively (data not shown). There are two possible explanations for these differences in competition: (1) Each single shifted band in the gel retardation assays (protein-DNA complex I, II, and III) results from the DNA fragment interacting with more than one protein. Some of the proteins are involved in the formation of all three complexes; (2) Each protein-DNA complex results from the DNA fragment interacting with only one common protein. This protein binds to all three fragments, but with different affinities. The following results suggest that the first possibility is the case.

2. Identification of multiple protein binding sites in the *ftz* proximal enhancer

To define protein binding sites in the ~500 bp region of the *ftz* proximal enhancer, DNaseI footprinting was carried out on the three complexes generated in the fragment gel retardation assays using a modified method as described essentially by Ueda, et al. (1990). The ^{32}P-labeled fragments were incubated with 0-12 hr *Drosophila* nuclear extracts. The reactions were then subjected to DNaseI digestion. The digestion conditions, including the amount of DNaseI, digestion buffer, time and temperature, were carefully optimized to reveal the footprint sites (Method 3). The protein-DNA complexes and the free probes were recovered, after they were electrophoresed through 4% native

polyacrylamide gels. They were then analyzed using 8% sequencing gels.

The footprinting was done on each DNA strand by labeling one strand at a time. As shown in Figure 3, three regions of each DNA fragment (F2, F3, F4) were protected from DNaseI digestion. These footprint sites were numbered 1 to 6 and 8 to 10. The borders of the footprint sites on both DNA strands were determined according to the strand on which the borders could be most clearly defined. The footprint sites (1, 2, 3) detected from complex I are shown in F4 on both strands; the footprint sites (4, 5, 6) detected from complex II are shown in F3 on both strands; the footprint sites (8, 9, 10) detected from complex III are shown in F2 on both strands. Thus, in each protein-DNA complex detected in the gel retardation assays (Fig. 2), the proteins actually interact with three separated DNA sequences generating three footprint sites in DNaseI footprinting assays.

To identify potential direct protein-DNA contact sites, methylation interference analysis was done on the three gel retardation complexes. Fragments F2, F3 and F4 were methylated after ^{32}P-labeling and incubated with 0-12 hr *Drosophila* nuclear extract (NE). The protein-DNA complexes and free probes were separately recovered from native polyacrylamide gels. They were then subjected to chemical cleavage at guanine sites, and the cleaved fragments were separated on sequencing gels. As shown in Figure 4, guanine (G) residues which are directly involved in contacting

proteins were detected on both DNA strands. Each of the nine footprint sites (1 to 6, 8 to 10) contains several guanines that interact directly with proteins; no such guanines were detected outside of footprinting sequences.

A summary of the *ftz* proximal enhancer footprinting and methylation interference results is shown in Figure 5. The nine sites cover about 300 bp of the *ftz* proximal enhancer. The lengths of the footprinted sequences varied from 10 bp to 23 bp. Some of the sites are close together and some of them are very well separated. The striking character is that the nine sites all have the sequence AGGA with the middle GG residues contacting proteins directly. However, one AGGA site at positions 2363 to 2366 was not footprinted. This site and a site at positions 2283 to 2286 were not detected by methylation interference. AGGA is a part of the DNA binding consensus sequence of tramtrack protein (TTK) (Harrison and Travers 1990; Read et al. 1990; Brown et al. 1991). In addition, the methylation interference patterns and DNA sequences lying outside of the AGGA core vary from site to site. Footprint sites 4, 6, 8 and 9 all have the sequence CAAGGA, which is a part of the DNA binding consensus sequence of FTZ-F1 (Ueda et al. 1990). Another interesting feature is that the TTK and FTZ-F1 binding sites overlap at footprint sites 4,6,8 and 9, which share AGGA sequence. Both TTK and FTZ-F1 were previously shown to be *ftz* trans-regulators. TTK protein bound to one site between the proximal and distal enhancers (Harrison and Travers 1990). It was also identified

as a protein that bound to five sites in the zebra element (Harrison and Travers 1990; Brown et al. 1991). FTZ-F1 interacted with two sites in the zebra element (Ueda, et al. 1990). It is quite likely both of these proteins also interact with the proximal enhancer (see below).

The known homeodomain binding core consensus sequence is ATTA (Smith et al. 1990). Previous studies showed that bacterially expressed *ftz* homeodomain bound to the upstream element many times (Pick et al. 1990). We found that only footprint site 2 contains the ATTA sequence. This site was also detected as a high affinity FTZ homeodomain binding site in the enhancer.

3. Identification of multiple specific protein-DNA complexes in the *ftz* proximal enhancer

Gel retardation assays with the *ftz* proximal enhancer fragments as probes (Fig. 2) revealed the formation of three related protein-DNA complexes (I, II, III) with 0-12 hr NE. Nine individual binding sites were identified by DNaseI footprinting and methylation interference assays on the three complexes. To determine how many proteins interact with these nine sites, we further examined their protein binding properties using oligonucleotide gel retardation analysis. Nine oligonucleotides, oligonucleotide 1 (O1) to oligonucleotide 6 (O6) and oligonucleotide 8 (O8) to oligonucleotide 10 (O10) corresponding to the nine sites shown in Figure 5 were synthesized. A control oligonucleotide

(O7) was synthesized. This sequence comes from a non-footprinted region in the *ftz* proximal enhancer. The protein binding properties of the ten oligonucleotide probes were tested in gel retardation assays.

The single strand oligonucleotides were annealed and labeled with α-^{32}P-dATP using DNA polymerase large fragment (Klenow). The oligonucleotide probes were incubated with the 0-12 hr *Drosophila* embryo nuclear extracts (NE). Conditions for each probe in the gel retardation assay were titrated to get clearly separable shifted bands. Typically, the conditions were similar to the conditions used in the DNA fragment gel retardation assay (Method 2). An important difference between the two conditions is that 0.1 µg/reaction of random single strand oligonucleotide had to be included in the binding reactions to eliminate a rampant single stranded DNA binding activity interacting with the dissociated single strand oligonucleotide probes. This single strand oligonucleotide binding activity was not abolished by 0.4M KCl, while all double strand oligonucleotide binding activity was. This suggests that single strand DNA-protein interactions have higher affinity than the double strand protein-DNA interactions. Furthermore, oligonucleotide probes are less stable than longer DNA fragment probes, which is why single strand oligonucleotide probes are formed.

The autoradiographies of four gel retardation assays using oligonucleotides are shown in Figure 6. The protein-oligonucleotide complexes are termed *ftz* enhancer binding

complexes (fEBC). fEBC1 (C1) and fEBC2 (C2) were detected in gel retardation assays with O2 (Panel A). fEBC3 (C3) was detected in gel retardation assays with O3 (Panel B). fEBC4 to fEBC7 (C4 to C7) were detected in gel retardation assays with O4 (Panel C). fEBC8 to fEBC10 (C8 to C10) were detected in gel retardation assays with O6 (Panel D). To test if the fEBCs result from sequence-specific protein-DNA interactions, and to distinguish the oligonucleotide binding activities (fEBCs) from each other, unlabeled oligonucleotides were used as competitors in the binding reactions. As shown in Figure 6, C1 and C2 were competed by O2 only (Panel A, lane 3), but not by any of the other oligonucleotides O3, O4 etc. C3 was competed by O3 only (Panel B, lane 4). For C4-C7, a more complex pattern emerged (Panel C). C4 was competed by O1 to O6 and O9 (lanes 2 to 6, 8); C5 was competed by O4 and O9 (lanes 4, 8); C6 and C7 were competed by O4 only (lane 4). As shown in Panel D, C8 and C9 were competed by O4, O6, O8 and O9 (lanes 4, 6 to 8); C10 was competed by the eight oligonucleotides except O10 (lanes 2 to 8). Thus, unique cross-competition patterns differentiated the different fEBCs.

At this stage, we presume that fEBC1 and fEBC2 are different protein-DNA complexes because of the different shift positions (Fig. 6A). The same situation applies to the other pairs of fEBC6 and fEBC7; fEBC8 and fEBC9. Thus, according to the different cross-competition patterns and, in some cases, to the different shift positions, ten different

DNA-binding activities (fEBCs) were identified in the 0-12 hr nuclear extract. Of course, C2 could also be a degraded form of C1; this could also apply to C6 and C7 as well as C8 and C9. Alternatively, proteins could be differentially modified. Further studies will distinguish these possibilities.

The other oligonucleotides (O1, O5, O8, O9, O10) were used in gel retardation assays. Only complex fEBC10 (O10) is formed with O1 and O5 (Fig. 9A, lane 3; B, lane 2); O8 generates fEBC8 (C8) and fEBC10 (C10) (Fig. 8B, lane 4); O9 interact with fEBC8, fEBC9 and fEBC10 (Fig. 8A, lane 3). No double strand oligonucleotide binding activities were detected with O10 or with the control O7 oligonucleotides (data not shown).

A summary of the ten fEBCs identified by gel retardation assays with oligonucleotide is shown in Figure 7. Complexes were formed with eight of the footprint sites identified in ~300 bp of the *ftz* proximal enhancer. Except for footprint site 10, with which no binding activity was detected, different binding activities were detected for the other sites (1 to 6, 8, 9). fEBC10 was detected with sites 1 and 5; fEBC1 and fEBC2 with site 2; fEBC3 with site 3; fEBC4 to fEBC7 with site 4; fEBC8, fEBC9 and fEBC10 with sites 6 and 9; fEBC8 and fEBC10 with site 8. Among the ten fEBCs, fEBC1 to fEBC7 bound to the proximal enhancer only once; fEBC8, three times; while fEBC9 bound twice; fEBC10, five times. Among the eight footprint sites, three of them (sites

1, 3, 5) interact with only one fEBC; while the others interact with multiple fEBCs.

4. FTZ-F1 and TTK interact with the *ftz* proximal enhancer

a) FTZ-F1 interacts with footprint sites 6, 8 and 9

Since fEBC8 and C10 each bound to multiple sites, consensus sequences could be deduced from their binding sites. They are: CCAAGGAC for fEBC8 and AGGA .for fEBC10. These consensus sequences suggested that fEBC8 contains FTZ-F1 and fEBC10 contains TTK.

FTZ-F1 was identified as a specific DNA-binding protein interacting with the *ftz* zebra element. Its cDNA sequence revealed that it is a member of the nuclear hormone receptor superfamily. It was also suggested that it is a transcriptional activator of the *ftz* gene (Ueda et al. 1990; Lavorgna et al. 1991). As fEBC8 and FTZ-F1 have a similar binding consensus sequences, we used oligonucleotide gel retardation assays to test their relationship. Since fEBC8 bound to O6, O8 and O9, these three oligonucleotides were ^{32}P-labeled and used as probes in gel retardation assays (Fig. 8). An oligonucleotide that contains the known FTZ-F1 binding site I (Ueda et al. 1990) was synthesized and used as an unlabeled competitor. As shown in Figure 8A, the FTZ-F1 oligonucleotide specifically competes for formation of both C8 and C9 but not C10 with probes O6 (lane 2). This suggests that both C8 and C9 contain FTZ-F1. Anti-FTZ-F1 polyclonal antibody (kindly provided by Dr. C. Wu) was added to the gel

retardation reactions. As shown in Panel B, the antibody abolished the formation of fEBC8 with probe O6 (lane 2), while control preimmune serum had no effect (lane 3). In addition, formation of complex C9 and C10 were not affected by inclusion of the antibody in the binding reactions (lane 2). Similar results were obtained with probe O8 (lane 4 to 6) and O9 (data not shown). Interestingly, both antibody and preimmune serum alone also bind to the oligonucleotide probes giving consist "shift" signals (Fig. 8B, lanes 7, 8). We did not further investigate this phenomenon. The results summarized above suggest that, first, fEBC8 is different from fEBC9, as suggested by the different shift positions (Figs. 6D, 8), and, second, that only fEBC8 contains FTZ-F1.

b) TTK interacts with multiple sites in the *ftz* proximal enhancer

tramtrack protein (TTK) was identified as a transcription factor binding to the *ftz* zebra element and the *ftz* upstream element. It is thought to act as a repressor of *ftz* gene expression. Its cDNA sequence showed that TTK is a Zinc-finger protein (Harrison and Travers 1990). It also binds to the *even-skipped* (*eve*) promoter and autoregulatory regions (Read et al. 1990; Brown et al. 1991). The DNA binding consensus sequence for TTK, AGGA, was deduced from three independent studies (Harrison and Travers 1990; Read et al. 1990; Brown et al. 1991). To test the relationship between TTK and fEBC10 (C10), anti-TTK polyclonal antibody

(kindly provided by Drs. L. Brown and C. Wu) was added to the gel retardation assays with O5 (Fig. 9A). The antibody abolished fEBC10 complex formation with probe O5 (lane 2), while control preimmune serum had no effect (lane 1).

We next compared the ability of TTK and nuclear extract to interact with potential TTK binding sites in the proximal enhancer. TTK protein synthesized in bacteria was kindly provided by Drs. D. Read and J. Manley. O1, O5 and O6 were used as wild type probes in gel retardation assays; the corresponding mutant oligonucleotides M1, M5, M6 (Fig. 12) were used as mutant probes. As shown in Figure 9, Panel B, the complex formed with the bacterial TTK protein (lane 1) comigrated with fEBC10 generated with the nuclear extract (lane 2) in the gel retardation assays with O1. As shown in Figure 9, Panel C, TTK interacts with O6 (lane 2), O5 (lane 4) and O1 (lane 6), but not with M6, M5 and M1 (lane 1, 3, 5). Similar results were obtained for fEBC10 (part of the data is shown in Panel D). These experiments reveal that fEBC10 from *Drosophila* embryo nuclear extracts contains TTK protein.

5. The *ftz* enhancer binding activities are present at different stages of embryonic development

The *ftz* gene is expressed not only in a unique spatial but also in a distinct temporal pattern. The *ftz* mRNA stripes are turned on individually within the first 3 hr of embryonic development. They then gradually fade away between 3 and 6 hr

of development (see Introduction). Accordingly, the identified fEBCs should be detected at different stages of embryogenesis if they have different roles in regulating transcription of the *ftz* gene. If the DNA-binding activities of any fEBCs are involved in the establishment and maintenance of *ftz* expression in 0-3 hr embryo development, these fEBCs may function as transcription activators; if the DNA-binding activities of any fEBCs are involved in the fading of the *ftz* mRNA in 3-6 hr of the embryo development, these fEBCs may be transcription repressors. To test this, preliminary experiments were carried out to examine the temporal patterns of the fEBCs by gel retardation assays with staged embryo nuclear extracts.

Nuclear extracts were prepared from 0-2 hr, 2-4 hr, 4-6 hr, 6-9 hr, 16-22 hr and 0-16 hr staged embryos that had been frozen in 50% glycerol at -20^{o}C (Result 7). We do not yet know if this freezing procedure has any effect on extraction of different DNA binding proteins at different developmental stages nor do we know how reproducible this procedure is. Four oligonucleotides were used for gel retardation assays, O2, O3, O4, and O6, since these four footprint sites interact with all the fEBCs detected (Fig. 7). The results of the gel retardation assays are shown in Figure 10. fEBC1, fEBC2 and fEBC11 were detected with O2 (Panel A); fEBC3 and fEBC12 were detected with O3 (Panel B); fEBC4 to fEBC7 were detected with O4 (Panel C); fEBC8 to fEBC10 were detected with O6 (Panel D). Thus, a total of twelve fEBCs were identified interacting

with the ~300 bp fragment of the *ftz* proximal enhancer (Fig. 7). Of the twelve fEBCs shown in Figure 10, fEBC11 was strongest in 0-2 and 2-4 hr nuclear extracts; fEBC3, fEBC4, fEBC8 and fEBC9, were detected in all staged nuclear extracts; fEBC1, fEBC2, fEBC12, fEBC7 and fEBC10 were strongest in 4-6 and 6-9 hr nuclear extracts; fEBC5 and fEBC6 were detected in 16-22 hr nuclear extract. The different temporal patterns of the fEBCs suggest that they may have different functions in embryo development. Except for fEBC5 and fEBC6, the other ten fEBCs were detected in the first 6 hr of embryo development, when *ftz* is expressed. The temporal overlap between the expression of the *ftz* gene and the detection of ten fEBCs interacting with the *ftz* proximal enhancer is consistent with the hypothesis that the ten fEBCs are involved in regulating *ftz* gene expression.

The DNA-binding activities of the ten fEBCs with potential roles in regulating *ftz* expression were different in different staged nuclear extracts. For example, in gel retardation assays with O2 (Fig. 10A), the most active DNA-binding activity in 0-2 and 2-4 hr nuclear extracts was fEBC11. While in the 6-9 hr stage, fEBC1 and fEBC2 were more active. This result suggests that the DNA-binding protein forming fEBC11 is a transcriptional activator, while the proteins forming fEBC1 and fEBC2 are transcriptional repressors. Similar analysis applies to the gel retardation assays with O3 (Fig. 10B), O4 (Fig. 10C), and O6 (Fig. 10D). These results are summarized in Table 1. Five fEBCs (fEBC11,

fEBC3, fEBC4, fEBC8 and fEBC9) were correlated with turning on and maintaining *ftz* expression, suggesting that they are candidate *ftz* activators; five fEBCs (fEBC1, fEBC2, fEBC12, fEBC7 and fEBC10) were correlated with turning off *ftz* expression, suggesting that they are candidate *ftz* repressors. This analysis is strengthened by the findings that a candidate *ftz* activator, FTZ-F1, corresponds to fEBC8 (Fig. 8), and that a candidate *ftz* repressor, TTK, corresponds to fEBC10 (Fig. 9).

In the gel retardation assay with the staged embryo nuclear extracts, twelve fEBCs were identified. These DNA-binding activities appear to be temporarily regulated. Ten of them are candidate *ftz* regulators (Table 1). It seems that four footprint sites in the *ftz* proximal enhancer can interact with both potential transcriptional activators and potential repressors. A further speculation is that *ftz* transcription in stripes is turned off when the repressors bind those sites, competing off the activators that established *ftz* stripes.

6. Functional analysis of the footprint sites in the *ftz* proximal enhancer using transgenic flies

The *in vitro* DNA-binding studies described above identified nine protein binding sites (nine footprinting sites) in a 323 bp region of the *ftz* proximal enhancer (Fig. 5). Those sites interact with multiple potential *ftz* regulators in the staged embryo nuclear extracts (Fig. 7). The results of *in vitro* studies suggested that those protein

binding sites may be important for regulating *ftz* expression *in vivo*. To test the functions of the nine footprinting site *in vivo*, *ftz* proximal enhancer-*lacZ* fusion genes were constructed and their expression patterns were analyzed in transgenic embryos.

All fusion genes were inserted into the vector HZ50PL (Hiromi and Gehring 1987) which contains a *rosy* gene as eye color marker and P-element borders that allow for insertion into the germline by P element-mediated transformation (Rubin and Spradling 1982; Spradling and Rubin 1982). In the plasmid, the *lacZ* reporter gene is transcribed from a heterologous minimal *Drosophila hsp70* promoter. *ftz* proximal enhancer fragments were fused upstream of this minimal promoter. The fusion gene constructs made in the HZ50PL plasmid are shown in Figure 13. After injection of enhancer-*lacZ* fusion genes into *Drosophila* embryos, multiple independent transformant lines were established (see Fig. 13). Fusion gene expression was monitored with anti-β-galactosidase antibodies.

The *ftz* proximal enhancer covering positions 2043 to 2575 (Fig. 13), which directs expression of *lacZ* gene in seven *ftz*-like stripes (Pick et al. 1990), was used as the starting fragment to make other DNA constructs. Deletion of the 5'-end of the proximal enhancer to position 2168 generated Prox 406 (Fig. 13). Deletion of the 3'-end of Prox 406 to the posterior border of footprint sites generated Prox 323. Both Prox 406 and Prox 323 direct *lacZ* expression in

seven *ftz*-like stripes (Figs. 13, 14A). This analysis suggests that a 323 bp fragment of the *ftz* proximal enhancer containing all nine footprinting sites is sufficient to direct fusion gene expression in stripes. Thus, the biochemically identified nine protein binding sites appear to be important in regulating *ftz* gene expression *in vivo*. To our knowledge, the 323 bp *ftz* proximal enhancer is the shortest cis-element identified for a *Drosophila* segmentation or homeotic gene, in that it can direct reporter gene expression in an endogenous gene-like spatial pattern.

To test the role of individual footprinting sites, site-specific mutations of these sites were generated. Nucleotide substitution mutations were made using the overlap extension method (Method 8). As shown in Figure 11, the mutant sites on two template DNA strands were introduced with mutant oligonucleotides in two polymerase chain reactions. The mutant products from the two PCRs were joined by the overlap extension in the third round of PCR (Method 8). A summary of the mutations for each footprint site is shown in Figure 12. The mutant sequences are indicated on top of the corresponding wild type footprinting sites. Nine pairs of mutant oligonucleotides (M1 to M6, M8 to M10) were synthesized (Method 8). They were tested in the gel retardation assays using 0-12 hr nuclear extract. None of them interacted with any proteins in the embryo nuclear extracts (data not shown). They were then used to construct thirteen mutant proximal enhancer *lacZ* fusion genes (Fig.

13). The mutant constructs were named according to which footprinting site was mutated. For example, in the M1 construct, the footprinting site 1 was mutated. Eight constructs: M1, M2, M3, M4, M6, M8, M9, M10, carry mutations in single footprinting sites (Fig. 13). DNA constructs carrying mutations in multiple footprinting sites (Fig. 13) were designed to mutate the sites which are bound by the same fEBCs repeatedly. For example, site 6 and 9 were bound by fEBC8, 9 and 10. Site 8 was bound by fEBC8 and 10 which also bound to site 6 and 9 (Fig. 7). Thus M6,9 and M6,9,8 were made (Fig. 13). To make the double mutant M6,9, the M6 construct was used as the template in the overlap extension method (Method 8). To make the triple mutant M6,9,8, the M6,9 construct was used as the template in the PCRs. In all the reactions, the corresponding mutant oligonucleotides were used as the primers.

The fusion genes were integrated into the *Drosophila* germ lines by P-element mediated transformation. multiple independent lines for each DNA construct were established (Fig. 13). The results of analysis of the expression of these fusion genes is summarized in Figure 13. Mutations of individual *ftz* enhancer elements did not affect expression of the reporter gene in seven *ftz*-like stripes. An example of this class of mutation [the site 10 mutation (M10)] is shown in Figure 14B. The site 6 and 9 double mutation (M6,9) dramatically reduced expression of the β-galactosidase. Of the four lines examined, expression in two lines was

undetectable (not shown) and expression in two lines was extremely weak (Fig. 14C). Further mutation of site 8 (M6,9,8) abolished expression of the reporter gene entirely in all lines examined (Fig. 14D). Furthermore, there was no detectable reporter gene expression in the constructs with mutations in additional footprint sites, such as M6,9,8,4, M6,9,8,1, and M6,9,8,1,5. Clearly, only multiple mutations of the footprint sites affected gene expression. These results indicate that the *ftz* gene is controlled by multiple cis-elements in a redundant fashion. Evolutionarily, this might render gene expression more resistant to mutagenic events in the life of the *Drosophila*.

The original design of mutations of the multiple enhancer elements was intended to test the contribution of individual fEBCs to the regulation of *ftz* gene expression. Since fEBC8 (contains FTZ-F1) binds to footprint sites 6, 8, and 9 (Fig. 7), loss of the reporter gene expression due to mutation of those three sites (M6,9,8) suggests that FTZ-F1 acts as a transcriptional activator in the regulation of the *ftz* gene expression. However, other proteins also bind to these sites (see below) which may also activate *ftz* expression.

In addition, the same sites also interacted with fEBC10 (TTK), a potential repressor of *ftz* expression. Since these TTK binding sites overlap activator binding sites, its ability to repress transcription via the *ftz* proximal enhancer cannot be assessed in these experiments.

No stripe specific cis-element was identified in the mutational analysis of the nine footprint sites in the *ftz* proximal enhancer. Mutation of only two sites (M6,9) affects the expression of all seven stripes. This suggests that the seven *ftz* stripes are coordinately regulated.

To test the function of footprint sites 1, 2, 3, and 4, combinations of mutations in those sites can be further studied. Although mutations did not individually affect reporter gene expression, the indications from the M6,9,8 mutations suggest that they could be also important by acting together, or with the other sites in directing the transcription of the *ftz* gene.

From the P-element transformation studies of the *ftz* proximal enhancer and the site-specific mutagenesis of the enhancer elements, the importance of the *in vitro* findings are established. The *in vitro* defined minimal *ftz* proximal enhancer acts like the longer proximal enhancer in directing the transcription of the reporter gene expression in seven *ftz*-like stripes. The footprint sites 6, 8, 9 are essential for the function of the minimal *ftz* proximal enhancer. The complexity of the organization of the *ftz* transcriptional machinery reflected by the *in vitro* protein-DNA binding studies are restrained by the redundant function of the footprint sites in the regulation of *ftz* transcription, as demonstrated by the *in vivo* mutagenesis analysis.

7. Purification of potential *ftz* proximal enhancer binding proteins

The protein-DNA binding studies described above identified nine footprint sites in the proximal enhancer (Fig. 5). Of the nine sites, sites 6, 8 and 9 were necessary for the function of the proximal enhancer *in vivo* (Fig. 13). Twelve fEBCs were identified with eight of the nine sites, except site 10 (Fig. 7). The proteins in the nuclear extract binding to the three necessary sites all interacted with site 6 (fEBC8, 9, 10; Fig. 7). In addition, a unique group of proteins in the nuclear extract interacted with site 4 forming fEBC4 to fEBC7 (Fig. 7). To purify these proteins interacting with sites 4 (O4) and 6 (O6), the two most interesting footprint sites, O4 and O6 DNA-affinity chromatography was performed in the purification.

A scheme of the steps used for purification of the potential *ftz* enhancer binding proteins (fEBPs) is shown in Figure 15. All steps were carried out in the unlabeled room.

Step 1. Embryonic nuclear extract The procedure used is modified from three reports (Soeller et al. 1988; Biggin and Tjian 1989; Wamplar et al. 1990). 0-12 hr embryos were harvested from four mass population cages of *Drosophila melanogaster* (Oregon R). The embryos were dechorionated in 3% clorox for 90 sec, washed sequentially with 0.7% NaCl/0.04% Triton X-100; 0.7% NaCl; and water. The embryos then were either processed directly at 4°C or stored in 50% glycerol at -20°C for up to several months. All of the following steps

were carried out at 4°C. All solutions used contained the following protease inhibitors: 50 µg/ml soybean trypsin inhibitor, 1 mM benzamidine, 1 mM DTT, 2 U/ml aprotinin, 1 µg/ml antipain, 1 µg/ml bacitricin. The dechorionated embryos were homogenized in a Yamato LH-21 homogenizer at the speed of 1500 rpm. They were passed through the homogenizer only once in homogenization buffer (0.35 M sucrose, 15 mM Hepes, pH7.6, 0.1 M KCl, 2.5 mM MgCl2, 1 mM EDTA, 10% glycerol) at a concentration of 4 ml buffer/g embryo. The homogenate was filtered through three layers of nylon cloth (6.3 mm mesh). The filtrate was centrifuged in GSA bottles for 10 min at 10,000 rpm. The whitish pellets were resuspended in lysis buffer (15 mM Hepes, pH7.6, 0.1 M KCl, 3 mM MgCl2, 0.1 mM EDTA, 10% glycerol) at a concentration of 1 ml buffer/g embryos in 70Ti tubes. One tenth volume of a saturated ammonium sulfate solution, pH7.9, was added to the 70Ti tubes. The tubes were rotated for 30 min. The lysate was centrifuged in a 70Ti rotor for 45 min at 36,000 rpm. The supernatant was precipitated by addition of solid ammonium sulfate to a final concentration of 30% saturation, followed by stirring for 30 min. The precipitate was centrifuged in GSA bottles for 20 min at 10,000 rpm. The pellet was resuspended in sample buffer (homogenization buffer without sucrose) at a concentration of 1 ml buffer/10 g embryo. The sample was dialyzed against the sample buffer without salt for 1 hr until the conductivity of the sample was ~0.2 M NaCl. The protein concentration was measured by the method of

Bradford (1976) using a Bio-Rad protein assay solution. The sample was stored at -70°C. The protein concentration of the extracts was typically 20-30 mg/ml. 1,900 mg protein was obtained from pooled nuclear extracts derived from 1,300 g embryos.

Step 2. Heparin-sepharose chromatography ~1584 mg of nuclear extract was diluted to 10 mg/ml with the 0.2 M NaCl HEG buffer (25 mM Hepes, pH7.6, 0.1 mM EDTA, 10% glycerol, protease inhibitors used as in the homogenization buffer). The diluted NE was loaded onto a 140 ml heparin-sepharose CL-6B column equilibrated with 4 column volumes of 0.2 M HEG. The column was washed with 280 ml of 0.2 M HEG and step eluted sequentially with 280 ml of 0.4 M HEG and 280 ml of 1.0 M HEG. The flow rate was 2 ml/min. Fractions of 9 ml were collected and analyzed for fEBCs activities using the gel retardation assay. The active 0.4 M fractions and 1.0 M fractions were pooled separately. The two fractions were precipitated by addition of solid ammonium sulfate to a final concentration of 40% saturation and resuspended in no salt HEG. The 0.4 M Heparin fraction (H0.4) contained 303 mg protein; and the 1.0 M Heparin fraction (H1.0) contained 99 mg protein. The two protein fractions were frozen immediately at -80°C. fEBC activities in the NE and the two Heparin fractions were examined using gel retardation assays with O2 and O3 (data not shown); with O4 and O6 (Figs. 16D-19D, lane 2; see below). The protein profiles of the NE and the two

Heparin fractions were analyzed by SDS-PAGE and silver staining (Figs. 16A-19A, lanes 2, 3; see below).

DNA affinity chromatography

The two Heparin fractions were purified further by three-rounds O4 and O6 DNA-affinity chromatography (Fig. 15). Purification of the H0.4 Heparin fraction is presented below. Purification of the H1.0 Heparin fraction through O4 and O6 affinity columns was essentially as described for the purification of the H0.4 fraction.

The two DNA-affinity columns were made by coupling the concatamerized oligonucleotide 4 (O4 affinity column), and oligonucleotide 6 (O6 affinity column) to sepharose 4B respectively (Method 5). The bed volume of the DNA-affinity columns was 2 ml. The flow rate of the affinity columns was 0.3 ml/min. The fraction size was 0.6 ml (fraction/2 min). The affinity columns were repeatedly used three times. They were regenerated by washing with 50 ml of 2 M NaCl HEG buffer, followed by equilibrated with 20 ml of 0.2 M NaCl HEG buffer between runs. These parameters were used for all three-rounds affinity chromatography, as described below.

Step 3. The first-round of DNA-affinity chromatography. The 0.4 M Heparin fraction (H0.4, ~303 mg) was diluted to 0.2 M NaCl with no salt HEG. Nonspecific DNA competitors of poly(dI-dC) (50 µg/mg protein) and single stranded oligonucleotides (O7 upper strand, 10 µg/mg protein) were

added to the H0.4 fraction. The amount of the competitors was estimated from the gel retardation assay, that is, 0.5 μg poly(dI-dC)/10 μg H0.4 protein; 0.1 μg O7 upper strand/10 μg H0.4 protein. The sample was incubated on ice for 10 min and clarified by centrifuging at 10,000 X g for 10 minutes. The supernatant was loaded onto the O4 column. The passthrough of the O4 column was directly loaded onto the O6 column. The two columns were then individually washed with 8 ml of 0.1 M HEG and eluted with 8 ml of 0.8 M HEG. The eluate fractions (~8 ml) from O4 and O6 column were pooled individually and diluted to 0.2 M NaCl. They were then termed as O4 first eluate and O6 first eluate respectively. The eluate fractions were pooled, because from a pretest we knew that the DNA-binding activities were spread in all eluate fractions as revealed by the gel retardation assay (data not shown). No DNA-binding activity in the washing fractions was detected in the pretest. The O4 and O6 first eluates were frozen immediately at -80°C. The DNA-binding activities of the two eluates were examined by using O4 and O6 as probes in the gel retardation assays after freezing (Figs. 16D-19D, lane 3).

The O4 and O6 first eluates were then handled one at a time with the same procedure as described below. An example of further purification of the O4 first eluate is given.

Step 4. The second-round of O4 DNA-affinity chromatography. The poly(dI-dC) and single strand

oligonucleotides used above were added to the O4 first eluate at a concentration of 10 μg/mg protein. The amount of competitors was suggested from titration using the O4 first eluate in the gel retardation assays with O4. The mixture was incubated on ice for 10 minutes. After it was clarified by centrifuging at 10,000 X g for 10 minutes, the supernatant was loaded onto the O4 affinity column a second time. The column was washed with 8 ml of 0.1 M HEG, and eluted with 8 ml of 0.8 M HEG. The eluate fractions were pooled and diluted to 0.2 M NaCl with no salt HEG. The diluted sample was termed O4 second eluate.

Step 5. The third-round of O4 DNA-affinity chromatography. The O4 second eluate was immediately mixed with nonspecific DNA competitors used above at 1 μg/mg protein. After incubation for 10 min on ice, the mixture was precipitated as above. The supernatant was loaded onto the regenerated O4 column for the third-round of purification. The column was washed with 8 ml of 0.1 M HEG, and eluted with 8 ml of 0.8 M HEG. The elution fractions were pooled (O4 third eluate) and stored at -80°C immediately without dilution.

The DNA-binding activity from frozen samples of the O4 second and third eluates was measured using gel retardation assays with O4 as probe (Figs. 16D-19D, lanes 4, 5). Samples of the O4 first, second, and third eluate were analyzed by SDS-PAGE at the end of the three-rounds of affinity

purification, and the proteins on the gels were detected by silver staining (Figs. 16A-19A, lanes 3, 4, 5).

Step 6. Preparation of the purified proteins for microsequencing. For microsequencing of the purified polypeptides, the O4 third entire eluates (~8 ml) was precipitated by four times cold acetone (-20°C) for 18 hr at -80°C. After centrifuging at 12,000 X g for 30 min, the pellet was resuspended with SDS sample buffer (6 mM EDTA, Tris-HCl, pH6.8, 6% SDS, 20% glycerol, 1% (W/V) Bromphenol Blue). The sample was electrophoresed through a 10% SDS polyacrylamide gel. After the gel was stained with Coomassie-Brilliant Blue in dH_2O, the stained protein bands were cut out. The gel slice was loaded into the wells of second 15% SDS polyacrylamide gel. The wells were then filled with V8 protease in the SDS sample buffer (for details, see Method 6). The protein in each gel slice was partially digested during electrophoresis, and the digested polypeptides were also separated by the SDS-PAGE. After the digested polypeptides were transferred to a PVDF membrane by electrophoresis, the membrane was stained with Coomassie-Brilliant Blue. The individual bands were cut out and sent for microsequencing. This step is discussed in detail for each fraction below.

Characterization of the purified polypeptides. Three methods were performed to test DNA-binding activities of the purified polypeptides before sequence analysis.

(1) UV crosslinking of proteins to nucleic acids. After incubating purified proteins with labeled O4 or O6, protein-DNA complexes were electrophoresed through native 4% polyacrylamide gels. The complexes were localized by autoradiography. They were excised and exposed to UV light, and then analyzed using SDS-PAGE. Different UV wavelengths and intensities were tested. Only one weak band (~50 kDa) was observed for the O4 third eluate of the Hep 0.4 M fraction (data not shown). This difficulty to link protein to DNA probe was perhaps because the probe was not modified with bromodeoxyuridine (BrdU). BrdU containing DNA is more sensitive than unsubstituted DNA to UV-induced crosslinking (Chodosh 1986).

(2) Southwestern method. Southwestern blotting was performed essentially as described by White etc. (1985). After the purified proteins were electrophoresed through a SDS polyacrylamide gel, they were transferred onto a Nitrocellulose membrane. The proteins in the membrane were denatured with 6 M guanidine-HCl and renatured by removing guanidine-HCl. ^{32}P-labeled O4 or O6 were then incubated with the membrane. Probe-protein complexes were detected by autoradiography. Although several bands were detected, no specific patterns were observed for different probes. It is

possible that the proteins were not renatured properly in the membrane.

(3) Denaturation and renaturation of purified proteins. This protocol is described in Method 7. Purified proteins were separated using SDS-PAGE. Different regions of the gel were excised. Proteins in the gel slices were extracted. They were then denatured with 6 M guanidine-HCl and renatured by dilution of the guanidine-HCl (for detail, see Method 7). The processed proteins were tested for their DNA-binding activities using gel retardation assay. The results of this analysis is described below.

a) Purification of P1, P2 and P3 from the Heparin 1 M fraction using O6 affinity chromatography

P1, P2 and P3 were purified by O6 affinity chromatography of the 1 M NaCl Heparin faction (Figs. 15, 16). The proteins in each purification step were revealed by silver staining of SDS polyacrylamide gels (Fig. 16A). Three major bands, indicated by P1, P2 and P3, were enriched. The apparent molecular masses of the three proteins are 29 kDa, 34 kDa, and 53 kDa for P1, P2 and P3 respectively.

After separating the third affinity fraction on a preparative SDS polyacrylamide gel, the proteins were visualized with Coomassie Brilliant Blue staining (Method 6). P1, P2, and P3 were excised and digested with V8 protease while they were electrophoresed through an SDS polyacrylamide gel. The digested peptides were transferred onto a PVDF

membrane, which was stained with Coomassie Brilliant Blue as shown in Figure 16B. The V8 digested peptides indicated by arrows were sequenced. Reliable sequences were obtained for P1, P2, and P3.

Two minor species were also detected in the preparative SDS polyacrylamide gel stained with Coomassie Brilliant Blue (Method 6), but they were not detected in the gel stained with silver (Fig. 16A). The reason for these difference is that in the preparative gel, all proteins in the final fraction were loaded, while in the silver stained gel only a very small portion of the final fraction (1%) was analyzed. Thus, a total of five proteins were recovered from the preparative gel. The two minor species are termed XNS (54 kDa) and QLL (40 kDa). No reliable sequences for them were obtained because of their low levels.

The peptide sequences obtained for P1, P2, and P3 were compared with protein sequences in available data bases. All fifteen amino acids of P2 perfectly matched to the first fifteen amino acids of Adf-1, which is a *Drosophila* transcription factor regulating the expression of the *alcohol dehydrogenase* gene and other homeotic genes (England et al. 1990). No sequences in the data bases matched the other two sequences of P1 and P3. Cloning of the gene encoding P1 will be described later.

To determine which of the proteins contributes to the DNA-binding activities of the O6 second eluate from the 1 M Hep fraction, the second affinity fraction was separated on

an the SDS polyacrylamide gel which was then cut into four slices as indicated on the right of Figure 12A. The DNA-binding activities of the denatured and renatured proteins from the four gel slices (Method 7) were tested in gel retardation assays. As shown in Figure 16C, the proteins from region 1 (R1, lane 1) and region 2 (R2, lane 2) failed to interact with probe O6. Proteins from region 3 (R3, lane 3, 4) and region 4 (R4, lanes 5, 6) did interact with probe 6 (lanes 3, 5), but very weakly with probe O2 (lanes 4, 6). The results suggest that P3 is the DNA-binding activity for region 3, and that either P1 or P2 is the activity for region 4.

A summary of the DNA-binding activities for each purification step using gel retardation assays with O6 as probe is shown in Figure 16D. The positions of fEBC8, fEBC9 in the nuclear extract (NE, lane 1) were enriched in the Heparin and affinity 1 fractions (lanes 2, 3). Then two shifted complexes were formed with the affinity 2 and 3 fractions (lanes 4, 5). The change in size of these complexes as compared to the NE might be due to a change in protein composition resulting from splitting the nuclear extract into two parts on the Heparin-sepharose column. fEBC10 in the NE (lane 1) was not detected in the Heparin fraction and thereafter (lanes 2-5). It was found in the 0.4 M Heparin fraction (see below Fig. 19C, lane 2). A shifted band below fEBC10 in lanes 2-5 is probably a degraded form of protein(s) involved in formation of fEBC8 and fEBC9.

In spite of the position changes of the fEBCs, the proteins in the affinity 3 fraction still bound to O6 specifically, since they were competed by a 50-fold molar excess of unlabeled O6 (Fig. 16D, lane 6), but not by O10 (Fig. 16D, lane 7) in the gel retardation assay. Since the renatured proteins migrated at different positions (Fig. 16C) compared with the native proteins in Figure 16D, it is possible that the fEBC8 and fEBC9 are formed by multiple polypeptide subunits, only one of which binds to DNA directly. This is also supported by the observation that fEBC8 and fEBC9 activities were partially separated in both 0.4 M and 1 M Heparin fractions (compare Fig. 16D, lane 2 with Fig. 19C, lane 2).

b) Purification of P3 from the Heparin 1 M fraction using O4 affinity chromatography

P3 was also purified by O4 affinity chromatography of the 1 M Heparin fraction (Fig. 17). Only one band was strongly stained by silver in the third affinity fraction, although other minor bands were detected (Fig. 17A). The strong band was processed for microsequencing. The protein was digested with V8 protease and transferred onto a PVDF membrane, which was stained with Coomassie Brilliant Blue as shown in Panel B. The amino acid sequence obtained was the same as the sequence of P3 which was purified by O6 affinity chromatography of the 1 M Heparin fraction.

The renatured P3 from region 3 (Panel A) was shown to bind to O4 but not O2 in gel retardation assays (Panel C, lanes 4, 5). Position of the shift in lane 4 was the same as that in Figure 16C, lane 3. This observation suggests that the two renatured proteins in O6 and O4 affinity eluates from the 1 M Hep fraction are the same. None of the proteins from other regions of the gel displayed DNA binding activity (Panel C, lanes 1 to 3).

DNA-binding activity was monitored for each purification step using gel retardation assays. As shown in Figure 17D, fEBC4-fEBC7 in the NE (lane 1) were not clearly resolved after heparin-sepharose chromatography (Heparin and affinity 1 fractions, lanes 2, 3). This may reflect loss of protein from complexes due to chromatographic separation or, alternatively, protein degradation during the purification. Two clearly resolved complexes were formed with O4 using the affinity 2 and 3 fractions (lanes 4, 5). The two bands bound to O4 specifically, since they were competed by 50-fold excess of unlabeled O4 but not by the same amount of O10 (lanes 6, 7). It is likely that the two bands are formed by P3 interacting with O4 as a monomer and as a dimmer. This is supported by the presence of an inverted repeat in the O4 sequence, which is AGGATATGTAGGA (Fig. 5). However, the renatured protein generated only one band in the gel retardation (Panel C, lane 4). Thus, it is possible that the dimerization potential of the protein was not renatured. Purification of P3 from the 1 M Heparin fraction with both O4

and O6 affinity chromatography is not a surprise, because there is a common nucleotide sequence (CAAGGA) in the two oligonucleotides. Renaturation of the same protein from both O6 and O4 purification fractions strongly suggests that P3 is the DNA-binding activity which was purified by both affinity columns.

c) Purification of P4, P5 and P6 from the Heparin 0.4 M fraction using O4 affinity chromatography

P4, P5, and P6 were co-purified by O4 affinity chromatography of the 0.4 M Heparin fraction (Figs. 15, 18). As shown in Figure 18A, proteins from each fraction of the purification (10-50 µg) were separated through SDS polyacrylamide gels. Silver staining of the gels revealed three major bands in the third affinity fraction. Major bands indicated as P4, P5 (a doublet), and P6 were prepared for micro-sequencing.

After separating the third affinity fraction on a SDS polyacrylamide gel, the proteins were visualized with Coomassie Brilliant Blue staining. P4, two bands of P5, and P6 were excised and digested with V8 protease, as described in Method 6. The digested peptides were transferred onto a PVDF membrane which was stained with Coomassie Brilliant Blue as shown in Figure 18B. The two bands of P5 gave the same V8 digestion pattern. Therefore, only the digested peptides from the lower P5 band are shown (Fig. 18B, lane 2). Protein sequences were obtained for P4, P5, and P6 from the peptides

indicated by arrows (Fig. 18B). The same sequence was obtained from each of the three peptides from P5 (Fig. 18B, lane 2). One of the V8 digested peptides from the upper band of P5 also gave an identical sequence (data not shown). Comparing sequences for P4, P5 and P6 with the data bases revealed that P4 and P6 are two new sequences in the data bases. P5 matched the sequence of TTK at positions 334 to 348 (Harrison and Travers 1990).

Since the predicted molecular weight of TTK is 69 kDa, and P5 is 45 kDa, as estimated from the SDS polyacrylamide gel, TTK appears to be purified as two degraded forms of TTK with the same N-terminus (the P5 doublet). This was confirmed by the results of micro-sequencing. The amino terminal sequence of the undigested P5, as well as the two digested peptides (Panel B, lane 2), were the same. The most amino terminal amino acid of P5 matched the 334th amino acid of TTK. The purification of the C-terminal half of the TTK protein by O4 affinity column is consistent with the localization of its DNA-binding domain in this region (Harrison and Travers 1990).

To localize the DNA-binding activities in the O4 eluate, the second O4 affinity fraction (100 µl) was separated through an SDS polyacrylamide gel. Proteins were recovered from four regions of the gel as indicated in Figure 18A. Renatured proteins from individual regions were tested for their ability to bind to probe O4 and O6. As shown in Figure 18C, protein(s) from region 4 bound to probe O4 strongly

(lane 4). Weak binding to O6 was also detected (lane 5). Proteins renatured from the other regions of the gel did not bind to probe O4 (lanes 1, 2, 3). This result indicates that not every DNA-binding protein can be renatured. For example, P5/TTK was not renatured from the region 3 (lane 3). Therefore, renaturation of purified proteins should not be the only criteria for evaluating if a purified protein is a DNA-binding protein. On the other hand, if purified proteins can be renatured, that is strong evidence supporting the idea that the protein is a DNA-binding protein.

The DNA-binding activities were monitored for each purification step using gel retardation assay. A shifted band of fEBC7 (Fig. 18D, lanes 1 to 5) and two bands below fEBC7 (Fig. 18D, lanes 2 to 5) were maintained during the purification. The other bands above fEBC7 were lost during the purification (Fig. 18D, lanes 1 to 5). According to the protein sizes of P4, P5 and P6, it seems that P4 could correspond to fEBC7, P5 and P6 could correspond to the two additional shifted bands below fEBC7 (lanes 2 to 5).

d) Purification of P5, P7 and P8 from the Heparin 0.4 M fraction with O6 affinity chromatography

Three major proteins were purified by O6 affinity chromatography of the 0.4 M NaCl Heparin faction (Figs. 15, 19). The proteins in each purification step were revealed by silver staining of SDS polyacrylamide gels (Fig. 19A). Three major bands, indicated by P7, P5 (see below) and P8, were

enriched. The apparent molecular masses of the three proteins are 51 kDa, 45 kDa, and 33 kDa for P7, P5 and P8 respectively.

After separating the third affinity fraction on an SDS polyacrylamide gel, the proteins were visualized with Coomassie Brilliant Blue staining (Method 6). P7, P5, and P8 were excised and digested with V8 protease as described in Method 6. The digested peptides were transferred onto a PVDF membrane, which was stained with Coomassie Brilliant Blue as shown in Figure 19B. The V8 digested peptides indicated by arrows were sequenced for P7, P5 and P8.

Two minor species were also detected in the preparative SDS polyacrylamide gel stained with Coomassie Brilliant Blue (Method 6), but they were not detected in the gel stained with silver (Fig. 19A). The reason for this difference is that in the preparative gel, all proteins in the final fraction were loaded, while in the silver stained gel, only a very small portion of the final fraction (1%) was analyzed. Thus, a total of five proteins were recovered from the preparative gel. The two minor species are a 48 kDa peptide (namely MN, not sequenced) and a 28 kDa (namely DFK) (Fig. 15).

The V8 digestion pattern of the middle band in Figure 19B, lane 2 matched that of P5 purified with O4 in Figure 18B, lane 2. Therefore, this band was also named P5 in Figure 19A. Its identity was confirmed by sequencing the middle peptide of the V8 digested P5 (Fig. 19B, lane 2). The

sequence matched the P5 sequence purified with O4 (Fig. 18). The peptide sequences of P7 and P8 were compared with protein sequences in the available data bases. No identical sequences in the data bases were found. The sequence for DFK was not reliable, because it was present at very low level. Renaturation of the purified proteins from this O6 affinity fraction was not done.

Purification of TTK by O6 affinity chromatography was expected, because TTK activity in the embryo nuclear extract had been detected in the gel retardation assays with probe O6 as fEBC10 (Figs. 7, 9). fEBC10 was maintained during the purification as shown in Figure 19C. P7 and P8 were co-purified with the TTK. The DNA-binding activities of P7 probably corresponds to the gel shifted band just below fEBC9 in Figure 19C, lane 2 to 5. The gel shifted band below fEBC10 in Figure 19C, lane 2 to 5, is probably the activity of P8. Both of the proteins might be components which form the gel retardation activities of fEBC8 and fEBC9 as discussed above. The specificity of the three purified activities was demonstrated in the gel retardation assay in Figure 19C, lane 5 to 7. The three binding activities in Lane 5 were competed by a 50-fold molar excess of O6 (lane 6), but not by the same amount of O10 (lane 7).

e) <u>Summary of fEBC purification</u>

The purification of potential *ftz* proximal enhancer binding proteins (fEBPs) is summarized in Table 2 and Figure

15. A total twelve proteins were recovered from preparative gels, eight major and four minor species. Eleven of them were partially sequenced. Only the eight major purified proteins generated reliable sequences. Two of the proteins are known transcription factors regulating gene expression in the early *Drosophila* embryo. TTK has been suggested to be a transcriptional repressor of the *ftz* gene, interacting with the *ftz* zebra element (Brown et al. 1991; Brown and Wu 1993). As expected, the purification of P5 (TTK) by O4 and O6 further confirmed the interaction of TTK with the *ftz* proximal enhancer as described earlier with gel retardation assays (Fig. 7). In addition to the five TTK binding sites - footprinting sites 1, 5, 6, 8, 9 (Fig. 7), site 4 also interacts with TTK/P5. Thus, six binding sites for TTK/P5 were identified in the 323 bp *ftz* proximal enhancer. The purification of Adf-1 by O6 affinity chromatography suggests that it regulates transcription of the *ftz* gene in addition as well as the *alcohol dehydrogenase* gene.

The protein-binding specificity for O4 and O6 are reflected by the purification of different set of proteins, which is consistent with the gel retardation assay with the 0-12 hr embryo nuclear extracts (Fig. 7). P1, P2, P7 and P8 were only purified by O6 affinity chromatography, while P4 and P6 were purified by O4 only. The differential protein-binding activities are due to the sequence differences between O4 and O6. On the other hand, the existence of some common sequence between O4 and O6 (AGGAAC), accounts for the

purification of the same set of proteins (P3 and P5/TTK) by both O4 and O6. The purification of multiple proteins by both O4 and O6 DNA-chromatography is in agreement with the observation of multiple protein-binding activities for each oligonucleotide (Fig. 7). Further isolation of the genes encoding the purified proteins will be necessary to confirm their DNA-binding activities, and to study their biological function(s) in the regulation of *ftz* gene expression.

8. Characterization of P1 protein

a) Isolation and sequencing of *P1* cDNA clones

P1 was purified from staged *Drosophila* embryo nuclear extracts with O6 affinity chromatography (Figs. 15, 16, Table 2). The O6 was a 19 bp sequence identified in *in vitro* protein-DNA binding assays in a region of the *ftz* proximal enhancer (Figs. 5, 7). Site-specific mutagenesis of this sequence, along with the other two similar sequences (footprint sites 8 and 9), abolished transcription of a *lacZ* reporter gene in transgenic *Drosophila* embryos (Figs. 13, 14). Several proteins, besides P1, bind to those sites (Fig. 7). However, the *in vivo* mutagenesis demonstrated that this site is necessary for expression, suggesting that P1 might be one of the proteins that plays an important role in the control of *ftz* gene expression.

Comparison of partial amino acid sequence from P1 to known gene sequences in data bases, P1 was found to be a novel protein. In order to examine the function of P1 in the

regulation of *ftz* gene expression, and perhaps other functions in early embryogenesis, cDNAs encoding P1 were isolated and sequenced.

To isolate cDNAs encoding P1, the partial amino acid sequence, obtained by sequencing the purified protein, was used to generate deduced nucleotide sequences (Method 10). Degenerate oligonucleotides were synthesized and ^{32}P-labeled. They were used as probes to screen a 2-14 hr *Drosophila* embryos cDNA library (Strategene). 3 x 10^5 plaques were screened at low stringency (Method 10). Three clones were isolated after three-rounds of screening. Sequence analysis revealed that all three cDNAs contain an identical open reading frame. Sequence encoding the amino acids obtained by sequencing of the purified P1 was found in the beginning 18 bp of the open reading frame. This strongly suggests that the P1 cDNA was cloned.

The lengths of the three P1 cDNA clones are 722 bp, 875 bp and 930 bp. All contain translation start and stop codons, polyadenelation signals and poly(A) tails (Fig. 20). The 5' untranslated regions of the three clones are different, which makes lengths of the clones different (Fig. 20). The 3' untranslated regions of the three clones are similar, except that a 5 bp sequence was not found in clone B as indicated in Figure 20. Those species appear to be generated by different splicing in the 5' untranslated regions. Alternatively, different promoters might also generate of different P1 mRNA species. This possibility can be further examined by analysis

of the P1 gene structure. It will also be interesting to find out if the generation of different P1 species is regulated differently at different stages of embryogenesis and/or in different cells and tissues.

The full length sequence of the P1 clone B is shown in Figure 21. A potential nuclear transport signal (ProArgArgArgPro) was noted in the end of the protein. The five amino acids with three positively charged arginine and two proline residues resembles a typical nuclear transport signal peptide of SV40 T antegen (Lanford and Butel 1984; Goldfarb et al. 1986). The calculated molecular weight of the 185 amino acid protein is 20.5 kDa, which is close to the 29 kDa of *Drosophila* P1 based on analysis using SDS/PAGE (Fig. 16A). The *P1* cDNA was transcribed and translated *in vitro,* using a reticulacyte lysate system (Promega). The translated product is 29 kDa as analyzed by using SDS/PAGE (Fig. 23C), which is consistent with the apparent molecular mass of the native protein.

b) Sequence comparison of P1 with the ribosomal protein L9 family

To learn more about P1, both protein and DNA data bases were searched for genes encoding similar sequences. The BLAST program at the National Center for Biotechnology Information (NCBI) was used for the search. The ribosomal protein L9 (RL9) was found to be the closest sequence to P1.

RL9 sequences from different species were compared with the rat RL9, the first RL9 identified (Suzuki et al. 1990), using the Pileup and Pretty programs (Genetics Computer Group (GCG), Inc). The sequence identities of the rat RL9 to the yeast (Jones et al. 1991), plant pea (Cohn 1992), C.elegan (McCombie et al. 1993) and human RL9 (Hori et al. 1993) are 49%, 55%, 56% and 97% respectively. This reveals that the closer the species, the better alignment between the sequences of their RL9s. The high degree of identity also suggests that the RL9s perform similar functions in those species, although nothing is known about the function of RL9. Surprising, the *Drosophila* P1 shows only 36% sequence identity to the rat RL9, which is even lower than the identity of yeast RL9 with rat RL9 (49%).

Consensus sequence of the RL9 proteins from human, rat, pea, C.elegan and yeast are indicated by the in shaded boxes in Figure 22. Sequences of P1 which are identical to the RL9 consensus sequences are indicated in open boxes (Fig. 22). The homologous region between RL9 family and P1 is limited almost entirely to the 73 amino acid domain from position 96 to 168 aa of P1 which is underlined. 65% of the amino acids in this domain are identical to the RL9 consensus sequence (Fig. 22). In the amino terminal region of the protein, only 24% of the amino acids are identical between the RL9 consensus sequences and P1. No conservation is found in the carboxy terminal amino acids.

The sequence analysis revealed that a 73 aa region of P1 is significantly similar to the RL9s with overall 65% identity and 71% similarity. This suggests that a similar function is conserved in this region of P1 and RL9s. The lack of sequence similarity to RL9 in the rest of P1 suggests that P1 may play a role in *Drosophila* which is different from that of the RL9s in other species.

c) Northern and Southern analysis

A Northern blot of 0-6 hr embryo mRNA was performed. The anti-sense probe was a 26 bp oligonucleotide, corresponding to the P1 cDNA at positions of 290-315 bp (Fig. 21). As shown in Figure 23B, a single broad band of approximately 900 bp was detected. The position of band in the Northern blot (Fig. 23B) is in agreement with the size of the P1 cDNAs (722-930 bp). This indicates that P1 cDNAs are roughly full length. The broadness of the band probably reflects the three mRNA species with similar sizes represented by the three cDNAs.

Southern blot analysis of *Drosophila* genomic DNA was performed. The probe was the same one used for the Northern blot. The genomic DNA was digested with several restriction enzymes. After argarose gel electrophoresis and transfer to a PVDF membrane (Millipore), hybridization and washing were carried out as described in Method 11. As shown in Figure 23A, only one band was detected in each lane, suggesting that the P1 gene is a single copy gene in the *Drosophila* genome.

d) Expression pattern of the *P1* gene in the *Drosophila* embryos

To help elucidate the function of the P1 protein, the spatial and temporal distribution of the mRNA in early embryogenesis was examined by in situ hybridization. Sense and anti-sense RNA P1 probes labeled with digoxigenin (DGG) were prepared by *in vitro* transcription reactions as described in the Method 12. Hybridization of the probes to different stages of embryos is shown in Figure 24. Hybridization with the sense probe serves as a negative control for the staining (Fig. 24A, B, C, D). With the antisense probe, the expression of the P1 mRNA was detected in all stages analyzed (0-16 hr old embryos). Stained embryos at the syncitial blastoderm, cellular blastoderm, germband extension and germband retraction stages are shown in E, F, G, H respectively in Figure 24. P1 mRNA appears to be expressed in all cells of these embryos. The expression of P1 in the youngest syncitial blastoderm embryos suggests that the mRNA is deposited in the embryo maternally (Fig. 24E). Zygotic expression of the mRNA was seen homogeneously in all older embryos (Fig. 24F, G, H). The overall expression of the *P1* gene throughout embryogenesis suggests that it is an important developmental gene in the regulation of early embryogenesis, although it may also function in other stages of *Drosophila* development.

The expression of *P1* overlaps with the expression of the *ftz* gene in both spatial and temporal fashions, which

suggests that it might be a positive regulator of the transcription of the *ftz* gene. The dramatic difference between the expression patterns of the *P1* and *ftz* genes indicates that *P1* also has other functions.

e) The DNA-binding specificity of P1

The DNA-binding activity of the *Drosophila* P1 was suggested by its purification with DNA-affinity chromatography. The molecular cloning of the *P1* cDNA made it possible to further characterize its biochemical properties. The *P1* cDNA was expressed in *E. coli* using a pET-3a system (Novegen). After induction of the protein expression with IPTG, the protein was partially purified from *E.coli* extracts using the O6 affinity column.

The DNA-binding activity of P1 was examined using the gel retardation assay with ^{32}P-labeled O6 as probe. The recombinant P1 interacted with O6 in a protein concentration dependent fashion (Fig. 25A). Two shifted bands were detected in the gel retardation assay (Fig. 25A). The two bands are probably formed by different forms (monomer and dimmer) of the P1 interacting with the probe O6. Alternatively, the two bands are formed by different degraded forms of the protein. However, this is unlikely (see discussion below).

To examine the specificity P1 binding, competition analysis was performed in the gel retardation assay with the ^{32}P-labeled O6 as probe (Fig. 25B). The P1 shifted bands were competed with a 100-fold excess of unlabeled O6 (Lane 3).

However, binding was not competed with a 100-fold excess of the mutant O6 (M6) (Lane 5). The M6 is the mutant form of O6 with three nucleotide substitutions (Fig. 12). As a control, protein was prepared from *E.coli* transformed with the pET-3a vector without insert in the same fashion as described for P1. No control proteins bound to probe O6, as shown in Figure 25B, Lane 6.

Since P1 contains a region homologous to RL9, random single-stranded oligonucleotides and tRNA were used in all the gel retardation assays shown. Generally, a thousand excess of the oligonucleotides and tRNA (~0.1 μg) was added to the binding reactions. No effects on the binding of P1 to probe O6 were observed. Therefore, P1 interacts specifically with double stranded DNA. It is not simply a ribosomal protein contaminant in the purified fraction.

The similarities among the nine footprint sites in the *ftz* proximal enhancer (Fig. 5) suggested that the P1 might interact with other sites in addition to site 6 (O6). The nine oligonucleotides corresponding to the nine footprint sites were synthesized, and labeled with α-^{32}P-dATP. Gel retardation assays were carried out by incubating the recombinant P1 with the nine labeled probes in the presence of a thousand-fold excess of random single-stranded oligonucleotide and tRNA and 100-fold excess of poly(dI-dC). Different types of shifted bands were observed for the reactions with different probes (Fig. 25C). (1) Two strong shifted bands were detected in reactions with O6 and O9 as

probes (lanes 6, 8); (2) One strong shifted band was observed in reactions with O2 and O5 as probes (lanes 2, 5); (3) One weak shifted band was seen in reactions with O1, O3, O4, O8 and O10 as probes (lanes 1, 3, 4, 7, 9). The control *E. coli* protein did not bind to the O10 probe (lane 10). These results demonstrate that P1 interacts with different DNA sequences with different affinities. This is not a simple function of length of oligonucleotides or of A/T content.

The two shifted bands in the gel retardation assays with O6 and O9 were not seen with the other seven probes (Fig. 25C). Therefore, this suggests that the two bands are not a result of degraded P1 forms in the *E.coli* extracts. The lower shifted band in Figure 25C, lane 6 and 8, together with the single shifted band in the other lanes is probably formed by a P1 monomer interacting with all nine probes, but with different affinities. P1 was also prepared from the *E.coli* extracts by denaturation (with 6 M guanidine-HCl) and renaturation. This preparation did not affect the formation of the two bands, suggesting that the two bands are not formed because of different conformations. Since the *E.coli* control protein did not interact with any DNA sequences tested, it is unlikely that a contaminant *E. coli* protein interacting with P1 lead to the formation of the upper band. Therefore, the upper shifted band in lane 6 and 8 is probably formed by a P1 homodimer interacting with probes O6 and O9.

The DNA-binding consensus sequence, (A/G)ANCCAAGGAC (N as any nucleotides; Fig. 25D), for the presumptive homodimer

form of P1 was deduced from the two binding sites O6 and O9. The DNA-binding consensus sequences, AGGA (Fig. 25D), for the presumptive monomer form of P1 was deduced from all nine binding sites. We propose that the homodimerization of P1 increases the selectivity of the protein in choosing its DNA targets.

In summary, the specific DNA-binding activity of P1 is supported by the following. First, P1 was purified from embryo nuclear extracts and *E.coli* extracts by O6 DNA-affinity chromatography. Second, P1 bound to O6 in a protein concentration dependent fashion. Third, P1 binding to O6 was competed with wild type, but not mutant O6 in gel retardation assay. Finally, multiple P1 binding sites were identified in the *ftz* proximal enhancer. P1 bound to those sites not only with different affinities, but also apparently as different isoforms.

f) The potential transcriptional activation function of P1

Since P1 binds to multiple sites of the 323 bp *ftz* proximal enhancer, its transcriptional function *in vivo* was examined in a yeast *Sacchoramyces cereviciae* system available in our laboratory (Yu and Pick, unpublished). The yeast strain carries a reporter gene *His3* under the control of the 323 bp *ftz* proximal enhancer. Activation of the *His3* gene allows for growth of the transformed yeast in plates containing with 3-aminotriazol (3-AT). 3-AT is a competitive inhibitor of the *His3* product (Kishore and shah 1988). The

higher the expression level of the *His3* gene, the better the growth in higher concentrations of the 3-AT. The *P1* cDNA was inserted into a yeast expression plasmid pADNS (provided by Dr. J. Gerst). The P1/pADNS was then transformed into the yeast strain. As a control, pADNS was also transformed into yeast. The growth of the transformants were tested on plates containing 3-AT (Fig. 26). P1/pADNS transformants grew well (Fig. 26, lane 1), while the pADNS transformants did not grow on the same plate (Fig. 26, lane 2). This results suggests that P1 can augment the expression of the *His3* gene.

To examine if the activation is DNA-sequence dependent, a yeast strain containing the *His3* gene under control of six copies of an *engrailed* protein binding site (NP6) was used as a non-specific DNA target site for P1 (Desplan et al. 1985). The *NP6-His3* fusion gene was activated by the expression of *ftz* protein in the yeast, since *ftz* protein binds to this site, as shown by the cell growth in the plate without histidine (Fig. 26, lane 3). In contrast, P1 did not activate fusion gene expression, since no yeast growth was observed in the histidine minus plate (lane 4). Nevertheless, we know that P1 might bind to NP6, since it binds to O2 which interacted with FTZ protein (Pick et al. 1990). This suggests that the activation of transcription by P1 is dependent upon the presence of the P1 binding sites in the 323 bp *ftz* proximal enhancer.

9. Characterization of P2 protein

a) P2 is a known transcription factor, Adf1

P2 was purified from the 1 M Heparin fraction using O6 affinity chromatography (Figs. 15, 16, Table 1). Sequence of fifteen amino acids from the N-terminal of P2 was obtained by peptide micro-sequencing. In a search for homologs using the BLAST Network Service at the National Center for Biotechnology Information, the fifteen amino acids were found to exactly match the N-terminus of Adf1. Furthermore, the molecular mass of P2 (Fig. 16A) and that of Adf1 (England et al. 1990) appear to be the same, 34 kDa estimated by SDS-PAGE. Adf-1 was shown to be a transcription factor interacting with the distal promoter of the *alcohol dehydrogenase* gene (*Adh*) (Heberlein et al. 1985; England et al. 1990; England et al. 1992). Other DNA-binding sites for Adf-1 were found in the promoters of *Antennapedia* (*Antp*) P1 and *dopa decarboxylase*. The Adf1 protein was detected uniformly in early stage embryos (0-12 hr). However, *Adh*, a potential Adf1 target, is expressed only in later stage embryos (after 10 hr) in the fat body. The appearance of distinct expression patterns for *Adf1* and *Adh*, together with the detection of multiple potential targets for Adf1 (*Antp*, *Ddc*), strongly suggest that it has multiple functions.

To examine the interaction of Adf1 with the *ftz* proximal enhancer, recombinant Adf1 was prepared. The *Adf1* cDNA in an *E. coli* expression vector, pET-3a (Novegen), was provided by Dr. R. Tjian. The Adf1/pET-3a plasmid was transformed into

the *E.coli* host strain BL21(DE3) plysS. Following induction with IPTG, Adf1 was partially purified with O6 affinity chromatography (Method 14). Proteins in the eluate from the O6 column were analyzed by SDS-PAGE. As shown in Figure 27A, one major band in the position of 34 kDa was enriched in the O6 fraction. To determine whether or not the 34 kDa band was Adf1, a Western blot was performed with anti-Adf1 antibody (provided by Dr. R. Tjian). As shown in Figure 27B, the 34 kDa band was stained with anti-Adf1 antibody (lane 1). No similar band was detected from the control proteins (lane 2), which were prepared from the pET-3a transformed *E.coli* using the same procedure used for the preparation of Adf1. Thus, the 34 kDa recombinant protein was Adf1.

The partially purified recombinant Adf1 protein was used in gel retardation assays with ^{32}P-labeled O6 as probe. Adf1 interacted with the probe as shown by gel retardation assays with antibodies (Fig. 27C). The shifted band was abolished by addition of anti-Adf1 antibody to the reactions (lane 2), but not by addition of the pre-immune serum (lane 3). The O6-binding activity of Adf1 is protein-dependent, as shown by the protein titration in the gel retardation assay (Fig. 27D). As the protein concentration in the reactions was increased, the amount of shifted probe increased accordingly. Therefore, the recombinant Adf1 interacts with O6 in a protein concentration dependent fashion.

b) The DNA-binding specificity of Adf1

To test the DNA-binding specificity of Adf1, competition analysis in gel retardation assays with O6 was performed. Unlabeled wild type O6 and mutant O6 (M6) with three nucleotides substitutions (Fig. 12) were used as competitors. An autoradiograph of the competition results is shown in Figure 28A. The Adf1-O6 complex in lane 1 was not abolished by the addition of unlabeled O6 and M6 at 50 and 100 fold molar excess over the probe (lanes 2, 3, 6, 7). However, it was abolished by the addition of unlabeled O6 at 250 to 500-fold excess over the probe (Fig. 28A, lanes 4, 5), it was not abolished by the addition of the same amount of M6 (Fig. 28A, lanes 8, 9). These results indicate that the protein bound to the O6 specifically. The control proteins used above in the Western blot were also tested in gel retardation assays. No DNA-binding activity was detected for the control proteins (Fig. 28A, lane 10).

To identify other binding sites for Adf1, gel retardation assays were carried out with the nine labeled footprint sites from the *ftz* proximal enhancer (Fig. 5). As shown in Figure 28B, five of the nine probes interacted with Adf1. Of the nine probes, O5, O6, O8, O9 and O10 were bound by the protein (lanes 5 to 9), the other four probes of O1, O2, O3 and O4 did not detectably interact with the protein (lanes 1 to 4). A possible consensus DNA-binding sequence was deduced from the five Adf1 binding sites, with the possible sequence of (C/T)CN(A/G)(G/A)G(A/G)N(G/A) (Fig. 28C). This

consensus sequence cannot be found in the other four probes which do not interact with the protein.

This consensus sequence is different from the one, GC(T/C)G(C/T)(C/T)G(C/T)CG(C/T)(C/T/A) reported by England, et al. (1990). The observation suggested that Adf1 might bind to the two sequences with different affinity. The O6 fraction containing *E.coli* recombinant Adf1 was used in the gel retardation assays with probe O6. An Adf1 binding site in the *Adh* distal promoter (termed Adh distal site) was synthesized (England et al. 1990). Different amounts of unlabeled O6 and Adh distal site were added to the reactions to compete Adf1 binding to probe O6. Autoradiograph of the competition results is shown in Figure 28D. The major Adf1-O6 complex (lane 1) was competed by titration of unlabeled O6 at 50, 100, and 200-fold molar excess over the probe O6 (lanes 2, 3, 4). It was competed equally well by titration of Adh distal site at 5, 10 and 20-fold molar excess over the probe O6 (lanes 5, 6, 7). The results indicate that Adf1 binds to the Adh distal site with 10 times higher affinity than it does to the footprint site 6 in the *ftz* proximal enhancer.

c) <u>The potential transcriptional activation function of Adf1.</u>

To determine whether of not the native Adf1 in embryo nuclear extract bound to O6, gel retardation assays were performed using an anti-Adf1 antibody. The 0-12 hr embryo nuclear extract was incubated with ^{32}P-labeled O6 in the presence of nonspecific competitor poly(dI-dC) and random

single stranded oligonucleotide. As shown in Figure 29A, three fEBCs (8, 9 and 10) were detected (lane 1), which were observed before (Fig. 6D). Addition of the anti-Adf1 antibody to the reactions specifically abolished the formation of fEBC8 with probe O6 (lane 2). Addition of the pre-immune serum to the reactions did not affect the formation of any of the fEBCs with O6 (lane 3). Since fEBC8 was also abolished by the anti-FTZ-F1 antibody (Fig. 8), it is likely that the fEBC8 is formed by the interaction of the two proteins with the probe. Protein-protein interaction may be involved in the formation of fEBC8. fEBC8 complex was formed with three footprint sites (O6, O8, O9) in the *ftz* proximal enhancer (Fig. 7). Consistently, Adf1 has been shown to bind to all these sites (Fig. 28B). It has been suggested that FTZ-F1 regulates *ftz* gene expression positively (Ueda et al. 1990). The observations suggest that, like FTZ-F1, Adf1 might positively regulate the transcription of the *ftz* gene by binding to its enhancer.

The recombinant Adf1 interacted with five of the nine footprint sites (O5, O6, O8, O9, O10; Fig. 28B). However, the native Adf1 only bound to three footprint sites (O6, O8, O9), as suggested by fEBC8 complex formation (Fig. 7). Therefore, the DNA-binding consensus sequence for the native Adf1 was deduced from these three sites, which is (A/G)AN(C/T)CAAGGA(C/T) (Fig. 29B). The sequence conservation among the three binding sites for the native Adf1 is much higher than that among the five binding sites for the

recombinant protein. This observation suggests that the native Adf1, together with FTZ-F1, binds to DNA sequences with more stringency than the recombinant protein does.

Since Adf1 binds to multiple sites in the 323 bp *ftz* proximal enhancer, its transcriptional function *in vivo* was examined in a yeast *Sacchoramyces cereviciae* system available in our laboratory (Yu and Pick, unpublished). The yeast strain carries a reporter gene *His3* under control of the 323 bp *ftz* proximal enhancer. Activation of the *His3* gene will be detected by the observation of growth of the transformed yeast on plates with 3-aminotriazole (3-AT). 3-AT is a competitive inhibitor of the *His3* product (Kishore and Shah 1988). The higher the expression of the *His3* gene, the better the growth of the cells in higher concentrations of 3-AT. The *Adf1* cDNA was cloned into a yeast expression plasmid pADNS (provided by Dr. J. Gerst). The P2/pADNS plasmid was then transformed into the yeast strain. As a control, pADNS was also transformed into the yeast strain. The growth of the transformants was tested on plates with 2.5 mM 3-AT. As shown in Figure 29C, three Adf1 transformants grew well on the 3-AT plate (lane 1). The three pADNS transformants (Control) did not grow on the same plate at all (lane 2). The results suggests that Adf1 can augment the expression of the *His3* gene.

To determine whether or not the activation was DNA-sequence dependent, a yeast strain containing the *His3* gene under control of six copies of the *engrailed* protein binding

sites (NP6) was used as a non-specific DNA target site for Adf1 (Desplan et al. 1985). The *NP6-His3* fusion gene was activated by the expression of *ftz* protein in the yeast, since the *ftz* protein binds to this site, as shown by the cell growth in the plate without histidine (Fig. 29C, lane 3). In contrast, Adf1 did not activate the fusion gene expression, since no yeast growth was observed in the histidine minus plate (lane 4). These observations suggest that the activation of transcription by Adf1 is dependent upon the presence of Adf1 binding sites in the 323 bp *ftz* proximal enhancer.

In summary, P2 was found to be a previously identified transcription factor, Adf1, by peptide sequence analysis. The recombinant P2 interacted specifically with five of the nine protein binding sites in the 323 bp *ftz* proximal enhancer, sites O5, O6, O8, O9 and O10. Therefore, P2 was termed *ftz* enhancer binding protein 2 (fEBP2). The involvement of the endogenous Adf1 in the formation of fEBC8 suggests that it regulates the *ftz* expression positively, as does FTZ-F1. Adf1 activated a reporter gene expression in yeast in a DNA sequence-dependent fashion. Thus, these studies have unraveled a possible function for Adf1 in the regulation of *ftz* gene expression through interacting repeatedly with the *ftz* proximal enhancer.

CHAPTER 4. DISCUSSION

An ~1 kb DNA fragment at positions 1502 to 2574 in the *ftz* upstream element was identified as the *ftz* proximal enhancer, which directed gene expression in seven *ftz*-like stripes via a heterologous promoter (Pick et al. 1990). We have taken a biochemical approach to understand the molecular basis for the function of the *ftz* proximal enhancer. The *ftz* proximal enhancer was analyzed by *in vitro* protein-DNA binding experiments. Nuclear extracts were prepared from staged *Drosophila* embryos, and they were used in DNaseI footprinting and methylation interference analysis. Nine DNA footprint sites were identified in a 323 bp region of the *ftz* proximal enhancer (Fig. 5). This 323 bp fragment directed *ftz-lacZ* fusion gene expression in seven *ftz*-like stripes in transgenic embryos (Figs. 13, 14). Site-specific mutations of the footprint site 6, 8 and 9 abolished the expression of the *ftz-lacZ* fusion gene in transgenic embryos (Figs. 13, 14). This demonstrates that these three sites are crucial for the function of the proximal enhancer.

To identify trans-factors binding to the *ftz* proximal enhancer, gel retardation assays were performed with the nine footprint sites. Twelve *ftz* enhancer binding complexes (fEBCs) were identified from staged embryo nuclear extracts. Of the twelve fEBCs, ten were found to be active in the early embryo when the *ftz* gene is expressed (Fig. 10), suggesting

that they are involved in regulation of *ftz* gene expression (Table 1).

To isolate the proteins binding to footprint sites 4 and 6, protein purification was carried out with DNA-affinity chromatography (Fig. 15). Twelve proteins were partially purified from the 0-12 hr nuclear extracts. Clear amino acid sequences were obtained for eight (P1 to P8) of the twelve proteins by protein microsequencing. Of the eight proteins, P2 and P5 are previously identified transcription factors, Adf-1 and TTK respectively (discussed below). Adf1 and FTZ-F1 are involved in the formation of fEBC8 (Figs. 8, 29A). P5/TTK is involved in the formation of fEBC10 (Fig. 9).

The *P1* cDNA was isolated by screening an embryo cDNA library using the deduced nucleotides from the partial amino acid sequence. The *P1* cDNA encodes 185 aa, with a domain of 73 aa in its C-terminal homologous to the ribosomal protein L9 (Fig. 22). However, P1 does not appear to be a contaminant ribosomal protein from the purification. The *E.coli* recombinant P1 was demonstrated to interact specifically with multiple footprint sites in the *ftz* proximal enhancer (Fig. 25). It also activated reporter gene expression in yeast in a DNA-binding site dependent fashion (Fig. 26).

1) Complicated organization of the *ftz* proximal enhancer resembles a typical eukaryotic enhancer

We have identified nine footprint sites and twelve protein-DNA complexes in the *ftz* proximal enhancer (Fig. 7).

Four trans-acting factors, P1, Adf-1 (P2), TTK (P5), and FTZ-F1, were shown to interact with multiple footprint sites specifically (Figs. 25, 28, 9, 8). The importance of three footprint sites (6, 8, 9) in the regulation of the *ftz* expression was further demonstrated by site-specific mutation analysis in the transgenic embryos (Fig. 13). Although a large number of footprinting sites and protein-DNA complexes have been identified for the *ftz* proximal enhancer, this complexity has been observed for several eukaryotic cis-acting elements.

The *ftz* zebra element located within 669 bp proximal to the *ftz* transcription start site was sufficient to direct the expression of *ftz*-like stripes in the mesoderm (Hiromi et al. 1985). Deletion analysis of the zebra element in transgenic embryos revealed the complexity of the element (Dearof et al. 1989a; Topol et al. 1991). Nine cis-acting elements were identified in the zebra element. Five of them activated reporter gene expression; four of them repressed gene expression. Three trans-acting factors have been identified that interact with the zebra element: GAGA factor bound to site fAE3 (Topol et al. 1991); TTK and FTZ-F1 each bound to the zebra element twice (Ueda et al. 1990; Brown et al. 1991).

The Immunoglobulin H intron is a 700 bp enhancer (ENH_{iH}) that acts as a potent B cell-specific enhancer (Staudt and Lenardo 1991). Sixteen protein binding sites and fifteen DNA-binding proteins have been identified. The proteins

interacted with eleven of the sixteen sites. Some of the sites were bound by multiple proteins. For example, the octomer site was bound by three proteins (oct-1, oct-2 and oct-3). Some of them interacted with only one protein, e.g. the E site was bound by only Ig/EBP-1. Of the fifteen proteins, most of them interacted with only one site, but NF-mNF interacted repeatedly with four binding sites. Mutation analysis indicated that the multiple sites were functionally redundant, and that they regulated gene expression both positively and negatively.

2) The biochemical approach

By genetic screens, many developmental genes have been identified (Lewis 1978; Nusslein-Volhard and Wieschaus 1980; Nusslein-Volhard et al. 1984). The expression patterns of those developmental genes, such as segmentation genes and homeotic genes (see Introduction), are regulated at the transcription initiation level. A transcriptional regulatory network among these genes has been suggested on the basis of genetic studies (Akam 1987; Ingham 1988; Johnston and Nusslein-Volhard 1992). However, at least two issues cannot be resolved by genetics. First, not all developmental genes can be identified with genetic screens, such as functionally redundant genes or genes with complementary maternal and zygotic functions. Many genes encoding transcription factors were missed in the genetic screens. They were subsequently identified by other methods. These genes include the

homeobox-containing gene *caudal* (Mlodzik and Gehring 1987), *Drosophila AP-1* (Perkins et al. 1990), *pox meso* and *pox neuro* (Bopp et al. 1989), *ttk* (Harrison and Travers 1990), and *FTZ-F1* (Lavorgna et al. 1991). Second, the regulatory interactions among developmental genes identified by genetic methods are not necessarily direct interactions. For example, the expression pattern of the *ftz* gene was altered in all gap gene mutant embryos (Carroll and Scott 1985). However, this result did not demonstrate if the gap gene products interact directly with the cis-acting elements of the *ftz* gene. Other gene products might be in the middle of the regulatory network mediating the effects of the gap genes on *ftz* gene expression. So far, no direct interaction has been demonstrated between the gap gene products and the *ftz* regulatory elements using biochemical approaches. To complement the genetic analysis, a systematic biochemical approach has been undertaken to study the regulatory mechanisms of the *ftz* proximal enhancer.

a) Identification of crucial cis-acting elements in the *ftz* proximal enhancer

Nine footprint sites in the *ftz* proximal enhancer were mapped by footprinting and methylation interference assays with stage-specific nuclear extracts. The nine sites were localized in a 323 bp region of the ~1 kb proximal enhancer. The 323 bp fragment was sufficient to direct reporter gene expression in seven *ftz*-like stripes (Figs. 13, 14). Site-

specific mutations of three footprint sites (6, 8, 9) abolished expression in all seven stripes (Figs. 13, 14). These experiments demonstrated that the *in vitro* identified footprint sites 6, 8, and 9 are necessary for the function of the 323 bp proximal enhancer.

Harrison and Travers (1988) conducted a direct DNase I footprinting assay with the whole *ftz* upstream element using crude embryo nuclear extracts. In the proximal enhancer region of the *ftz* upstream element, nearly all of the sequence (70%) was footprinted. Some footprinting sites were over 100 bp long. However, no further evidence (e.g. gel retardation assay) was provided to show that the individual footprint sites interact with proteins specifically. The biological function of those sites were not examined. Eight of the nine footprint sites (except site 10) identified in our study lie within the sites identified by them. This sequence overlap suggests that the footprint sites reported in this study are the core sequences in mediating the regulation of *ftz* gene expression.

b) Protein complexes determine DNA-binding specificity

The *in vitro* biochemical approach allows for the identification not only of important cis-acting elements, but also of trans-acting factors. It was surprising that three footprint sites were detected from each protein-DNA complex identified in the gel retardation assays with three DNA fragments (Figs. 2, 3, 4, 5). This observation was explained

by the gel retardation assays with individual footprint site (Fig. 6), in that each footprint site, except site 10, interacted with at least one protein from the nuclear extracts (Fig. 7). Therefore, each DNA fragment (F2, F3, F4 in Fig. 1) interacted with three groups of proteins leading to the identification of three footprint sites. Thus, the protein-binding specificity for each of the three DNA fragments in the *ftz* proximal enhancer was determined by a unique combination of different proteins in the 0-12 hr embryo nuclear extract.

Each footprint site was analyzed in gel retardation assays. A total of twelve *ftz* enhancer binding complexes (fEBCs) were identified (Figs. 6, 7). Most of the footprint sites interacted with multiple protein complexes. Therefore, every DNA fragment (F2, F3, F4 in Fig. 1) actually interacts with a unique combination of fEBCs (Fig. 7). F2 interacts with fEBC8, 9, 10; F3 interacts with fEBC4, 5, 6, 7, 8, 9, and 10; F4 interacts with fEBC1, 2, 3, 10, 11, 12. This observation suggests that the right combination of different proteins in the 0-12 hr nuclear extracts determines where they interact with the *ftz* proximal enhancer.

c) Purification and microsequencing of several potential *ftz* proximal enhancer binding proteins

DNA-affinity chromatography was used to purify proteins interacting with two of the nine footprint sites in the *ftz* proximal enhancer (O4 and O6, Fig. 15). Similar purification

strategies have been successfully used in the purification of a number of *Drosophila* transcription factors, including Zeste which interacts with the *Ubx* promoter (Biggin et al. 1988), FTZ-F1 which interacts with the *ftz* zebra element (Ueda et al. 1990), GAGA factor which interacts with a number of cis-elements of developmental genes (Biggin and Tjian 1988; Soeller et al. 1993), E1f-1/NTF-1 which interacts with the cis-elements of *Ubx*, and *Ddc* (Bray et al. 1989; Dynlacht et al. 1989), and DTF1 and DF2 which interact with the *Antp* P2 promoter (Perkins et al. 1988). The genes encoding most of these transcription factors have been cloned.

Both O4 and O6 interacted with multiple proteins in gel retardation assays. Four (fEBC4-7) and three (fEBC8-10) protein-DNA complexes were detected in the gel retardation assay with O4 and O6 respectively (Figs. 6, 7). Each protein-DNA complex is formed by at least one protein. Therefore, there should be a minimum of four proteins interacting with O4; three proteins interacting with O6. Not surprising, we have purified twelve proteins using O4 and O6 affinity chromatography (Figs. 15-19, Table 2). Eight of them were present at high enough levels to obtain clear amino acid sequences.

Proteins purified form the nuclear extracts by the DNA-affinity chromatography may belong to one of the several categories: first, specific DNA-binding proteins; second, proteins interacting with the DNA-binding proteins; finally non-specific DNA-binding proteins. Of the eight major

species, three proteins P1, Adf1, and P5/TTK are specific DNA-binding proteins interacting with O6 specifically (Figs. 25, 28, 9; Table 2). The renatured gel slice containing P3 and P6 interacted with O4 and O6 specifically (Figs. 16-18, Table 2). Therefore, five of the eight major species are likely to belong to the first category of purified protein. P4 was not renatured for its DNA-binding activity (Fig. 18, Table 2). P7 and P8 were not tested by renaturation. Thus, which category these three proteins belong to is unclear. The available partial protein sequences will be helpful to isolate cDNAs for the five proteins (P3, P4, P6, P7, P8). Studies of the recombinant proteins will eventually determine whether or not they are specific DNA-binding proteins.

3) FTZ-F1, Adf1 and P1 are potential trans-activators in the regulation of *ftz* gene expression

Both FTZ-F1 and Adf1 were shown to be involved in the formation of fEBC8 with three footprint sites (6, 8, 9) (Figs. 6, 7, 8, 29A). Therefore, fEBC8 contained at least two proteins: FTZ-F1 and Adf1. The similar overall expression patterns of the two proteins in early embryogenesis (England et al. 1992; Ohno and Petkovich 1992) further suggests that they might interact with each other, forming a heterodimer. Heterodimerization between the nuclear hormone receptors and transcription factors of different families has been examined. For example, the glucocoticoid receptor (GR) was proposed to interact directly with Fos or Jun from the

results of cross-linking and coimmunoprecipitation experiments (Diamond et al. 1990; Yang-Yen et al. 1990). Interaction between FTZ-F1 and Adf1 will be studied in the future.

Detection of fEBC8 in the 16-22 hr nuclear extract (Fig. 10D) suggests that FTZ-F1 and Adf1 also have other functions. FTZ-F1 is expressed in the proventriculus and anal plate in later staged embryos (Ohno and Petkovich 1992). In addition, it also bound to the distal enhancer of *alcohol dehydrogenase* gene (*Adh*) (Ayer and Benyajati 1992). Other DNA-binding sites for the *E.coli* recombinant Adf-1 were found in the promoters of *Antp* P1 and *dopa decarboxylase* (England et al. 1990). It seems that both proteins have multiple functions in embryogenesis.

P1 was purified from embryo nuclear extracts (Fig. 16). Its cDNA was isolated by screening an embryo cDNA library. It bound to DNA sequences specifically in *in vitro* assays (Fig. 25). It also activated gene expression *in vivo* in a binding site dependent fashion (Fig. 26). A region of 73 amino acids of P1 shows high homology to the ribosomal protein L9 (RL9) (Fig. 22). The sequence conservation indicates a functional conservation for this region. Since P1 does not contain any previously identified DNA-binding domains, we propose that the 73 aa region of P1 might be a new class of nucleic acid binding domain. The functional dissection of P1 and examination of the DNA-binding activities of the RL9 proteins

are necessary to confirm if that region is a DNA-binding domain.

Since FTZ-F1, Adf1 and P1 all bound to these three sites, we propose that all three proteins are important in the regulation of the *ftz* gene expression. The similar overall expression patterns of the three genes in the early staged embryos also suggests their roles in the regulation of *ftz* expression. However, further analysis of mutations in these genes is necessary to determine their *in vivo* functions.

4) <u>The mechanism underlying the autoregulatory function of the *ftz* proximal enhancer</u>

It has been proposed that FTZ protein has an autoregulatory function that maintains high level gene expression after the initiation of *ftz* striped expression at low levels (Hiromi and Gehring 1987) (Yu and Pick, unpublished). The cis-element mediating FTZ autoregulation was identified as the *ftz* upstream element (the *ftz* enhancer) (Hiromi and Gehring 1987). The shortest cis-element shown to mediate autoregulation was a 531 bp fragment in the *ftz* proximal enhancer (Pick et al. 1990). Multiple FTZ binding sites were identified in this region (Pick et al. 1990). It was postulated that the FTZ protein and germlayer specific transcription factors interact with the *ftz* proximal enhancer to mediate *ftz* autoregulation (Pick et al. 1990).

Our results further revealed the molecular basis for the *ftz* autoregulation. The fragment of the proximal enhancer was narrowed down to 323 bp which contained nine nuclear factor binding sites and five previously identified FTZ binding sites (Figs. 7, 30). The autoregulatory function of this fragment might contain functionally equivalent protein binding sequences. We have shown that footprint sites 6 and 9 are identical sequences interacting with the same set of nuclear proteins, including FTZ-F1, Adf1, P1, TTK, and fEBC9. Footprint site 8 interacted with the same set of proteins except for fEBC9. There was a dramatic decrease in reporter gene expression as the number of mutant sites was increased from one to three (Figs. 13, 14). Therefore, we favor a mechanism that the same set of proteins binds to those neighboring site cooperatively to increase their binding affinities. For example, Adf1 bound to those sites with at least 10-fold lower affinity than it bound to the Adh promoter sequences (Fig. 28D). Cooperative binding may occur between any two of the three sites. This seems to be sufficient to maintain the fusion gene expression, because mutation of individual sites had no affect on the gene expression (Fig. 13). Consistent with, Adf1 bound to a footprint site 6 dimmer much stronger than it bound to the monomer site 6 (data not shown).

A 700 bp fragment upstream from the transcription start site of a pair-rule gene *even-skipped* (*eve*) was identified as the *eve* autoregulatory element (Goto et al. 1989; Harding et

al. 1989). Multimerizing a 100 bp sequence from this element restored full *eve* autoregulatory activity (Jiang et al. 1991). This observation suggests that multiple functionally equivalent sequences may exist in the 700 bp *eve* autoregulatory element. We have identified at least three redundant sequences which exist endogenously in the *ftz* proximal enhancer. Other functional sequences in the 700 bp *eve* autoregulatory element, except the 100 bp fragment, have not been determined. Our observation of these three tandem repeated sequences in the *ftz* gene structure is a reasonable justification for multimerizing regulatory sequences in the transcriptional studies (Topol et al. 1991; Jiang and Levine 1993; Zeng et al. 1994). Multimerization of cis-elements often increases reporter gene expression levels, but does not affect expression patterns.

Five previously identified FTZ binding sites are localized in the 323 bp fragment of the *ftz* proximal enhancer (Pick et al. 1990). One high-affinity site overlaps with footprint site 2. We have attempted to determine whether or not native FTZ protein interacted with this site in the gel retardation assays with anti-FTZ monoclonal and polyclonal antibodies. However, neither fEBC1 nor fEBC2 interacting with site 2 was affected in the experiments (data not shown). We do not know whether or not FTZ protein is involved in the formation of fEBC1 and fEBC2, though FTZ protein is present in the nuclear extracts, as suggested by Western blot (data not shown).

At least three footprint sites (6, 8, 9) are essential for *ftz* autoregulation mediated by the proximal enhancer (Figs. 13, 14). Site-specific mutation of those sites did not change any of the five FTZ binding sequences. Sites 8 and 9 partially overlap with one medium and one low-affinity FTZ binding site (Fig. 30A). Mutation of site 9 did not affect gene expression (Fig. 13). Additional mutation of site 6 which did not overlap with any of the FTZ binding sites dramatically decreased the gene expression (Figs. 13, 14). These results indicate that nuclear factors FTZ-F1, Adf1, P1 and fEBC9 which bind to those sites are crucial for the *ftz* autoregulatory function. The importance of those factors does not rule out the essential role of FTZ protein in the autoregulation, since the reporter gene was not activated in the *ftz* interstripe regions (Fig. 30C). Therefore, FTZ protein together with other nuclear factors are involved in *ftz* autoregulation.

Deletion analysis of the FTZ binding sites in the proximal enhancer has been reported (Schier and Gehring 1992). Three of the four deletion constructs also deleted part of our footprint sites. A 78 bp deletion, namely ΔA, removed two medium-affinity and one high affinity FTZ binding sites, along with footprint sites 1 and 2. Another 24 bp deletion, namely ΔC, removed a medium-affinity FTZ binding site, along with our footprint site 8. The three mutations, when tested individually, did not significantly affect the enhancer function. Double mutations of ΔA and ΔC dramatically

reduced the enhancer activity. The double mutation changed three of the five FTZ binding site, along with footprint sites 1, 2 and 8. This catastrophic decrease of the gene expression suggests that FTZ protein, fEBC11, FTZ-F1 and Adf1 act cooperatively to maintain the *ftz* expression (Fig. 30A). Cooperative interactions among these proteins might allow binding of FTZ protein to the proximal enhancer with high affinity.

Similarly, the *eve* autoregulatory element was shown to interact with at least three proteins: EVE, DENF1 and DENF2 (Jiang et al. 1991). Mutations of three EVE binding sites, or DENF1 and DENF2 binding sites reduced reporter gene expression (Jiang et al. 1991). Based on the DENF2 binding sequences, DENF2 was proposed to be TTK (Jiang et al. 1991). Although the TTK binding consensus sequence contains AGGA (Brown et al. 1991), AGGA containing sequences might interact with more proteins than TTK or may not interact with TTK at all. We have shown that only five of the nine AGGA containing footprint sites interacted with TTK, and that three of these five TTK binding sites (site 6, 8, 9) interacted with multiple proteins in addition to TTK (Figs. 7, 9). Therefore, it is hard to predict identities of nuclear factors based upon their binding sites. The proteins identified in our studies may also interact with these sites in the *eve* autoregulatory element.

A 120 bp element (module E) in the *Deformed* (Dfd) autoregulatory enhancer contained one Dfd binding site and an

imperfect inverted repeat sequence (Regulski et al. 1991; Zeng et al. 1994). Each sequences is necessary for the autoregulatory function of the module E. A putative cofactor was proposed to cooperate with the Dfd protein to bind to module E (Zeng et al. 1994).

5) <u>The potential function of the *ftz* proximal enhancer in establishing the *ftz* stripe expression</u>

Of the three *ftz* cis-elements, the zebra element, neurogenic element and upstream element (Hiromi et al. 1985), the zebra element directs gene expression in the mesoderm only (Hiromi et al. 1985). The establishment of striped expression in both germlayers requires both the zebra element and the upstream element (Hiromi et al. 1985; Hiromi and Gehring 1987), which is independent of FTZ protein (Hiromi and Gehring 1987). Quantitatively the upstream element is also required for high level expression of the *ftz* gene (Hiromi and Gehring 1987; Dearof et al. 1989a). The *ftz* upstream element contains two independent enhancers (Pick et al. 1990). Only the *ftz* proximal enhancer can direct reporter gene expression in both ectoderm and mesoderm (Pick et al. 1990), as does the endogenous *ftz* gene.

Nuclear proteins with potential transcriptional activation function have been identified in the *ftz* proximal enhancer (Fig. 30A). In addition to their potential functions involved in *ftz* autoregulation, they are also potential transcription activators in the initiation of the *ftz* striped

expression. Mutation of fEBC3, fEBC4 and fEBC11 binding sites individually did not affect the enhancer function (Fig. 13). These three fEBCs were detected in as early as 0-2 hr nuclear extract (Fig. 10). Thus, these fEBCs might be involved in the initiation of *ftz* expression. Other factors (FTZ-F1, Adf1, P1, fEBC9) which appear to be involved in *ftz* autoregulation as discussed above, might be also involved in the establishment of the *ftz* seven stripes, since the genes encoding FTZ-F1, Adf1 and P1 are expressed very early in the syncytium blastoderm.

6) Transcriptional repression function of the *ftz* proximal enhancer

The *ftz* expression is maintained via a positive autoregulatory feedback loop. This autoregulation is mediated by the ftz enhancers including the proximal enhancer (Pick et al. 1990). However, the *ftz* mRNA stripes decay very quickly before end of the germband extension. The rapid disappearance of the *ftz* mRNA suggests that the *ftz* transcription maintained by autoregulation is actively turned off. A molecular mechanism for this repression of *ftz* expression is suggested by our studies.

Mutation analysis of the nine footprint sites did not detect any miss-expression or prolonged expression of the *ftz-lacZ* fusion gene (Figs. 13, 14). This result would appear to suggest that no footprint site mediates repression of gene expression outside of the seven stripes. However, analysis of

fEBCs with stage-specific nuclear extracts indicated that the same footprint site could interact with both potential transcriptional activators and repressors (Table 1, Fig. 10). Mutation of those sites would therefore abolish the activation function of the proximal enhancer as discussed above. This would preclude the possibility of detecting prolonged expression of the reporter gene in this analysis. fEBC1, C2, C7, C10 and C12 were proposed to be potential repressors of *ftz* gene expression. It is quite likely that the transcriptional repressors competitively replace the activators by binding to the same sites in *ftz* proximal enhancer to repress autoregulation in an active fashion (Fig. 30B).

This mechanism has been suggested for the regulation of *eve* stripe 2 expression. Three cis-elements, a stripes 2 and 7 element, a stripe 3 element and an autoregulatory element, were identified for the pair-rule gene *even-skipped* (*eve*) (Goto et al. 1989; Harding et al. 1989). The maternal and gap gene products of *bicoid* (*bcd*), *hunchback* (*hb*), *giant* (*gt*) and *Kruppel* (*Kr*) have been shown to directly interact with the stripe 2 element in footprint assays with *E. coli*-synthesized proteins (Small et al. 1991). It has been proposed that *eve* stripe 2 is activated by the *bcd* and *hb* products, and repressed by the *gt* and *Kr* products at its anterior and posterior borders. The repression is perhaps mediated by the competition of *gt* and *Kr* proteins with the *bcd* and *hb*

products for binding to overlapping or neighboring sites in the stripe 2 element (Small et al. 1991).

This competition mechanism underlying the repression of the *ftz* autoregulation is strongly supported by the analysis of the footprint sites 6, 8, 9. First, FTZ-F1, Adf1 and P1 interacted repeatedly with these sites. fEBC8 (consisting of FTZ-F1 and Adf1) was detected in the nuclear extract from 0-22 hr embryos, including the time when the *ftz* gene is expressed. Second, TTK interacted repeatedly with the same sites. The activity of fEBC10 (TTK) reached the highest level in the nuclear extracts from 4-9 hr embryos, when the *ftz* gene is turned off. Third, mutations of the three sites abolished gene expression. Finally expression of TTK earlier in the embryos significantly decreased the *ftz* expression (Read et al. 1992; Brown and Wu 1993). These observations indicate that TTK replaces FTZ-F1 and Adf1 binding to the three essential sites to repress the *ftz* gene expression actively. Lack of the detectable changing of the *ftz* expression in the *ttk* null mutant embryos (Xiong and Montell 1993) suggests that TTK itself is not sufficient to stop the *ftz* autoregulatory feedback loop. Other potential repressors are fEBC1, C2, C7 and C12, as discussed above.

CHAPTER 6. FIGURES AND FIGURE LEGENDS

List of Tables

Table 1. Stage specificity of the *ftz* enhancer binding complexes (fEBCs)

Different staged nuclear extracts were incubated with four footprint sites (O2, O3, O4, O6) in the *ftz* proximal enhancer in gel retardation assays. The protein-DNA complexes (fEBCs) were detected in different staged nuclear extracts interacting with different probes as summarized in the Table. The temporal activities of fEBCs are compared with the time of *ftz* expression.

Probes	fEBCs most active in	
	0-4 hr NE	4-9 hr NE
O2	fEBC11	fEBC1, 2
O3	fEBC3	fEBC12
O4	fEBC4	fEBC7
O6	fEBC8, 9	fEBC10
ftz	On	Off

Table 2. Summary of the purified proteins

Twelve proteins were purified from the nuclear extracts of 0-12 hr *Drosophila* embryos by DNA-affinity chromatography (Fig. 15). Reliable amino acid sequences were obtained by microsequencing for eight of them. These eight proteins are summarized in the Table.

Proteins	kDa	Purified with oligo-nucleotide	Renaturation	cDNA	Potential functions	References
P1 (fEBP1)	29	O6	Yes	*P-1*		this work
P2 (fEBP2)	34	O6	No	*Adf-1*	Activator of *Adh*	England et al. this work
P3	52	O4, O6	Yes			this work
P4	54	O4	No			this work
P5 (fEBP5)	45	O4, O6	No	*ttk*	Repressor of *ftz*	Harrison et al. Brown et al. this work
P6	40	O4	Yes			this work
P7	51	O6	No			this work
P8	33	O6	No			this work

List of Figures

Figure 1. Preparation of the *ftz* proximal enhancer fragments

The ~6 kb 5'-flanking sequence of the *ftz* gene containing three cis-elements is shown on the top (Hiromi and Gehring 1987). The distal and proximal enhancer are derived from the upstream element. The ~1 kb proximal enhancer, an XbaI fragment from 1502 to 2574 position, was cloned into a plasmid construct 5'D (Pick et al. 1990). Subclones derived from 5'D are indicated at the right of their respective containing fragments. Seven fragments were indicated from F1 to F7. (X) XbaI; (B) BglII; (A) AvaII; (H) HincII; (E) EcoRV.

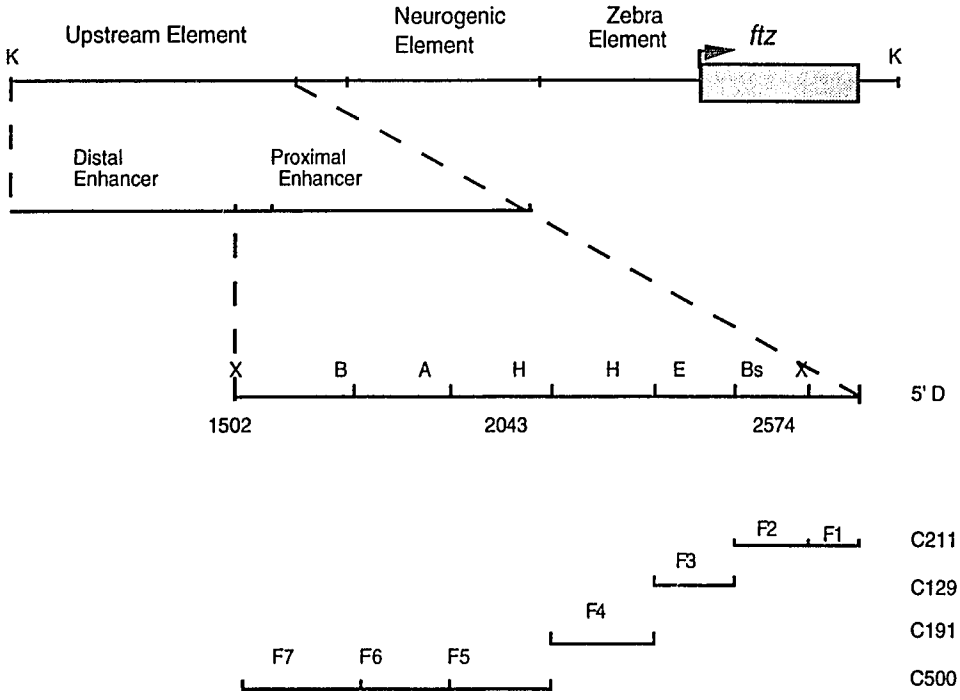


Figure 2. Interactions of nuclear proteins with the *ftz* proximal enhancer fragments

Nuclear extracts from 0-12 hr *Drosophila* embryos (NE) were incubated with the *ftz* proximal enhancer fragments (F2, F3, F4) in gel retardation assays. **Panel A:** Gel retardation assays with F4 as probe. 4 µg NE was incubated with 10 fmole ^{32}P-labeled F4 and 4 µg poly(dI-dC) for 1 hr at 4°C. The competitors indicated above the lanes were added to the binding reactions at a 50-fold molar excess over F4 probe. The reactions were electrophoresed through a native 4% polyacrylamide gel. (I) Protein-F4 complex I; (F4) Free probe F4. Lane 1, no NE; Lane 2, no competitor; Lanes 3-8, *ftz* proximal enhancer fragments F2-F7; Lane 9, a 230 bp AvaII fragment from Bluescript KS+. **Panel B:** Gel retardation assays with F3 as probe. Reactions were carried out as above, but with 6 µg NE and 6 µg poly(dI-dC). (F3) free probe F3; (II) protein-F3 complex II. Lane 1, no NE; Lane 2, no competitor; Lanes 3-10, the competitors shown in Panel A. **Panel C:** Gel retardation assays with F2 as probe. Reactions were carried out as above, with 2 µg NE and 2 µg poly(dI-dC). (F2) free probe F2; (III) protein-F2 complex III. Lane 1, no competitor; Lanes 2-8, competitors shown above the lanes; Lane 9, no NE. **Panel D:** Titration of competitors in gel retardation assays with F3 as probe. Reactions were carried out as in Panel B. Competitors (F3, F2, F4) were used in the reactions at 5-fold, 10-fold and 20-fold molar excess over probe F3 as indicated (lanes 2-9). (F3) free probe F3; (II) protein-F3 complex II. Lane 1, no competitor; Lane 10, no NE.

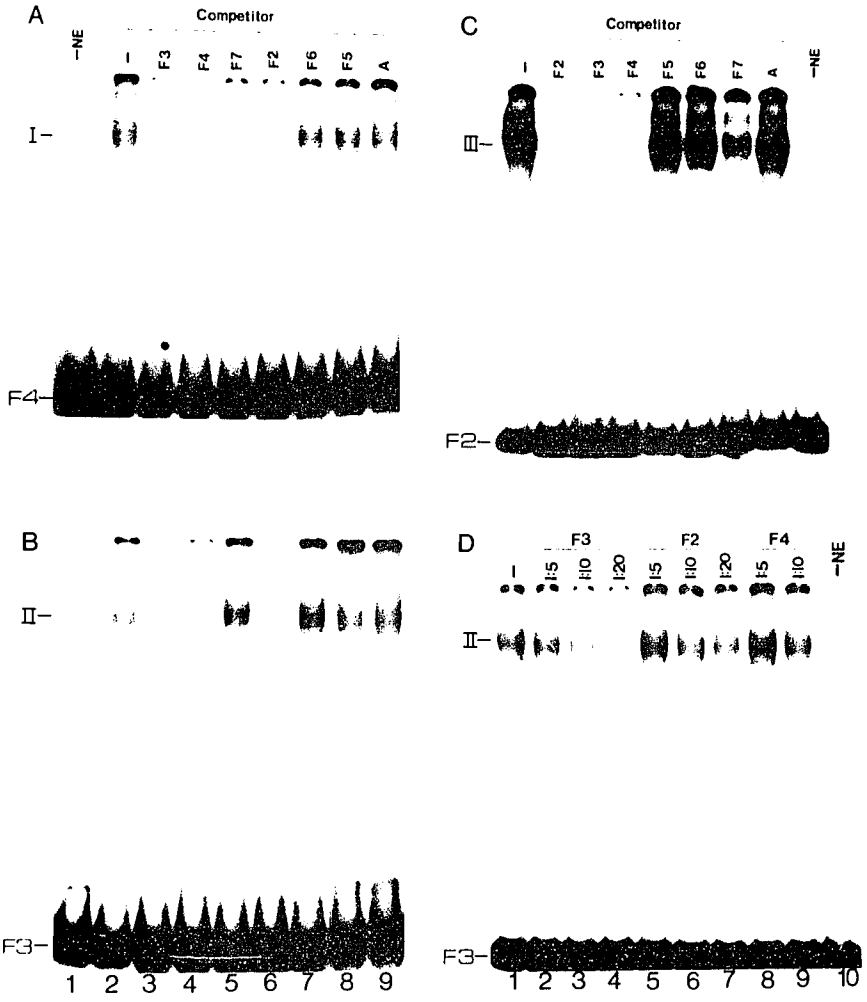

A
Competitor
-NE
-
F3
F4
F7
F2
F6
F5
A
I-
F4-
B
II-
F3-
1 2 3 4 5 6 7 8 9
C
Competitor
-
F2
F3
F4
F5
F6
F7
A
-NE
III-
F2-
D
F3
F2
F4
-
1:5
1:10
1:20
1:5
1:10
1:20
1:5
1:10
-NE
II-
F3-
1 2 3 4 5 6 7 8 9 10

Figure 3. Identification of nine footprint sites in the *ftz* proximal enhancer

The *ftz* enhancer fragments F2, F3 and F4 were ^{32}P-labeled on either the upper (U) or lower (L) strand and used as probes in DNase I footprinting assays. Each probe was incubated with *Drosophila* 0-12 hr embryo nuclear extract (NE) under the conditions described in Figure 2. DNase I digestion and analysis of the digested DNA products are described in Method 3. Autoradiographs of the results are shown for each probe in both strands. The digested DNA fragments recovered from the complexes are shown in Lanes B (Bound), and from the free probes in Lanes F (Free). (G) Guanines of the probes were obtained using the sequencing method of Maxam and Gilbert (Maxam and Gilbert, 1980). The footprint sites are indicated by numbered bars along both strands.

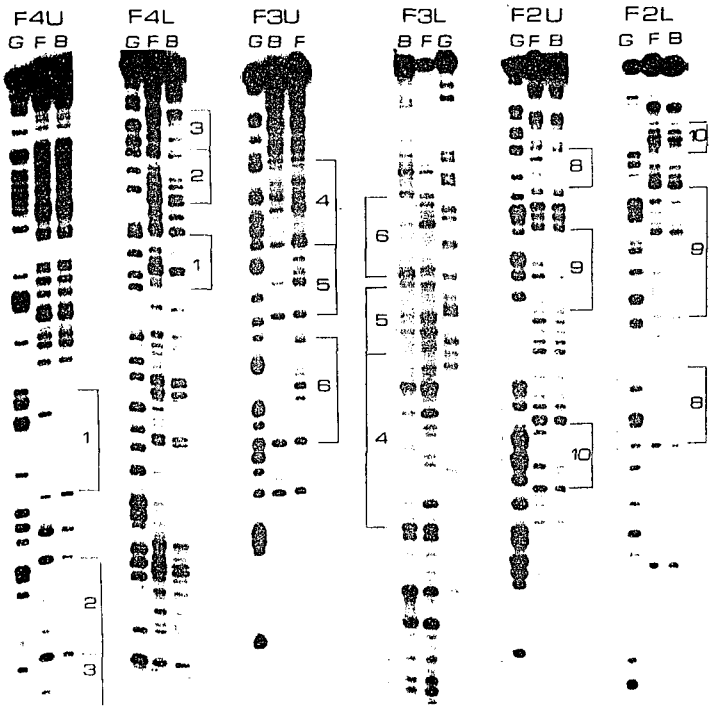


Figure 4. Protein-contact sites in the *ftz* proximal enhancer

DNA fragments F2, F3 and F4 were ^{32}P-labeled on the upper (U) and lower (L) strands. Partially methylated probes were incubated with 0-12 hr *Drosophila* embryo nuclear extracts as described (Method 4). The resulting protein-DNA complexes and free probes were separated through 4% native polyacrylamide gels. After they were recovered from the gel, they were subjected to piperidine digestion. The digested fragments from the protein-DNA complexes (B) and the free probes (F) were separated on 8% sequencing gels. Guanines which, when methylated, interfere with protein binding are indicated with lines and dotted lines. The dotted lines are weak interference sites. The numbers beside the lines indicate the respective footprint sites (Fig. 3) in which the detected guanines are located.

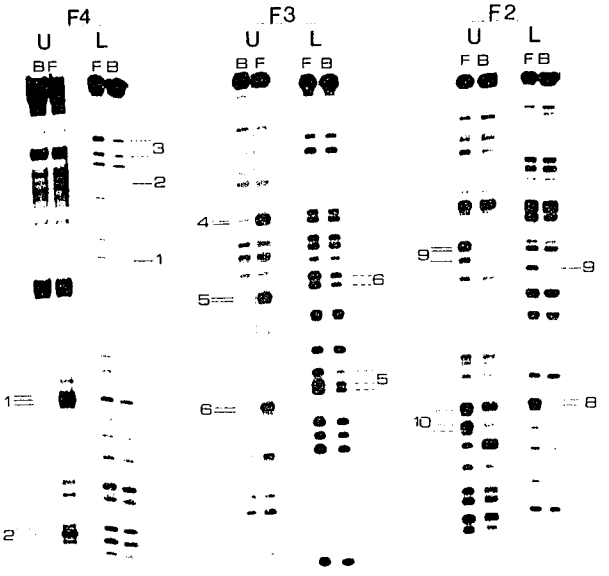


Figure 5. Summary of the nine footprint sites in the *ftz* proximal enhancer

A 400 bp sequence of the *ftz* proximal enhancer from positions 2121 to 2520 is shown (Harrison and Travers 1988). The nine footprint sites are shown in the shaded boxes with numbered in circles above them (1-6 and 8-10). These sites were identified by DNaseI footprinting assays of this region with 0-12 hr embryo nuclear extracts (Fig. 3). Site 7 at positions 2121 to 2140, a randomly chosen nonfootprinted region, was used as a negative control in gel retardation assays (see below). This site is not indicated. The dots indicate the guanines on both DNA strands which were detected by the methylation interference assay. The exact border between footprint sites 4 and 5 was assigned on the basis of results of gel retardation assays (data not shown). Sites 4 and 5 were the only tested oligonucleotides from this region that generated clear protein-DNA complexes.

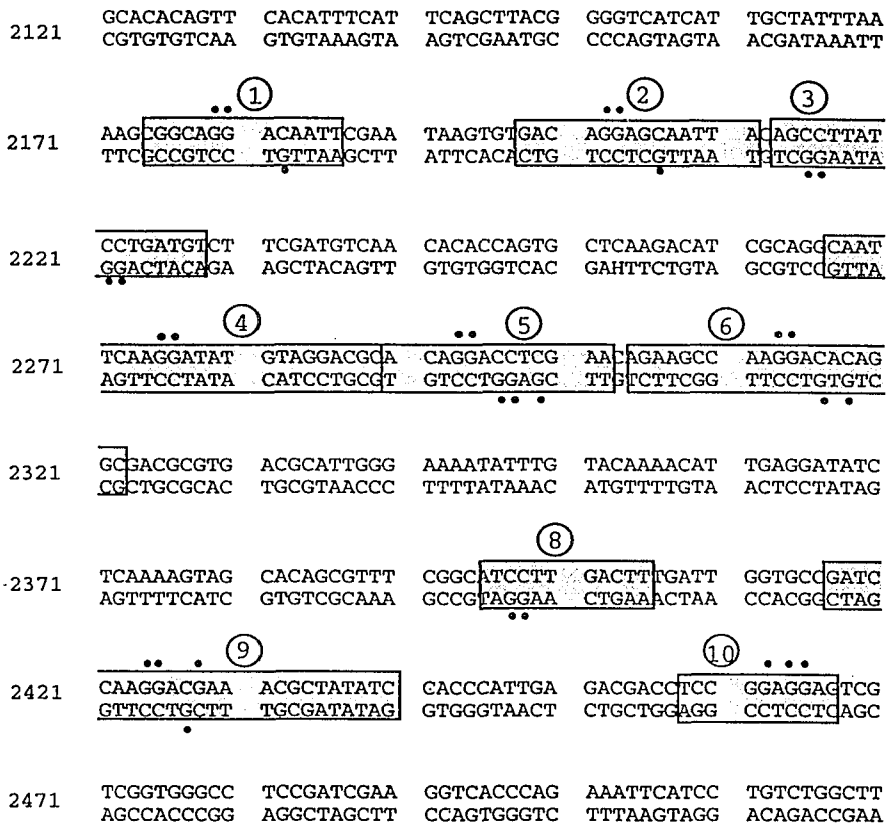

2121
GCACACAGTT CACATTTCAT TCAGCTTACG GGGTCATCAT TGCTATTTAA
CGTGTGTCAA GTGTAAAGTA AGTCGAATGC CCCAGTAGTA ACGATAAATT
2171
AAGCGGCAGG ACAATTCGAA TAAGTGTGAC AGGAGCAATT ACAGCCTTAT
TTCGCCGTCC TGTTAAGCTT ATTCACACTG TCCTCGTTAA TGTCGGAATA
2221
CCTGATGTCT TCGATGTCAA CACACCAGTG CTCAAGACAT CGCAGGCAAT
GGACTACAGA AGCTACAGTT GTGTGGTCAC GAHTTCTGTA GCGTCCGTTA
2271
TCAAGGATAT GTAGGACGCA CAGGACCTCG AACAGAAGCC AAGGACACAG
AGTTCCTATA CATCCTGCGT GTCCTGGAGC TTGTCTTCGG TTCCTGTGTC
2321
GCGACGCGTG ACGCATTGGG AAAATATTTG TACAAAACAT TGAGGATATC
CGCTGCGCAC TGCGTAACCC TTTTATAAAC ATGTTTTGTA ACTCCTATAG
2371
TCAAAAGTAG CACAGCGTTT CGGCATCCTT GACTTTGATT GGTGCCGATC
AGTTTTCATC GTGTCGCAAA GCCGTAGGAA CTGAAACTAA CCACGGCTAG
2421
CAAGGACGAA ACGCTATATC CACCCATTGA GACGACCTCC GGAGGAGTCG
GTTCCTGCTT TGCGATATAG GTGGGTAACT CTGCTGGAGG CCTCCTCAGC
2471
TCGGTGGGCC TCCGATCGAA GGTCACCCAG AAATTCATCC TGTCTGGCTT
AGCCACCCGG AGGCTAGCTT CCAGTGGGTC TTTAAGTAGG ACAGACCGAA

Figure 6. Specific proteins interact with nine footprint sites in the *ftz* proximal enhancer

Nine double-stranded oligonucleotides (O1-O6, O8-O10) were synthesized according to the DNA sequences of the nine footprint sites (Fig. 5). The oligonucleotides were named according to the numbers of their respective footprint sites (Fig. 5). They were ^{32}P-labeled and incubated with 0-12 hr *Drosophila* nuclear extracts (NE) before electrophoresis through 4% native polyacrylamide gels. Four representatives of the nine oligonucleotide gel retardation assays are shown. The same competitors and lane numbers were used for Panel A and B and for Panel C and D. Lane 1 in each panel has no competitor; Lane 10, no NE; Lanes 2-9, a 50 fold molar excess of the indicated unlabeled oligonucleotide was added to the binding reactions as competitor. O45 is an oligonucleotide containing the sequences covering footprint sites 4 and 5; O23 covers sites 2 and 3. O2, O3, O4 and O6 on the left of the Panels indicate the free oligonucleotide probes. *ftz* enhancer binding complexes fEBC1 (C1) to fEBC10 (C10) are ten different protein-oligonucleotide complexes. Two nonspecific complexes were detected in Panel C. One, below fEBC4, is not competed by any oligonucleotide, including O4. The other, between fEBC5 and fEBC6, represents a residual single stranded DNA-binding activity, described in the text. Similarly, a complex between fEBC9 and fEBC10 is also the residual single stranded DNA-binding activity (Panel D).

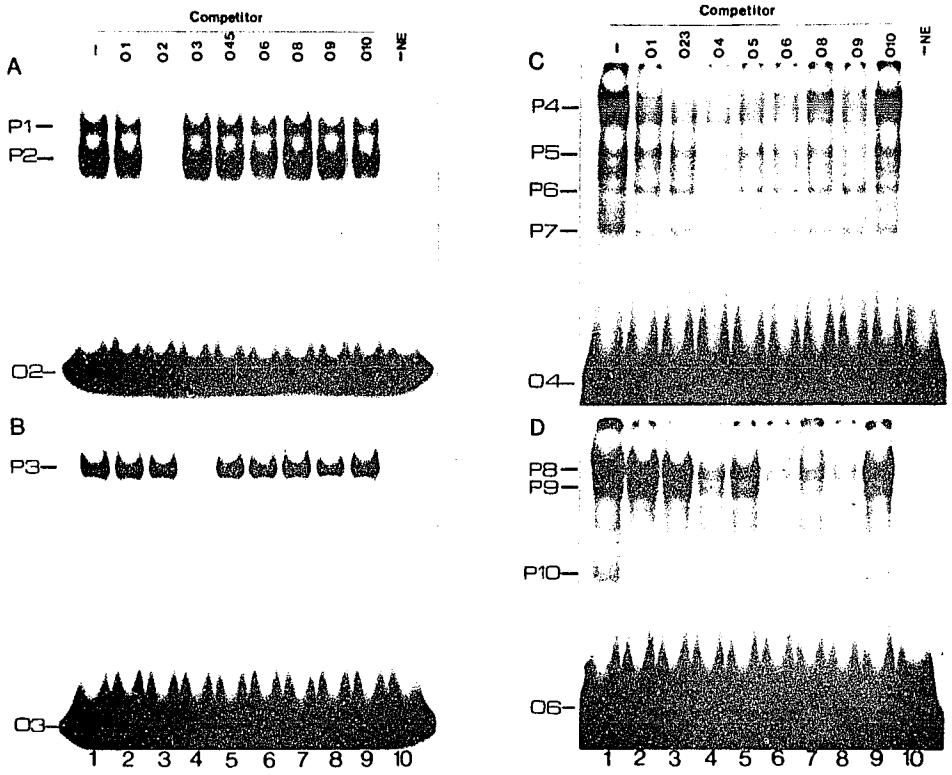

Competitor
A
-
01
02
03
045
06
08
09
010
-NE
P1
P2
O2
B
P3
O3
1 2 3 4 5 6 7 8 9 10
Competitor
C
-
01
023
04
05
06
08
09
010
-NE
P4
P5
P6
P7
O4
D
P8
P9
P10
O6
1 2 3 4 5 6 7 8 9 10

Figure 7. A diagrammatic representation of fEBCs interacting with the *ftz* proximal enhancer

The relative positions of three DNA fragments (F2, F3, F4) in the 300 bp *ftz* proximal enhancer are indicated. The nine footprint sites are represented by nine boxes, and named by 1 to 6 and 8 to 10. The numbers in the boxes indicate the lengths of the footprint sites. Twelve fEBCs (C1 to C12) are grouped below their respective binding sites. No binding activities were detected at footprint site 10 (?) by gel retardation assays. The asterisks indicate that fEBC8 may contain FTZ-F1 (Ueda et al. 1990; Lavorgna et al. 1991) and fEBC10 may contain TTK protein (Harrison and Travers 1990).

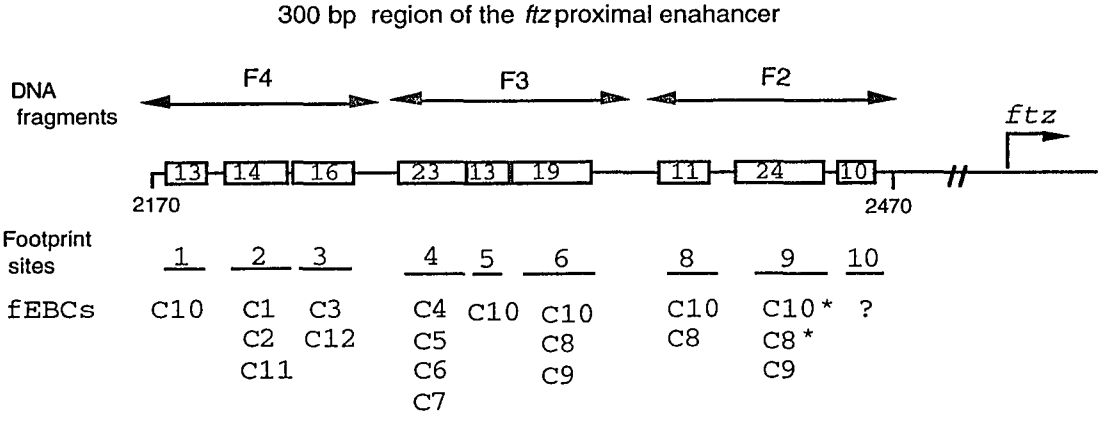


Figure 8. FTZ-F1 is involved in formation of fEBC8

A) Probe O6 was ^{32}P-labeled and incubated with 0-12 hr *Drosophila* nuclear extract (NE) (Lane 1). A 50-fold molar excess of an oligonucleotide containing *FTZ*-F1 binding site I (5'-GTCGACGCAGCACCGTCTCAAGGTCGCCGAGTAGGAGAA-3') (Ueda et al. 1990) specifically competes fEBC8 (C8) and fEBC9 (C9) binding to O6 in gel retardation assays (lane 2).

B) Gel retardation assays with probes O6 and O8 were carried out (lanes 1, 4) in the presence of 2 µl anti-FTZ-F1 antibody (α-FF1; lanes 2, 5, 7) or 2 µl preimmune serum (PiS; lanes 3, 6, 8). Anti-FTZ-F1 antibody, but not preimmune serum, specifically abolished fEBC8 binding to O6 and O8. fEBC9 was not affected by addition of antibodies to gel retardation assays with O6 as probe. -NE in lanes 7, 8, no nuclear extract added. fEBC8 (C8), fEBC9 (C9) and fEBC10 (C10) refer to both Panels.

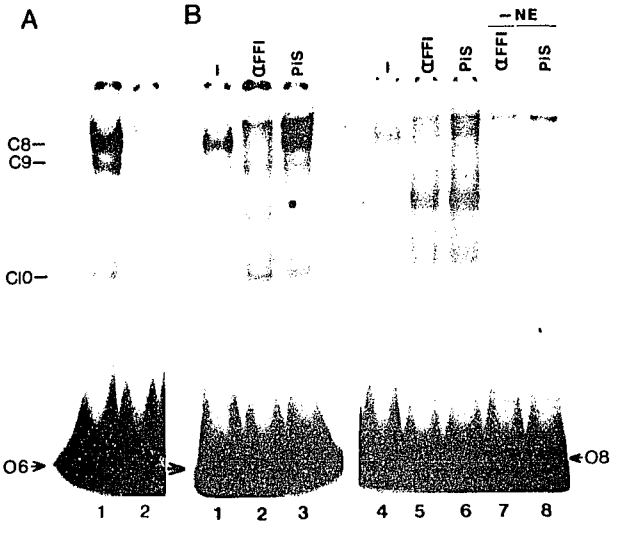


Figure 9. TTK is involved in formation of fEBC10

A) Probe O5 was used in the gel retardation assays with 0-12 hr nuclear extract. fEBC10 (C10) was detected (lane -). Addition of 2 µl of anti-TTK antibody (αTTK) abolished fEBC10 (C10) binding to O5. Addition of 2 µl preimmune serum (PiS) did not affect fEBC10 binding.

B) Probe O1 was incubated with: 0.2 µg bacterially expressed TTK (TTK); 4 µg of 0-12 hr *Drosophila* nuclear extract (NE). Similar migrating protein-DNA complexes were detected in the gel retardation assay.

C) Bacterially expressed TTK (0.2 µg) was incubated with ^{32}P-labeled wild type O6, O5 and O1 or corresponding mutant oligonucleotides M6, M5 and M1 (for mutant sequences, see Fig. 12) as indicated. TTK bound to the wild type oligonucleotides (O6, O5, O1) but not the mutant forms (M6, M5, M1) in gel retardation assays.

D) 4 µg nuclear extracts (NE) was incubated with probes O5 and M5. fEBC10 (C10) was detected in gel retardation assays with O5, but not with M5.

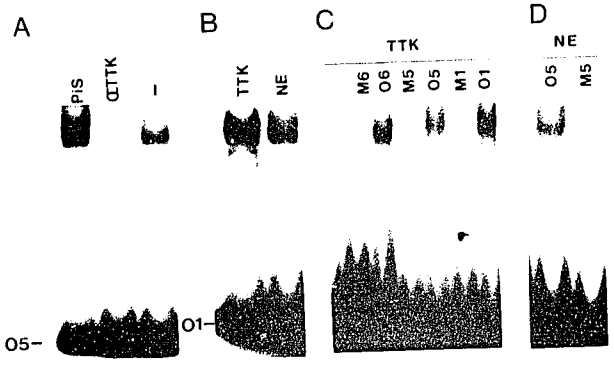


Figure 10. Stage specificity of fEBCs in the *Drosophila* embryos

Nuclear extracts were prepared from 0-2 hr, 2-4 hr, 4-6 hr, 6-9 hr, 16-22 hr, and 0-16 hr staged *Drosophila* embryos. Nuclear extracts were incubated with four oligonucleotide probes in gel retardation assays (Method 2). The extract used is indicated on top of the lanes. The results of gel retardation assays with O2, O3, O4 and O6 are shown in Panel A, B, C and D. Shifted protein-DNA complexes are indicated on the left of the Panels as fEBC1 to fEBC12. Different DNA-binding activities (twelve fEBCs) were detected in different staged nuclear extracts.

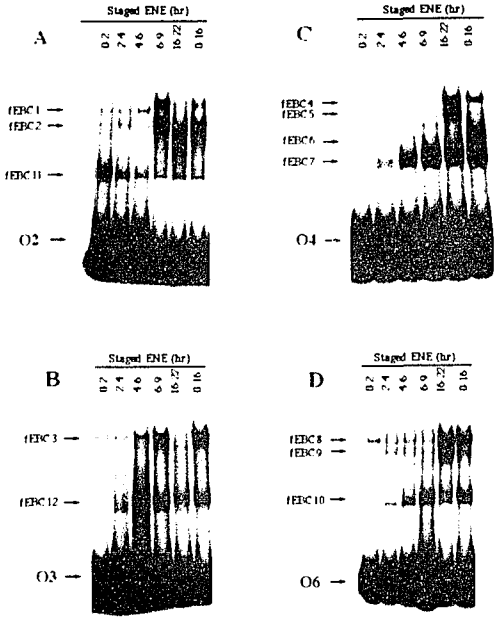


Figure 11. Site-specific mutagenesis by the method of overlap extension using PCR

The experimental design is described in detail in Method 8.

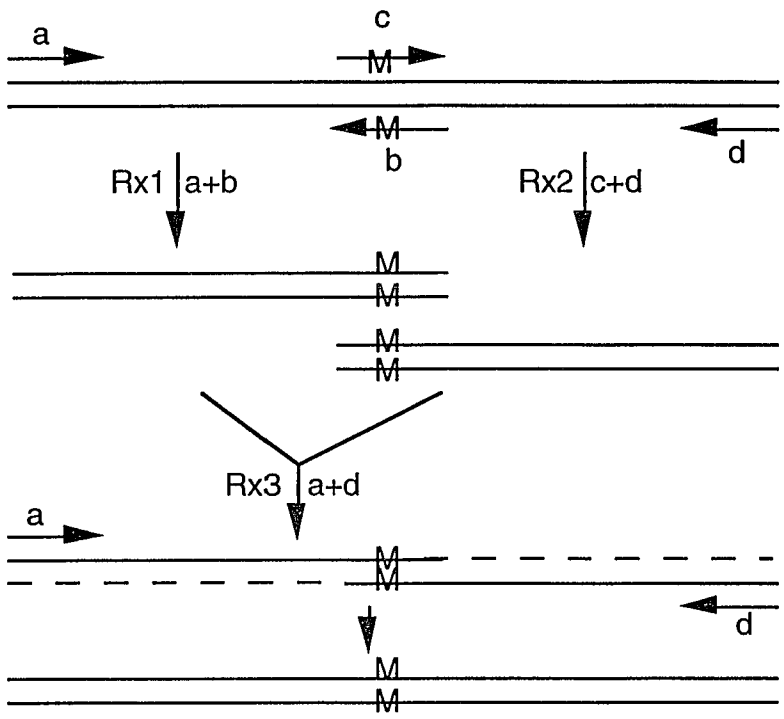


Figure 12. Site-specific mutagenesis of nine footprint sites in the *ftz* proximal enhancer

Site-specific substitution mutations were designed to abolish nuclear protein binding to the footprint sites in the *ftz* proximal enhancer. Guanidines in the footprint sites appear to make direct protein contacts were detected in methylation interference assays (Figs. 4, 5). Therefore, those guanidines were primary targets of mutations. The mutant sequences are indicated on top of the corresponding wild type sequences. Nine pairs of mutant oligonucleotides were synthesized and named M1 to M6, M8 to M10 according to the numbers of the footprint sites.

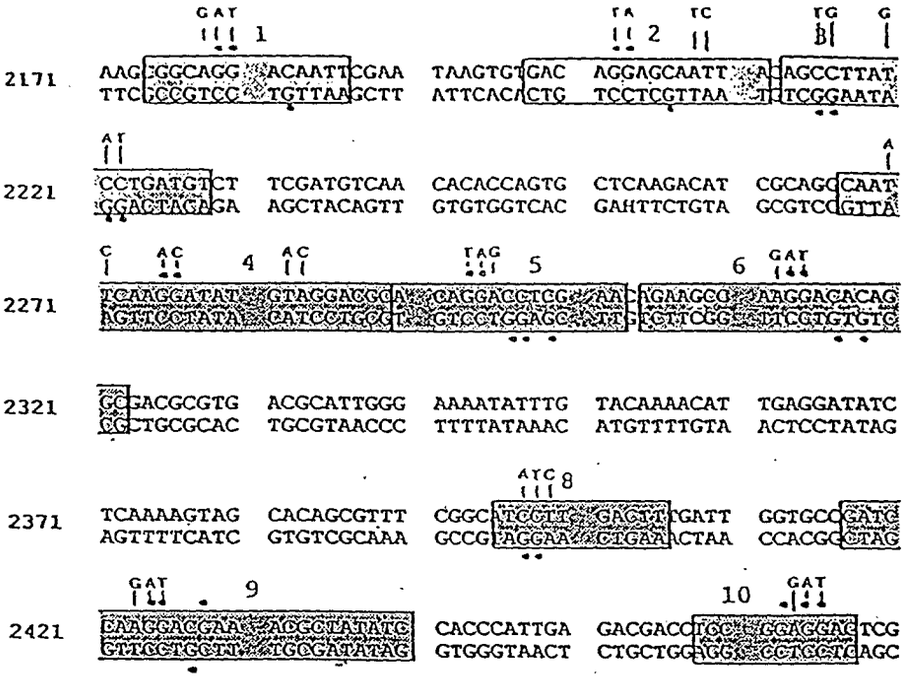


Figure 13. Summary of the *ftz* proximal enhancer-*lacZ* fusion genes

A 406 bp DNA fragment from the *ftz* proximal enhancer (at positions of 2168-2574 bp) contains all nine footprint sites. The nine sites are represented by nine boxes, and numbered above them. Different fillings in the boxes indicated different binding proteins (Fig. 7). This fragment (Prox 406) was narrowed down from its 3'-end to 2491 (Prox 323). Mutant *ftz* proximal enhancers were constructed containing mutations in different footprint sites in fragment Prox 406 (Method 8). The mutated site is indicated by the number (e.g. M1 contains a mutation in site 1, see Fig. 17). Mutations are indicated by squares filled with crosses. The different *ftz* proximal enhancers were fused to an *hsp70* minimal promoter and the reporter gene *lacZ* in HZ50PL vectors (Hiromi and Gehring 1987). All fifteen DNA constructs are in the forward orientation. After the fusion genes were introduced into the embryo genome by P-element mediated transformations, their expression patterns were monitored with immunohistochemical staining using anti-β-galactosidase antibodies (Method 9). Numbers of established transformant lines for each DNA construct are indicated in parentheses. Strong β-galactosidase expression in the seven *ftz*-like stripes is indicated with ++; weak expression, with +-; no expression, with -.

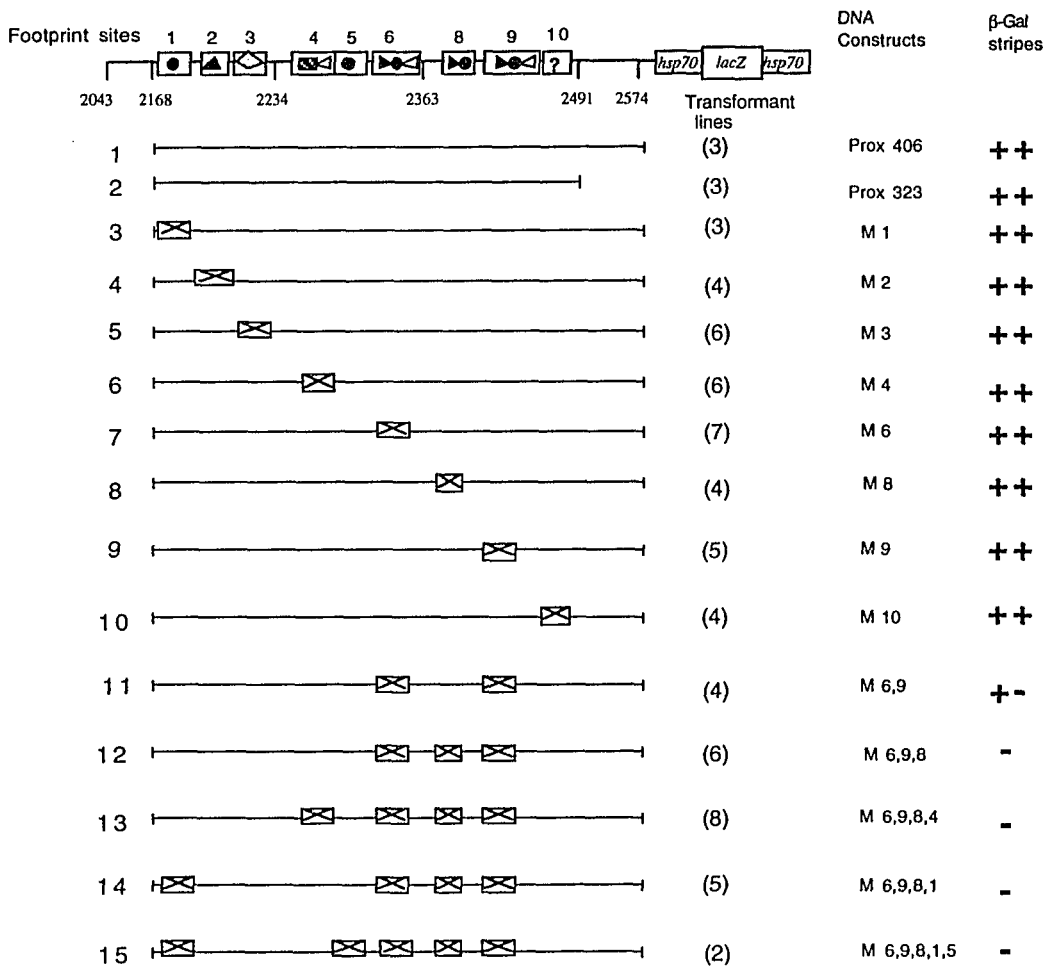

Footprint sites
1
2
3
4
5
6
8
9
10
?
hsp70
lacZ
hsp70
2043
2168
2234
2363
2491
2574
Transformant lines
DNA Constructs
β-Gal stripes
1 (3) Prox 406 ++
2 (3) Prox 323 ++
3 (3) M 1 ++
4 (4) M 2 ++
5 (6) M 3 ++
6 (6) M 4 ++
7 (7) M 6 ++
8 (4) M 8 ++
9 (5) M 9 ++
10 (4) M 10 ++
11 (4) M 6,9 +-
12 (6) M 6,9,8 -
13 (8) M 6,9,8,4 -
14 (5) M 6,9,8,1 -
15 (2) M 6,9,8,1,5 -

Figure 14. Expression of selected *ftz* proximal enhancer-*lacZ* fusion genes

The expression of *ftz* proximal enhancer-*lacZ* fusion genes was monitored immunohistochemically in transformant embryos using anti-β-galactosidase antibodies. Representative embryos carrying four different fusion genes are shown. They are termed under their transforming DNA constructs (Fig. 13). Embryos are oriented: anterior, left; dorsal side, up. Positions of the seven *ftz*-like strips are indicated in Panel A. Fusion genes Prox 323 (Panel A) and M10 (Panel B) are expressed in seven *ftz*-like stripes. Fusion gene M6,9 (Panel C) is expressed very weakly in six stripes (no stripe 1 is detectable). No expression of M6,9,8 (Panel D) was detected.

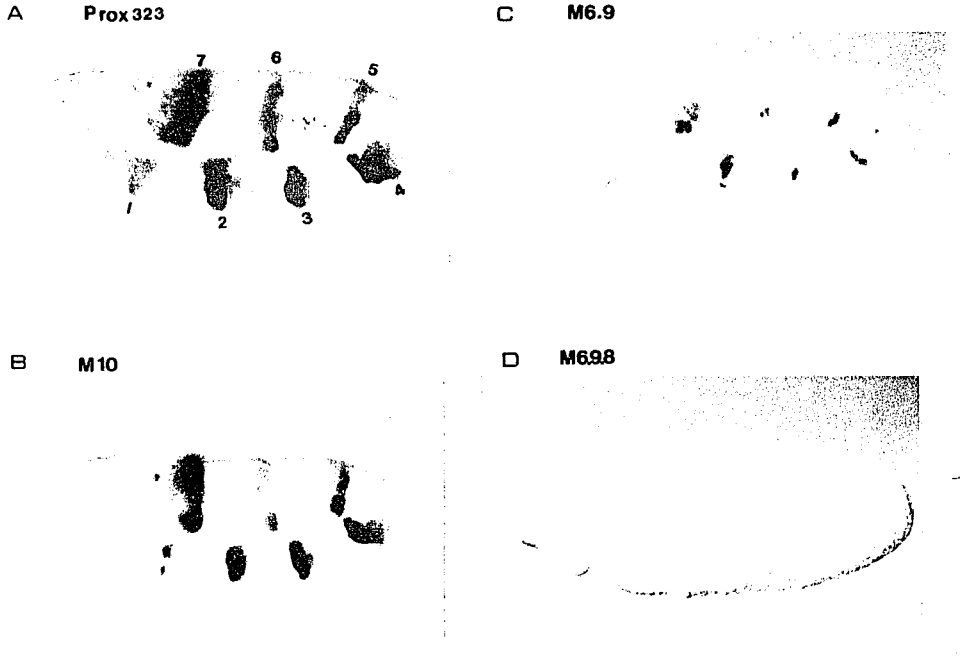
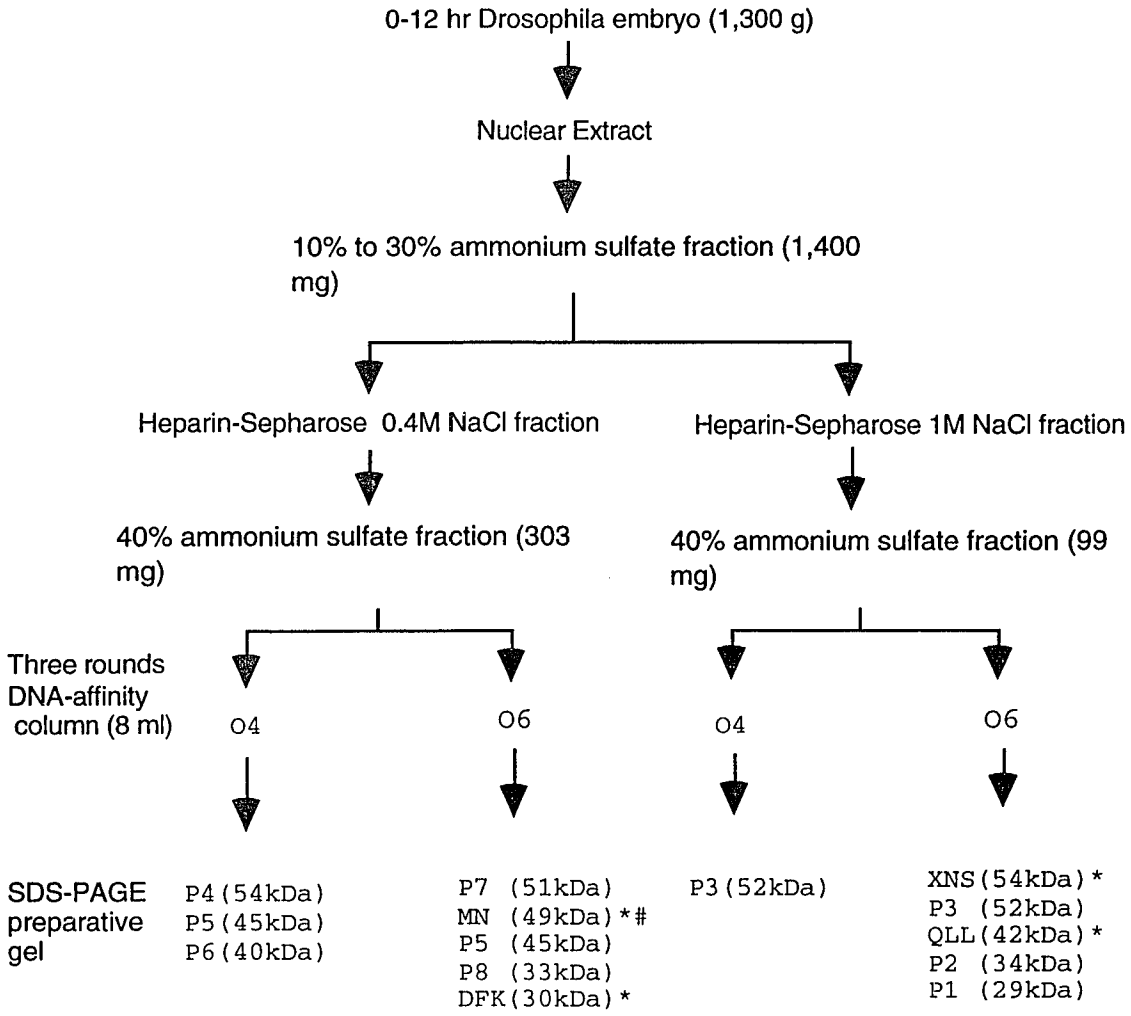


Figure 15. A summary of the purification of potential *ftz* enhancer binding proteins

Nuclear extract was prepared from 0-12 hr *Drosophila* embryos. It was purified by a Heparin-sepharose column. Proteins in two Heparin fractions were further purified by O4 and O6 DNA-affinity chromatography. Three-rounds of purification with the two columns were performed as described in Result 7. Twelve purified polypeptides are indicated with P1 to P8, XNS, QLL, MN and DFK.



*Minor bands in SDS gel
Peptide sequences are not reliable.

Not sequenced.

Figure 16. Purification of P1, P2, and P3 from the 1 M Heparin fraction using O6-affinity chromatography

A) Proteins from each purification step were electrophoresed through a 10% SDS polyacrylamide gel. A photograph of the silver stained gel is shown. NE, 10 µg; Hep 1M, 10 µg 1 M Heparin fraction; Affinity 1, 20 µl; Affinity 2, 50 µl; Affinity 3, 100 µl. M, protein markers. P1, P2 and P3 were enriched. The dark staining of the bottom of lane 5 is from the protease inhibitors used in each purification step. Proteins in regions 1, 2, 3, and 4 were extracted, denatured and renatured (Method 7).

B) P1, P2 and P3 were recovered from a preparative gel and digested with V8 protease (Method 6). Resulting peptides were transferred onto a PVDF membrane and stained with Coomassie-Brilliant Blue. Arrows indicate the peptides which were partially sequenced.

C) Renatured proteins (8 µl) from region 1 (R1), 2 (R2), 3 (R3), 4 (R4) (indicated in Panel A) were used in gel retardation assays with probes O2 and O6 as indicated on top of the lanes. Proteins from R3 and R4 interacted with O6 strongly (lanes 3 ,5), with O2 very weakly (lanes 4, 6). No DNA-binding activities were detected with proteins from R1 and R2 (lanes 1, 2).

D) To monitor the purification steps, gel retardation assays with O6 were performed with different protein samples and different amounts of poly(dI-dC). Lane 1, 4 µg nuclear extract and 2 µg poly(dI-dC); Lane 2, 2 µg 1 M Heparin fraction and 0.2 µg poly(dI-dC); Lane 3, 2 µl O6 first-eluate and 5 ng poly(dI-dC); Lane 4, 2 µl O6 second-eluate and 1 ng poly(dI-dC); Lanes 5-7, 2 µl O6 third-eluate and no poly(dI-dC). A 50-fold molar excess of unlabeled O6 and O10 were used in lanes 6 and 7 respectively. Two specific DNA-binding complexes were purified. One is between fEBC9 (C9) and fEBC10 (C10). The other is below fEBC10 (for details, see text).

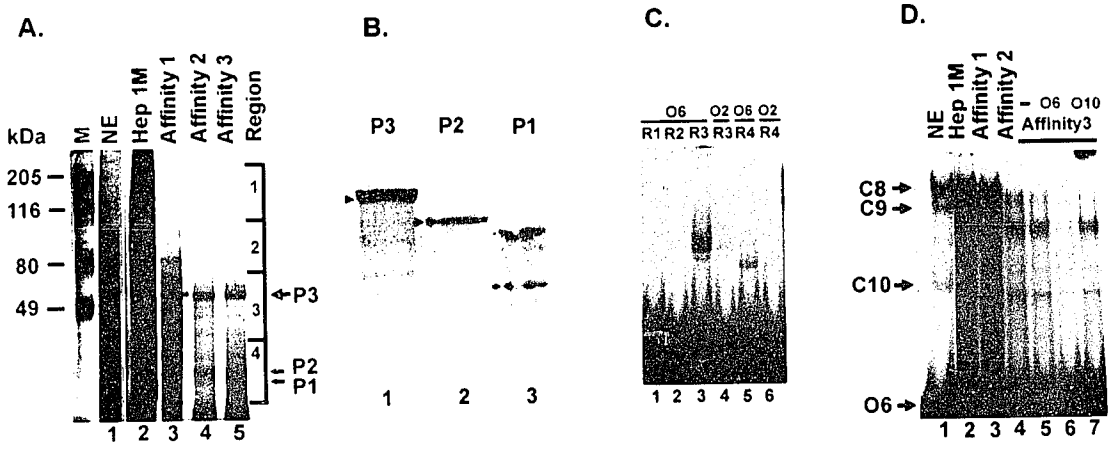

A.
B.
C.
D.
kDa
M
NE
Hep 1M
Affinity 1
Affinity 2
Affinity 3
Region
205
116
80
49
1
2
3
4
5
P3
P2
P1
1 2 3
P3 P2 P1
O6 O2 O6 O2
R1 R2 R3 R3 R4 R4
1 2 3 4 5 6
NE
Hep 1M
Affinity 1
Affinity 2
O6 O10
Affinity3
C8
C9
C10
O6
1 2 3 4 5 6 7

Figure 17. Purification of P3 from the 1 M Heparin fraction using O4-affinity chromatography

A) Proteins from each purification step were electrophoresed through a 10% SDS polyacrylamide gel. A photograph of the silver stained gel is shown. NE, 10 µg; Hep 1M, 10 µg 1 M Heparin fraction; Affinity 1, 20 µl; Affinity 2, 50 µl; Affinity 3, 100 µl. M, protein markers. P3 was enriched. Proteins in each of the regions 1, 2, 3, and 4 were extracted, denatured and renatured (Method 7).

B) P3 was recovered from the preparative gel and digested with V8 protease (Method 6). The resulting peptides were transferred onto a PVDF membrane and stained with Coomassie-Brilliant Blue. Arrows indicate the peptides which were partially sequenced. Sequences obtained from the two peptides were identical.

C) Renatured proteins (8 µl) from region 1 (R1), 2 (R2), 3 (R3), 4 (R4) (indicated in Panel A) were used in gel retardation assays with probes O2 and O4 as indicated on top of the lanes. Protein from R3 interacted with O4 only (lane 4). No DNA-binding activities were detected with the proteins from R1, R2 and R4 (lanes 1, 2, 3, 5).

D) To monitor the purification steps, gel retardation assays with O4 were performed with different protein samples and different amount of poly(dI-dC). Lane 1, 4 µg nuclear extract and 2 µg poly(dI-dC); Lane 2, 2 µg 1 M Heparin fraction and 0.2 µg poly(dI-dC); Lane 3, 2 µl O4 first-eluate and 5 ng poly(dI-dC); Lane 4, 2 µl O4 second-eluate and 1 ng poly(dI-dC); Lanes 5-7, 2 µl O4 third-eluate and no poly(dI-dC). A 50-fold molar excess of unlabeled O4 and O10 were used in Lane 6 and 7 respectively. Two specific DNA-binding complexes were purified (for details, see text). C4, C5, C6 and C7 are abbreviations of fEBC4, fEBC5, fEBC6 and fEBC7 respectively. O4 indicates the free probe.

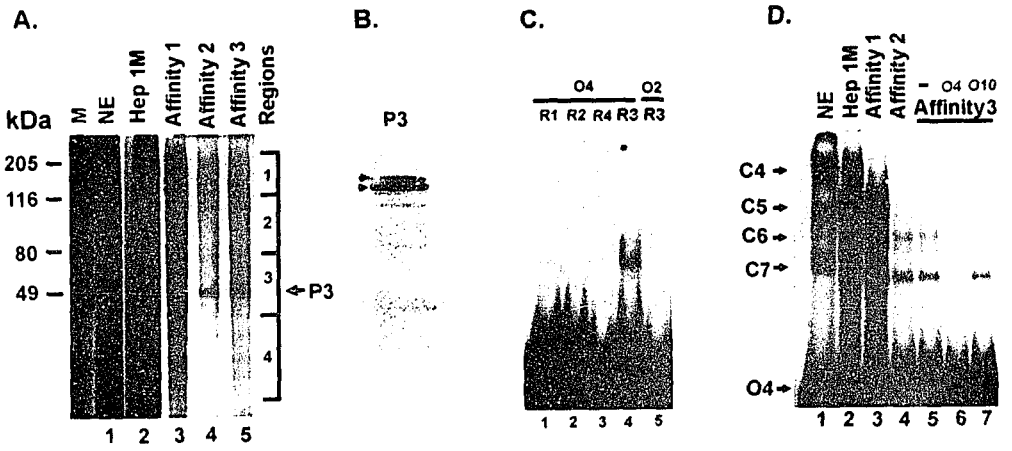

A.
kDa
M
NE
Hep 1M
Affinity 1
Affinity 2
Affinity 3
Regions
205
116
80
49
P3
1
2
3
4
1 2 3 4 5
B.
P3
C.
O4
O2
R1 R2 R4 R3 R3
1 2 3 4 5
D.
NE
Hep 1M
Affinity 1
Affinity 2
- O4 O10
Affinity3
C4
C5
C6
C7
O4
1 2 3 4 5 6 7

Figure 18. Purification of P4, P5, and P6 from the 0.4 M Heparin fraction using O4-affinity chromatography

A) Proteins from each purification step were electrophoresed through a 10% SDS polyacrylamide gel. A photograph of the silver stained gel is shown. NE, 10 μg; Hep 0.4M, 10 μg 0.4 M Heparin fraction; Affinity 1, 20 μl; Affinity 2, 50 μl; Affinity 3, 100 μl. M, protein markers. P4, P5 and P6 were enriched. The dark staining of the bottom of Lanes 3, 4, 5 are from the protease inhibitors used in each purification step. Proteins in regions 1, 2, 3, and 4 were extracted, denatured and renatured (Method 7).

B) P4, P5 and P6 were recovered from the preparative gel and digested with V8 protease (Method 6). Resulting peptides were transferred onto a PVDF membrane and stained with Coomassie-Brilliant Blue. Arrows indicate the peptides which were partially sequenced. Sequences obtained from the three bands of P5 in Lane 2 were identical.

C) Renatured proteins (8 μl) from region 2 (R2), 3 (R3), 4 (R4) (indicated in Panel A) were used in gel retardation assays with probes O4 and O6 as indicated on top of the lanes. A protein from R4 interacted with O4 strongly (lane 4), with O6 weakly (lane 4). No DNA-binding activities were detected with the proteins from R1, R2 and R3 (lanes 1, 2, 3).

D) To monitor the purification steps, gel retardation assays with O4 were performed with different protein samples and different amounts of poly(dI-dC). Lane 1, 4 μg nuclear extract and 2 μg poly(dI-dC); Lane 2, 2 μg 0.4 M Heparin fraction and 0.2 μg poly(dI-dC); Lane 3, 2 μl O4 first-eluate and 5 ng poly(dI-dC); Lane 4, 2 μl O4 second-eluate and 1 ng poly(dI-dC); Lane 5, 2 μl O4 third-eluate and no poly(dI-dC). Three DNA-binding complexes were purified (lane 5) (for details, see text). The shifted bands below C7, indicated by the arrows, were not clearly resolved in this experiment. C4, C5, C6 and C7 are abbreviations of fEBC4, fEBC5, fEBC6 and fEBC7 respectively. O4 indicates the free probe.

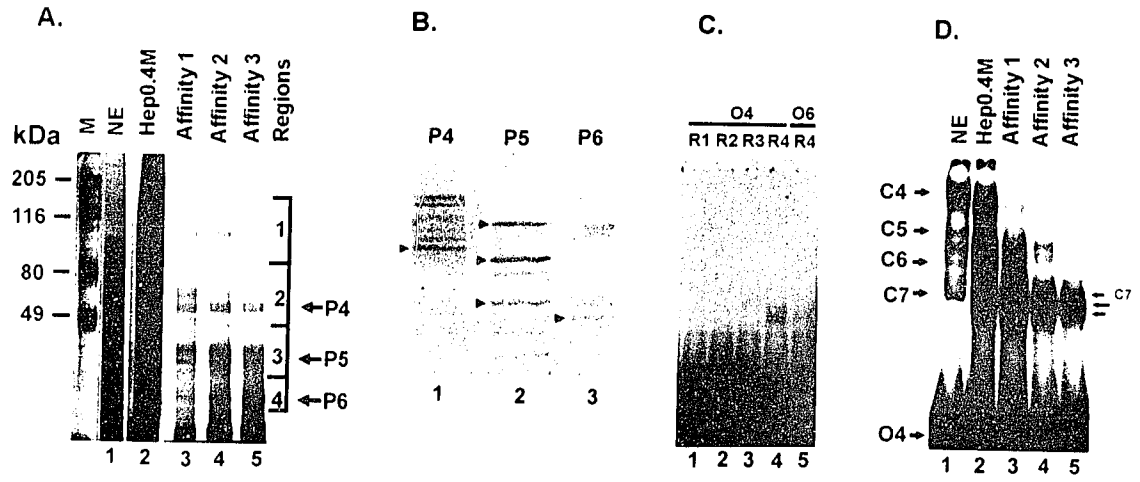

A.
kDa
M
NE
Hep0.4M
Affinity 1
Affinity 2
Affinity 3
Regions
205
116
80
49
1
2
3
4
P4
P5
P6
1 2 3 4 5
B.
P4
P5
P6
1
2
3
C.
O4
O6
R1 R2 R3 R4 R4
1 2 3 4 5
D.
NE
Hep0.4M
Affinity 1
Affinity 2
Affinity 3
C4
C5
C6
C7
C7
O4
1 2 3 4 5

Figure 19. Purification of P5, P7, and P8 from the 0.4 M Heparin fraction using O6-affinity chromatography

A) Proteins from each purification step were electrophoresed through a 10% SDS polyacrylamide gel. A photograph of the silver stained gel is shown. NE, 10 µg; Hep 0.4M, 10 µg 0.4 M Heparin fraction; Affinity 1, 20 µl; Affinity 2, 50 µl; Affinity 3, 100 µl. M, protein markers. P7, P5 and P8 were enriched three major bands. The dark staining of the bottom of Lanes 4, 5 are from the protease inhibitors used in each purification step.

B) P7, P5 and P8 were recovered from a preparative gel and digested with V8 protease (Method 6). The resulting peptides were transferred onto a PVDF membrane and stained with Coomassie-Brilliant Blue. Arrows indicate the peptides which were partially sequenced.

C) To monitor the purification steps, gel retardation assays with O6 were performed with different protein samples and different amounts of poly(dI-dC). Lane 1, 4 µg nuclear extract and 2 µg poly(dI-dC); Lane 2, 2 µg 0.4 M Heparin fraction and 0.2 µg poly(dI-dC); Lane 3, 2 µl O6 first-eluate and 5 ng poly(dI-dC); Lane 4, 2 µl O6 second-eluate and 1 ng poly(dI-dC); Lanes 5-7, 2 µl O6 third-eluate and no poly(dI-dC). A 50-fold molar excess of unlabeled O6 and O10 were used in lane 6 and 7 respectively. Three specific DNA-binding complexes were purified in the O6 third-eluate (lane 5). One is between fEBC9 and fEBC10; the second is fEBC10 (C10); the third is below fEBC10 (for details, see text). C8, C9 and C10 are abbreviations of fEBC8, fEBC9 and fEBC10 respectively. O6 indicates the free probe.

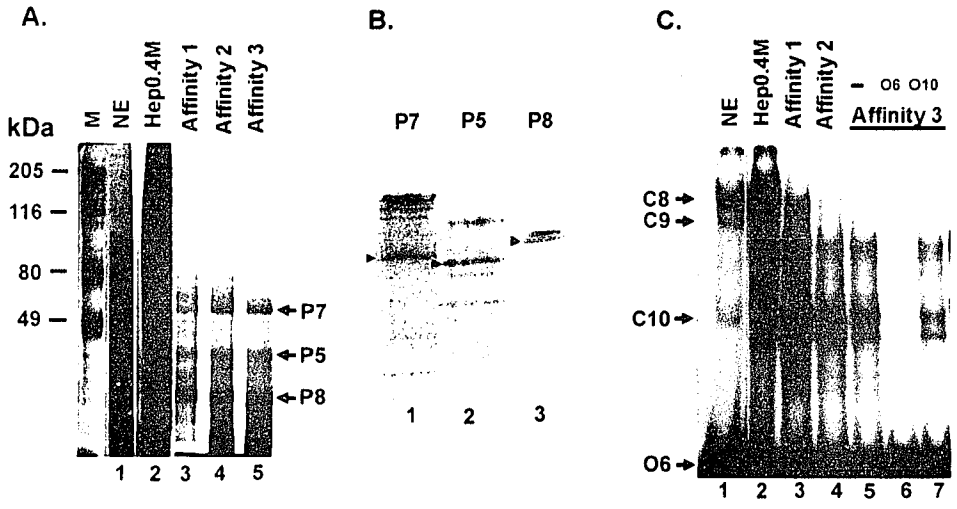

A.
kDa
M
NE
Hep0.4M
Affinity 1
Affinity 2
Affinity 3
205
116
80
49
P7
P5
P8
1 2 3 4 5
B.
P7
P5
P8
1 2 3
C.
NE
Hep0.4M
Affinity 1
Affinity 2
O6 O10
Affinity 3
C8
C9
C10
O6
1 2 3 4 5 6 7

Figure 20. Comparison of three *P1* cDNAs

A) Three *P1* cDNAs (clone A, B, C) are shown schematically. All three clones have the same open reading frame of 555 bp encoding 185 amino acids. Each cDNA contains a unique 5'-leader sequence. The identical portions of 5'-leader sequences in different cDNAs are indicated with filled boxes. The 16 bp of the 5'-leader is present in all three clones; the second 30 bp is present only in clone A and B. The most 5'-end sequences of the three clones are all unique. The 3'-trailers of the three clones are almost identical, but clone B missed 5 bp as indicated. All three clones end with polyadenines.

(B) The synthesized oligonucleotides were used as primers for sequencing *P1* cDNAs by "primer walking". They are indicated schematically. Both panels are not drawn to scale.

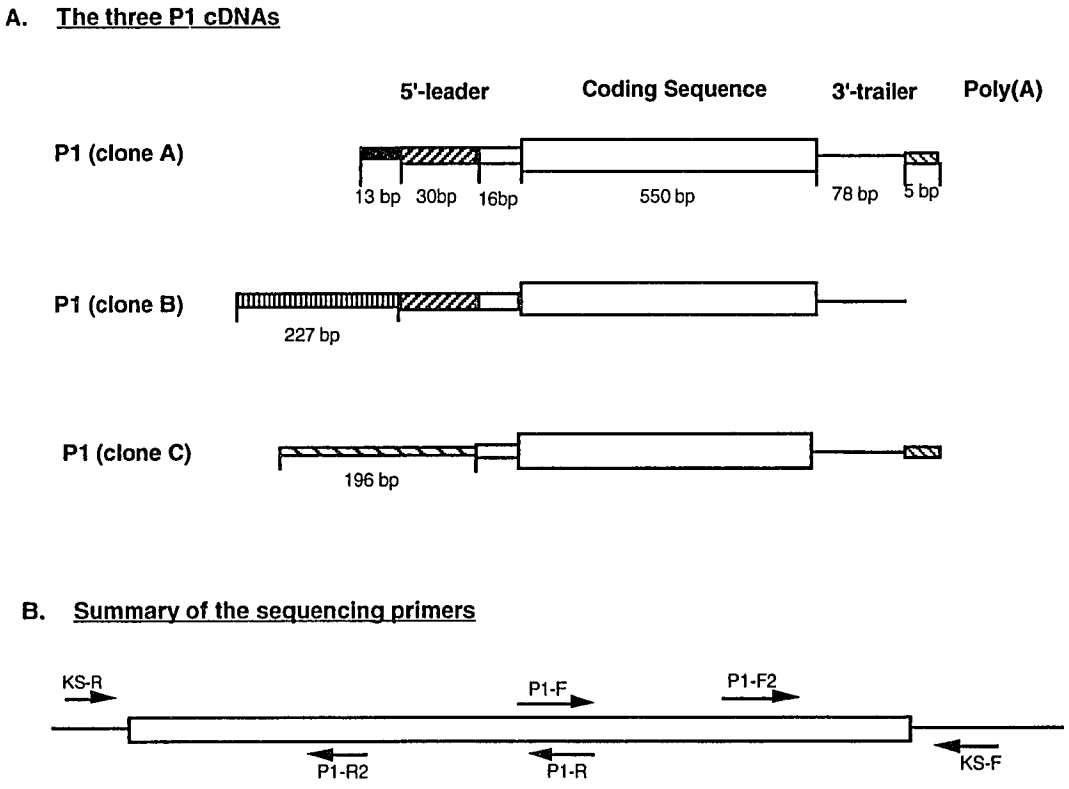


Figure 21. Nucleotide and putative amino acid sequences of P1

The sequence of *P1* cDNA clone B (Fig. 20) is shown. The cDNA is 930 bp long with the following features: 1) A consensus translation sequence (CTS, CAAG) for *Drosophila* genes (Cavener 1987) is indicated at positions 270 to 274 bp; 2) A translation start codon (ATG) is found immediately following the CTS; 3) A translation stop codon (TGA) is found at positions 828 to 830 bp; 4) A polyadenylation signal (PAS, AATAAA) is located at positions 878 to 883 bp; 5) A poly(A) tail is at the end of cDNA. Translated sequence is shown below the open reading frame of cDNA. The first six amino acids of the protein were obtained from micro-sequencing of the purified P1. The deduced nucleotides from these amino acids were used as probes for screening a *Drosophila* cDNA library. A region of 73 amino acids (positions 96 to 168 aa) shows high homology to a part of the ribosomal protein L9 (Fig. 22). A typical nuclear translocation signal (NTS, ProArgArgArgPro) is indicated at the end of the protein.

```
GAAAACTTAAAAGCAAACTTAAACACATACTCAAACCCAAATCAAGCAACTTGGCGGTGC 60

AACAAAAAAGGAATGCTTGAGTATTGCGCAACAGGCTCCAAGATGCGGATTCGGCTTAAG 120

GATTTCGTTAGCAGCTTTTGCAGTTAGCATGTAGCATGTAGCACACAGCAAATAGCAATT 180

GCCCGATTCCAAAAGAACTGGCAAGTGCAGAGGCCTCAAAAAAAAAAACAGAATTGTGTCA 240
                              CTS Start
TACAAACGAGGGACTAACGGAGCTGTCATCAAGATGCGTACCATTAATTCCAACCAGTGC 300
                                   MetArgThrIleAsnSerAsnGlnCys 9
                                         Probe
GTGAAGATCCCTAAGGATATCAAGGCCTTCGGTGAAGGCCCGTGTGGTGACCATCACCGG 360
ValLysIleProLysAspIleLysAlaPheGlyGluGlyProCysGlyAspHisHisArg 29

CACCCGCGGCACCCTGAAGCGCAGCTTCAAGCACTTGGCTCTGGACATGTACATGCCCGA 420
HisProArgHisProGluAlaGlnLeuGlnAlaLeuGlySerGlyHisValHisAlaArg 49

CAAGCGCACCCTGAAGGTGGAGAAGTGGTTCGGAACCAAGAAGGAGTTGGCCGCCGTCCG 480
GlnAlaHisProGluGlyGlyGluValValArgAsnGlnGluGlyValGlyArgArgPro 69

TACCGTCTGCAGTCACATCGAGAACATGATCAAGGGAGTCACGTTTGGATTCCAGTACAA 540
TyrArgLeuGlnSerHisArgGluHisAspGlnGlySerHisValTrpIleProValGln 89

GATGCGTGCTGTGTACGCCATTTCCCCATCAACTGTGTCACCTCCGAGAACAACACGGTC 600
AspAlaCysCysValArgHisPheProIleAsnCysValThrSerGluAsnAsnThrVal 109
                        RL9 homologous region
ATTGAGATCCGTAACTTCTTGGGTGAGAAGTACATCCGTCGTGTGGAGATGGCTCCTGGC 660
IleGluIleArgAsnPheLeuGlyGluLysTyrIleArgArgValGluMetAlaProGly 129

GTCACCGTGGTCAACTCCACTGCCCAGAAGGACGAACTTATCGTGGAGGGAAACGATATT 720
ValThrValValAsnSerThrAlaGlnLysAspGluLeuIleValGluGlyAsnAspIle 149

GAGTCTGTCTCCGGATCTGCCGCCCTCATCCAGCAGTCCACGACCGTCAAGAACAAGATA 780
GluSerValSerGlySerAlaAlaLeuIleGlnGlnSerThrThrValLysAsnLysIle 169

TCCGTAAGTTCCTTGACGGTCTGTACGTGTCCGAGAAGACGACCGTTGTGAAGCTAGAGT 840
SerValSerSerLeuThrValCysThrCysProArgArgArgProLeu Stop         185
                                        NTS
CTTAGACTTTTGGAATGTGTTTCAATCTAAAATGTTAATAAAAAAAAAACCAACGGCCGAA 900
                                       PAS
ACTTTCAAAAAAAAAAAAAAAAAAAAAAA                                  930
            Poly(A) tail
```

Figure 22. Sequence comparison of P1 with the ribosomal protein L9 from different species

The ribosomal protein L9s (RL9s) from different species (human, rat, pea, yeast) were compared using the Pileup program (GCG). A consensus sequence was deduced if three of the four RL9s had the same amino acid. The consensus sequence was compared to P1. The conserved amino acids between P1 and the RL9 consensus sequence were boxed. The underlined P1 region, 73 amino acids, shows 65% identity and 71% similarity to the RL9 consensus sequence.

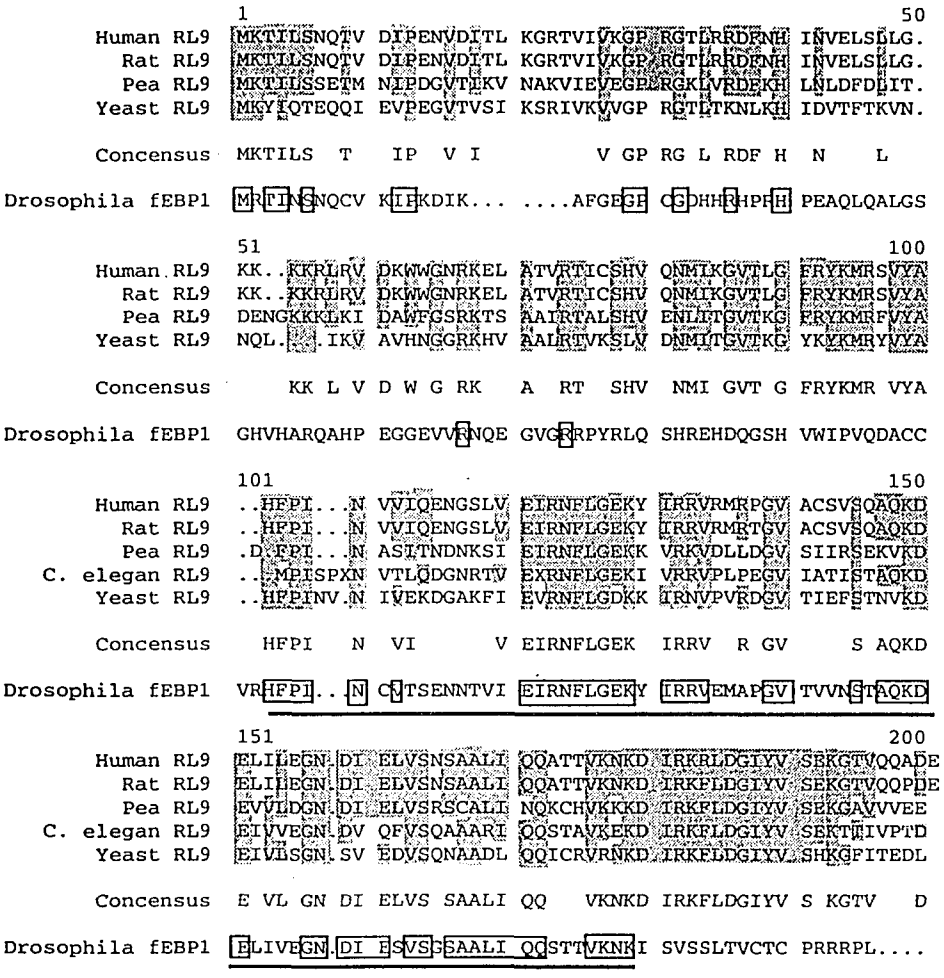

1 50
Human RL9 MKTILSNQTV DIPENVDITL KGRTVIVKGP RGTLRRDFNH INVELSLLG.
Rat RL9 MKTILSNQTV DIPENVDITL KGRTVIVKGP RGTLRRDFNH INVELSLLG.
Pea RL9 MKTIISSETM NIPDGVTIKV NAKVIEVEGP RGKLVRDFKH LNLDFDLIT.
Yeast RL9 MKVIQTEQQI EVPEGVTVSI KSRIVKVVGP RGTLTKNLKH IDVTFTKVN.
Concensus MKTILS T IP V I V GP RG L RDF H N L
Drosophila fEBP1 MRTINSNQCV KIEKDIK... ....AFGEGE GGDHHEHPEH PEAQLQALGS
51 100
Human RL9 KK..KKRLRV DKWWGNRKEL ATVRTICSHV QNMIKGVTLG FRYKMRSVYA
Rat RL9 KK..KKRLRV DKWWGNRKEL ATVRTICSHV QNMIKGVTLG FRYKMRSVYA
Pea RL9 DENGKKKLKI DAMFGSRKTS AAIRTALSHV ENLITGVTKG FRYKMRFVYA
Yeast RL9 NQL..I.IKV AVHNGGRKHV AALRTVKSLV DNMITGVTKG YKYKMRYVYA
Concensus KK L V D W G RK A RT SHV NMI GVT G FRYKMR VYA
Drosophila fEBP1 GHVHARQAHP EGGEVVRNQE GVGRRPYRLQ SHREHDQGSH VWIPVQDACC
101 150
Human RL9 ..HFPI...N VVIQENGSLV EIRNFLGEKY IRRVRMRPGV ACSVSQAQKD
Rat RL9 ..HFPI...N VVIQENGSLV EIRNFLGEKY IRRVRMRTGV ACSVSQAQKD
Pea RL9 .D.FPI...N ASITNDNKSI EIRNFLGEKK VRKVDLLDGV SIIRSEKVKD
C. elegan RL9 ..MPISPXN VTLQDGNRTV EXRNFLGEKI VRRVPLPEGV IATISTAQKD
Yeast RL9 ..HFPINV.N IVEKDGAKFI EVRNFLGDKK IRNVPVRDGV TIEFSTNVKD
Concensus HFPI N VI V EIRNFLGEK IRRV R GV S AQKD
Drosophila fEBP1 VRHFPI..N CNTSENNTVI EIRNFLGEKY IRRVEMAPGV TVVNSTAQKD
151 200
Human RL9 ELIIEGN.DI ELVSNSAALI QQATTVKNKD IRKRLDGIYV SEKGTVQQAD E
Rat RL9 ELIIEGN.DI ELVSNSAALI QQATTVKNKD IRKFLDGIYV SEKGTVQQPD E
Pea RL9 EVVIDGN.DI ELVSRSCALI NQKCHVKKKD IRKFLDGIYV SEKGAVVVEE
C. elegan RL9 EIVVEGN.DV QFVSQAAARI QQSTAVKEKD IRKFLDGIYV SEKTNIVPTD
Yeast RL9 EIVISGN.SV EDVSQNAADL QQICRVRNKD IRKFLDGIYV SHKGFITEDL
Concensus E VL GN DI ELVS SAALI QQ VKNKD IRKFLDGIYV S KGTV D
Drosophila fEBP1 ELIVEGN.DI ESVSSSAALI QQSTTVKNKI SVSSLTVCTC PRRRPL....

Figure 23. Characterization of the *P1* gene

A) Southern blot of the P1 gene. Genomic DNA isolated from *Drosophila* embryos was digested with six restriction enzymes as indicated above the lanes. Hybridization to the digested genomic DNA was performed with the *P1* probe as described in Method 14. An autoradiograph of the Southern blot is shown. One band was detected from each digested genomic DNA sample. They range from 0.5 kb to 8 kb. "kb" indicates DNA markers. The *P1* gene is a single copy gene in the *Drosophila* genome.

B) Northern blot of the *P1* gene. 2 μg mRNA isolated from 0-6 hr *Drosophila* embryos was used for hybridization (Method 11). Only one band was detected at a position of ~0.8 kb as indicated on the right of the Panel. The size of the detected mRNA matches the full length of the *P1* cDNAs.

C) The P1 protein was synthesized using an *in vitro* transcription and translation system. The *P1* cDNA (clone A, Fig. 20) in the Bluescript SK vector was used as template for synthesizing the protein P1. The Bluescript SK vector served as a control template (Cntrl). The ^{35}S-methionine labeled proteins were analyzed by SDS-PAGE. An autoradiograph is show. A 29 kDa peptide was detected with the *P1* cDNA as template (lane 1). Two nonspecific polypeptides were detected with the control template (lane 2).

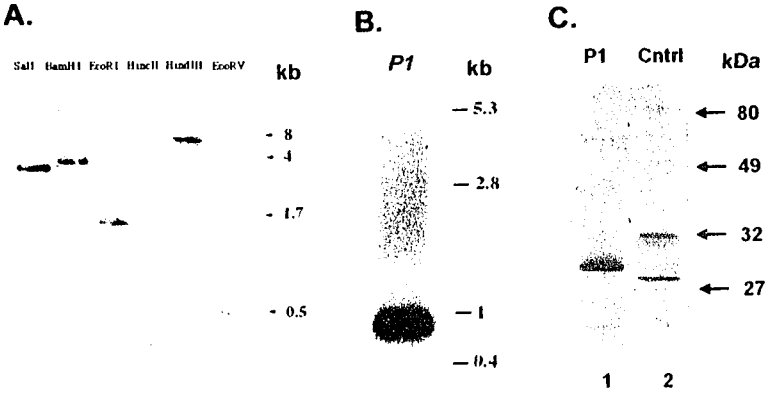


Figure 24. Expression pattern of the *P1* gene in *Drosophila* embryos

The digoxigenin labeled sense and anti-sense *P1* RNA probes were hybridized to 0-16 hr *Drosophila* embryos. Photographs of the stained embryos are shown. The embryos hybridized with the sense probes are shown on the left column (A, B, C, D); those hybridized with the anti-sense probes are shown on the right column (E, F, G, H). Hybridization with the anti-sense probes demonstrated that the *P1* gene is expressed uniformly in all stages of embryos examined. Stages: (A, E) syncytium blastoderm; (B, F) cellular blastoderm; (C, G) germband extension; (D, H) germband retraction.

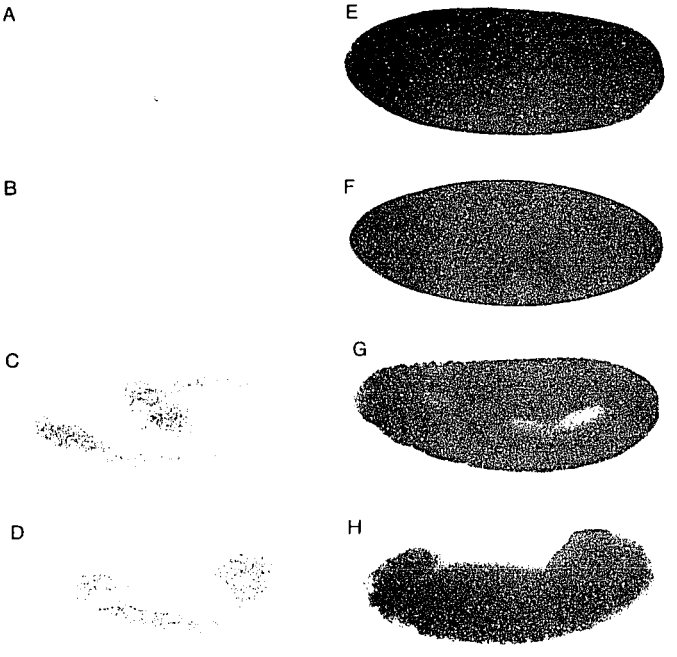


Figure 25. P1 interacts with DNA specifically

A) The *E.coli* recombinant P1 was purified with O6 affinity chromatography. Gel retardation assays with probe O6 were performed with different amounts of the O6 fraction, as indicated on top of the lanes. Two shifted bands were indicated as P1. O6 indicates the free probe.

B) The recombinant P1 (~1.5 μg) was used in the gel retardation assays with probe O6. Different amounts of unlabeled O6 and M6 (mutant O6) were added to the binding reactions as competitors. The two P1-O6 complexes (lane 1) was not affected by addition of either competitors at a 50-fold molar excess over the probe (lanes 2, 4). However, they were abolished by addition of a 100-fold molar excess of unlabeled O6 (lane 3), but not by a 100-fold molar excess of M6 (lane 5). As a control, proteins were purified with O6 affinity chromatography from *E.coli* transformed with the pET3a vector only. No shifted band was detected when the control protein (~2 μg) was used in the gel retardation assays with the probe O6 (Lane 6).

C) The recombinant P1 (~1 μg) was used in the gel retardation assays with the nine oligonucleotide probes corresponding to the nine footprint sites in the proximal enhancer (Fig. 5). The same amount of probe was used in each reaction (~10 fmole). The reactions were electrophoresed through a native 6% polyacrylamide gel. Of the nine probes, seven probes, O1, O2, O3, O4, O5, O8, and O10 interacted with P1 forming one shifted band. However, the protein has different affinities for different oligonucleotides (lanes 1-5, 7, 9). P1 bound to O2, O3 and O10 more strongly than it bound to O1, O4 and O10. P1 strongly interacted with O6 and O9 forming two shifted bands (lanes 6, 8).

D) The P1 monomer DNA-binding consensus sequence was deduced from the nine footprint sites (O1 to O6, O8 to O10). The strand and orientation of the sequences are indicated. The consensus sequence is AGGA.

The P1 dimmer DNA-binding consensus sequence was deduced from the two footprint sites O6 and O9. The strand and orientation of the sequences were indicated. The consensus sequence is PuANCCAAGGAC. Pu, purine (A or G); N, any nucleotide.

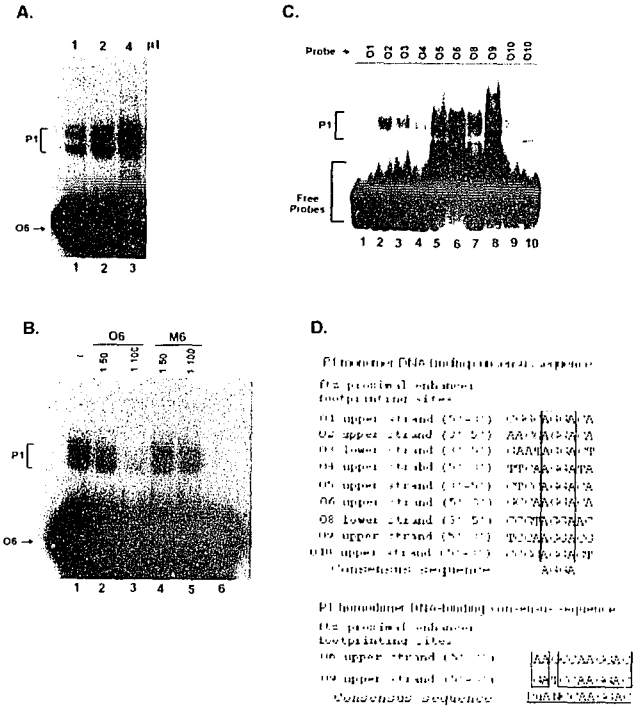

A.
1 2 4 µl
P1
O6 →
1 2 3
C.
Probe
P1
Free Probes
1 2 3 4 5 6 7 8 9 10
B.
O6
M6
P1
O6 →
1 2 3 4 5 6
D.

Figure 26. The transcription activation potential of P1

The *P1* cDNA was cloned into the yeast expression vector pADNS. The yeast strain was transformed with both P1/pADNS and pADNS. This strain carried a 323 bp *ftz* proximal enhancer-*His3* fusion gene as reporter as indicated (lanes 1, 2). Three P1 transformants grew well on the plate (lane 1). Three pADNS transformants (control) did not grow on the same plate (lane 2). The plate lacked leucine (Leu, the pADNS plasmid selection marker) and histidine (His, the reporter gene product), and contained 2.5 mM 3-aminotriazole (3-AT, a competitive inhibitor of His3 product).

The *ftz* cDNA in a yeast expression vector Ycp50 (ftz/Ycp50) and the P1/pADNS were also transformed into another yeast strain. This strain carried six copies of the *engrailed* protein binding site fused to the *His3* reporter as indicated (lanes 3,4). Three FTZ transformants grew well on the plate without histidine and uracil (ura, Ycp50 plasmid selection marker) (lane 3). Three P1 transformants did not grow on the plate without histidine and leucine (lane 4).

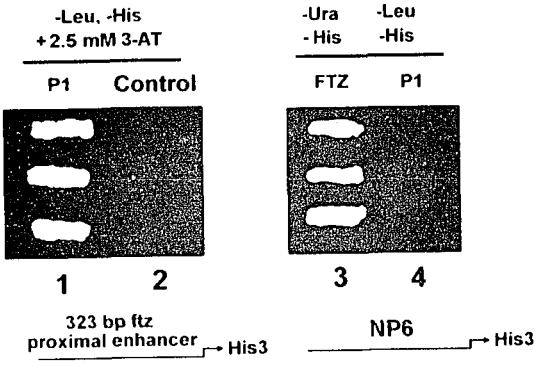


Figure 27. Analysis of the recombinant Adf1 protein

A) The *E.coli* expressed Adf1 was partially purified with O6 affinity chromatography. The protein samples of the *E.coli* extract (10 µg, lane 1) and the partially purified O6 fraction (50 µl, lane 2) were electrophoresed through a 10% SDS polyacrylamide gel. A photograph of the gel stained with Coomassie-Brilliant Blue is shown. A 34 kDa band was enriched in the O6 fraction. MW, protein size markers.

B) The Adf1 fraction (~6 µg, lane 1) and the control protein (~6 µg, lane 2) were prepared as described in Method 14, and electrophoresed through a 10% SDS polyacrylamide gel. The proteins in the gel were then transferred onto a PVDF membrane. Immunohistochemical staining of the proteins in the membrane was performed with the anti-Adf1 antibody as described in Method 14. A photograph of the stained membrane is shown. A 34 kDa protein was stained in the Adf1 fraction only. The other minor species above the Adf1 band in lane 1 could be non-specific staining.

C) The recombinant Adf1 (~0.3 µg) was used in a gel retardation assay with probe O6. An autoradiograph is shown. The Adf1-O6 complex in Lane 1 was abolished by addition of 1 µl anti-Adf1 antibody (α-Adf1, lane 2), but it was not affected by addition of 1 µl preimmune serum (Preim, lane 3). The band below Adf1 in all three lanes was a nonspecific species. Free probe is indicated by O6.

D) Titration of the recombinant Adf1 was carried out in a gel retardation assay with probe O6. An autoradiograph of is shown. As the amount of the O6 fraction was increased in the binding reaction (as indicated on top of the lanes), the amount of shifted probe O6 increased accordingly. The weak bands below Adf1 indicated by arrows could be other forms of Adf1 in the solution interacting with O6 to form different Adf1-O6 complexes.

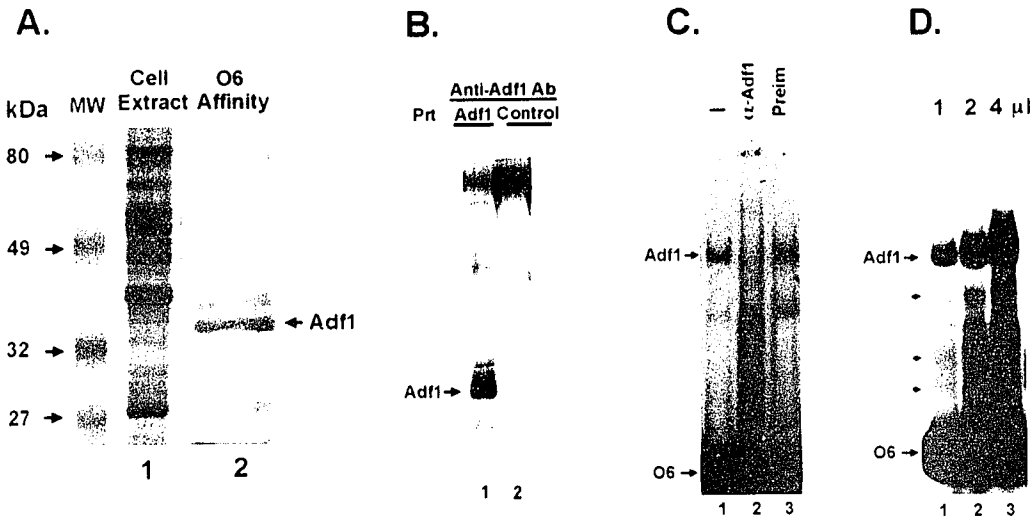

A.
kDa
MW
Cell Extract
O6 Affinity
80
49
32
27
Adf1
1
2
B.
Pre
Anti-Adf1 Ab
Adf1
Control
Adf1
1 2
C.
–
α-Adf1
Preim
Adf1
O6
1 2 3
D.
1 2 4 µl
Adf1
O6
1 2 3

Figure 28. Recombinant Adf1 protein interacts with multiple sites in the *ftz* proximal enhancer

A) The recombinant Adf1 (~0.3 µg) was used in a gel retardation assay with probe O6. Different amounts of unlabeled O6 and M6 (mutant O6) were added to the binding reactions as competitors. An autoradiograph of the competition results is shown. The major Adf1-O6 complex (lane 1) indicated by the arrow was abolished by addition of unlabeled O6 at a 250 to 500-fold molar excess over the probe O6 (lanes 2, 3). It was not competed as well by the same amounts of M6 (lanes 4, 5). As a control, proteins were purified with O6 affinity column from *E.coli* transformed with the pET3a vector only. They did not interact with the probe O6 (lane 6).

B) The recombinant Adf1 (~0.6 µg) was used in a gel retardation assay with the nine oligonucleotide probes corresponding to the nine footprint sites in the proximal enhancer (Fig. 5). The same amount of probe was used in each reaction (~10 fmole). An autoradiograph is shown. Of the nine probes, five probes O5, O6, O8, O9 and O10 interacted with Adf1. However, the protein has different affinities for different oligonucleotides (lanes 5-9). Four probes O1, O2, O3 and O4 did not interact with Adf1 (lanes 1-4).

C) The recombinant Adf1 DNA-binding consensus sequence was deduced from the five footprint sites (O5, O6, O8, O9, O10). The strand and orientation of the sequences are indicated. The consensus sequence is PyCNPuPuGPuNPu. Py, pyrimidine (T or C); Pu, purine (A or G); N, any nucleotides.

D) The recombinant Adf1 (~0.3 µg) was used in a gel retardation assay with probe O6. An Adf1 binding site from the *Adh* distal promoter (termed Adh distal site) was synthesized according to the reported sequence 5'-AGCTG CTGCT GCATC CGTCG ACGTC GACTG CACTC GCCCC-3'(England 1990). Different amounts of unlabeled O6 and Adh distal site were added to the binding reactions as competitors. An autoradiograph of the competition results is shown. The major Adf1-O6 complex (Lane 1) was competed by a titration of unlabeled O6 at 50, 100, and 200-fold molar excess over the probe O6 (lanes 2, 3, 4). It was competed as well by a titration of Adh distal site at 5, 10 and 20-fold molar excess over the probe O6 (lanes 5, 6, 7).

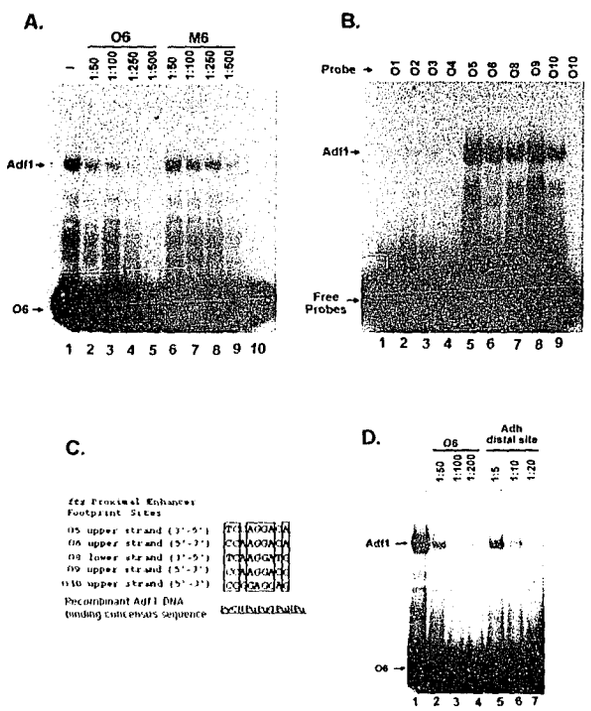

A.
O6
M6
–
1:50
1:100
1:250
1:500
1:50
1:100
1:250
1:500
Adf1→
O6→
1 2 3 4 5 6 7 8 9 10
B.
Probe →
O1
O2
O3
O4
O5
O6
O8
O9
O10
O10
Adf1→
Free Probes →
1 2 3 4 5 6 7 8 9
C.
fzy Proximal Enhancer Footprint Sites
O5 upper strand (3'-5')
O6 upper strand (5'-3')
O8 lower strand (3'-5')
O9 upper strand (5'-3')
O10 upper strand (5'-3')
Recombinant Adf1 DNA binding consensus sequence
D.
O6
Adh distal site
1:50
1:100
1:200
1:5
1:10
1:20
Adf1 →
O6 →
1 2 3 4 5 6 7

Figure 29. The transcription activation potential of Adf1

A) 0-12 hr embryo nuclear extract was used in a gel retardation assay with probe O6. An autoradiograph is shown. fEBC8, 9 and 10 were detected from the nuclear extract as before (lane 1, Fig. 6). fEBC8 was abolished by addition of anti-Adf1 antibody (α-Adf1) to the reaction (lane 2). It was not affected by addition of preimmune serum (Preim, lane 3). fEBC9 and 10 were not affected specifically by either antibody or preimmune serum. No protein was added in lane 4. Thus, the native Adf1 in the nuclear extracts is involved in the formation of fEBC8.

B) Since fEBC8 bound to three footprint sites (Fig. 7), the native Adf1 also interacted with the same three sites. Parts of the sequences of the three binding sites (O6, O8, O9) for the native Adf1 were aligned. The consensus sequence, PuANPyCAAGGAPy, was deduced. Pu, purine (A or G); Py, pyrimidine (T or C); N, any nucleotide.

C) The *Adf1* cDNA was cloned into the yeast expression vector pADNS. The yeast strain was transformed with both Adf1/pADNS, and pADNS. This strain carried a 323 bp *ftz* proximal enhancer-*His3* fusion gene as a reporter as indicated (lanes 1, 2). Three Adf1 transformants grew well on the plate (lane 1). Three pADNS transformant (control) did not grow on the same plate (lane 2). The plate lacked leucine (Leu, the pADNS plasmid selection marker) and histidine (His, the reporter gene product), and contained 2.5 mM 3-aminotriazole (3-AT, a competitive inhibitor of His3 product).

The *ftz* cDNA in a yeast expression vector Ycp50 and Adf1/pADNS were also transformed into another yeast strain. This strain carried six copies of the *engrailed* protein binding site fused to the *His3* reporter as indicated (lanes 3, 4). Three FTZ transformants grew well on the plate without histidine and uracil (ura, Ycp50 plasmid selection marker) (lane 3). Three Adf1 transformants did not grow on the plate without histidine and leucine (lane 4).

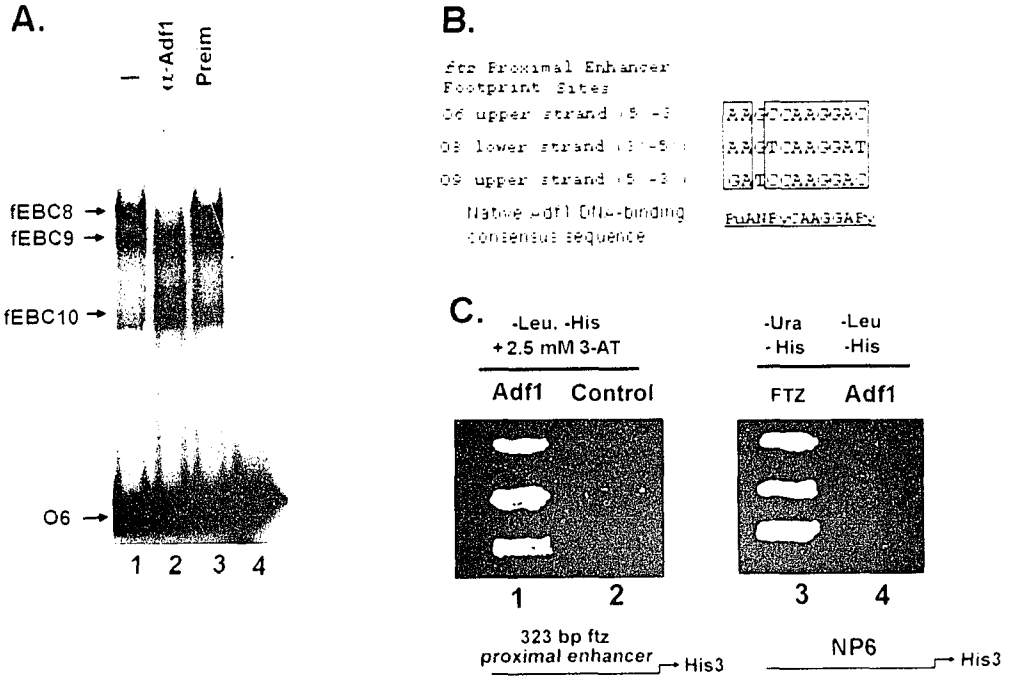

A.
—
α-Adf1
Preim
fEBC8
fEBC9
fEBC10
O6
1 2 3 4
B.
ftz Proximal Enhancer
Footprint Sites
C.
-Leu, -His
+2.5 mM 3-AT
Adf1
Control
-Ura
-His
-Leu
-His
FTZ
Adf1
1 2 3 4
323 bp ftz
proximal enhancer
His3
NP6
His3

Figure 30. The transcriptional regulatory mechanisms of the *ftz* proximal enhancer

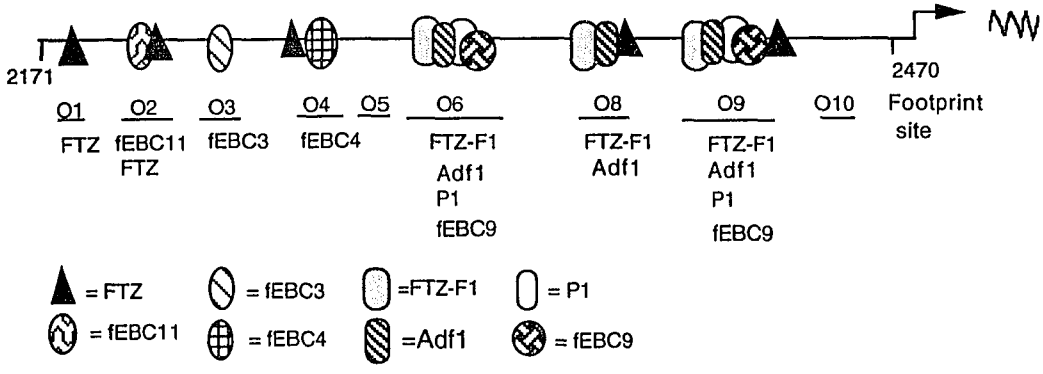
A, B) A 300 bp fragment from 2171 to 2470 bp in the *ftz* proximal enhancer is shown. The nine footprint sites used as probes in gel retardation assays are underlined (O1 to O6, O8 to O10).

A) Activation of gene expression. The potential transcriptional activators (FTZ, FTZ-F1, Adf1, P1) for *ftz* gene expression are listed below their binding sites. fEBC11, 3, 4, and 9 detected in the gel retardation assays with the 0-4 hr embryo nuclear extracts are also shown under their binding sites. FTZ interacted with five sites in this region (Pick et al. 1990). The high-affinity site (2198 to 2215) overlaps with O2; of the three medium-affinity sites, one (2177 to 2195) overlaps with O1; one (2263 to 2278) overlaps partially with O4, and one (2400 to 2426) overlaps partially with O8. The low-affinity site (2435 to 2416) overlaps partially with O9. Thus, multiple transcriptional activators interact repeatedly with the *ftz* proximal enhancer to activate the gene expression.

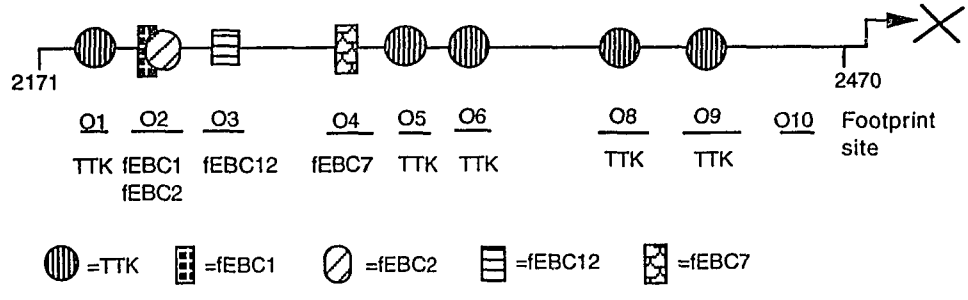
B) Repression of gene expression. fEBC1, 2, 12, 7 were detected with 4-9 hr embryo nuclear extract. They were probably formed by unidentified transcriptional repressors interacting with O2, O3, and O4 in the gel retardation assays. Those fEBCs and the potential transcriptional repressor TTK are listed under their binding sites. Thus, multiple transcriptional repressors competitively occupy the protein binding sites in the *ftz* proximal enhancer to turn off gene expression.

C) Spatial gene expression. A schematic depiction of a embryo with two *ftz* stripes in black and one interstripe in white is shown. The reporter gene *lacZ* directed by the proximal enhancer can not be activated in the interstripe regions after the *ftz* seven stripes are established (the *ftz* autoregulation phase). The other potential transcriptional activators (fEBC3, 4, 9, 11, FTZ-F1, Adf1 and P1) are not sufficient to activate the gene expression to a detectable level in this phase without the FTZ protein. Thus, the reporter gene expression can be only activated to a detectable level in the regions where both FTZ protein and the other potential transcriptional activators are available.

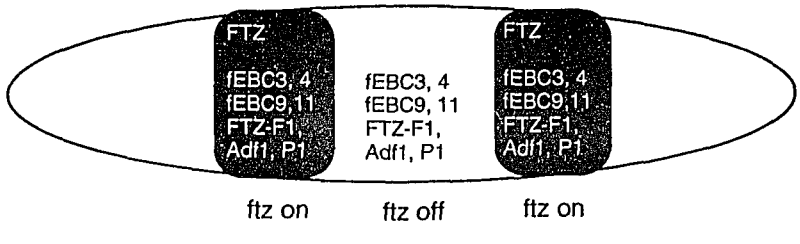
A. Activation of gene expression



B. Repression of gene expression



C. Spatial gene expression



CHAPTER 5. BIBLIOGRAPHY

Akam, M. 1987. The molecular basis for metameric pattern in the *Drosophila* embryo. *Development* **101:**1-22.

Anderson, K. 1987. Dorsal-ventral embryonic pattern genes of *Drosophila*. *Trends in Genetics* **3:**91-97.

Ayer, S., and C. Benyajati. 1992. The binding site of a steroid hormone receptor-like protein within the *Drosophila Adh* adult enhancer is required for high levels of tissue-specific alcohol dehydrogenase expression. *Mol. Cell. Biol.* **12:**661-673.

Biggin, M. D., S. Bickel, M. Benson, V. Pirrotta, and R. Tjian. 1988. *Zeste* encodes a sequence-specific transcription factor that activates the *Ultrabithorax* promoter in vitro. *Cell* **53:**713-722.

Biggin, M. D., and R. Tjian. 1988. Transcriptional factors that activate the *Ultrabithorax* promoter in developmentally staged extracts. *Cell* **53:**699-711.

Biggin, M. D., and R. Tjian. 1989. Transcription factors and the control of *Drosophila* development. *Trends Genet.* **53:**377-383.

Bopp, D., E. Jamet, S. Baumgartner, M, Burri, and M. Noll. 1989. Isolation of two tissue-specific *Drosophila* paired box genes, *Pox meso* and *Pox neuro*. *EMBO J.* **8:**3447-3457.

Bradford, M. M. 1976. A rapid and sensitive method for measuring of microgram quantities of protein utilizing the principle of protein-dye binding. *Anal. Biochem.* **72:**248-253.

Bray, S. J., B. Burke, N. H. Brown, and J. Hirsh. 1989. Embryonic expression pattern of a family of *Drosophila* proteins that interact with a central nervous system regulatory element. *Genes & Dev.* **3:**1130-1145.

Briggs, M. R., J. T. Kadonaga, S. P. Bell, and R. Tjian. 1986. Purification and biochemical characterization of the promoter-specific transcription factor Sp1. *Science* **234:**47-52.

Brown, J. L., S. Sonoda, H. Oeda, M. P. Scott, and C. Wu. 1991. Repression of the *Drosophila fushi tarazu* (*ftz*) segmentation gene. *EMBO J.* **10:**665-674.

Brown, J. L., and C. Wu. 1993. Repression of *Drosophila* pair-rule segmentation genes by ectopic expression of *tramtrack*. *Development* **117:**45-58.

Carroll, S. B., and M. P. Scott. 1985. Localization of *fushi tarazu* protein during *Drosophila* embryogenesis. *Cell* **43:**47-57.

Carroll, S. B., and M. P. Scott. 1986. Zygotically active genes that affect the spatial expression of the *fushi tarazu* segmentation gene during early *Drosophila* embryogenesis. *Cell* **45:**113-126.

Carroll, S. B., and S. H. Vavra. 1989. The zygotic control of *Drosophila* pair-rule gene expression II. Spatial repression by gap and pair-rule gene products. *Development* **107:**673-683.

Carroll, S. B., G. M. Winslow, T. Schupbach, and M. P. Scott. 1986. Maternal control of *Drosophila* segmentation gene expression. *Nature* **323:**278-280.

Carthew, R. W., L. A. Chodosh, and P. A. Sharp. 1985. An RNA polymerase II transcription factor binds to an upstream element in the adenovirus major late promoter. *Cell* **43:**439-448.

Cavener, D. R. 1987. Comparison of the consensus sequence flanking translational start sites in *Drosophila* and vertebrates. *Nucleic Acids Res.* **15:**1353-1361.

Chodosh, L. A., R. W. Carthew, and P. A. Sharp. 1986. A single polypeptide possesses the binding and transcription activities of the adenovirus major late transcription factor. *Mol. Cell. Biol.* **6:**4723-4733.

Cohn, N. S. 1992. P. sativum GA mRNA (clone F) similar to ribosomal protein L9. Direct submission to Gene Bank. Unpublished.

Dearof, C. R., J. Topol, and C. S. Parker. 1989. Transcriptional control of *Drosophila fushi tarazu* zebra stripe expression. *Genes & Dev.* **3:**384-398.

Desplan, C., J. Theis, and P. H. O'Farrell. 1985. The *Drosophila* developmental gene, *engrailed*, encodes a sequence-specific DNA binding activity. *Nature* **318:**630-635.

Diamond, M. I., J. N. Miner, S. K. Yoshinga, and K. R. Yamamoto. 1990. c-Jun and c-Fos levels specify positive or negative glucocorticoid regulation from a composite GRE. *Science* **249:**1266-1272.

Dynan, W. S. 1989. Modularity in promoters and enhancers. *Cell* **58:**1-4.

Dynlacht, B. D., L. D. Attardi, A. Admon, M. Freeman, and R. Tjian. 1989. Functional analysis of NTF-1, a developmentally

regulated *Drosophila* transcription factor that binds neuronal *cis* elements. *Genes & Dev.* **3:**1677-1688.

England, B. P., A. Admon, and R. Tjian. 1992. Cloning of *Drosophila* transcription factor Adf-1 reveals homology to Myb oncoprotein. *Proc. Natl. Acad. Sci. U.S.A.* **89:**683-687.

England, B. P., U. Heberlein, and R. Tjian. 1990. Purified *Drosophila* transcription factor, *Adh* distal factor-1 (Adf-1), binds to sites in several *Drosophila* promoters and activates transcription. *J. Biol. chem.* **265:**5086-5094.

Fitzpatrick, V. D., and C. J. Ingles. 1989. The Drosophila *fushi tarazu* polypeptide is a DNA binding transcriptional activator in yeast cells. *Nature* **337:**666-668.

Frohnhofer, H. G., and C. Nusslein-Volhard. 1987. Maternal genes required for the anterior localization of *bicoid* activity in the embryo of *Drosophila*. *Genes & Dev.* **1:**880-890.

Gehring, W. J. 1987. Homeoboxes in the study of development. *Science* **236:**1245-1252.

Goldfarb, D. S., J. Gariepy, G. Schoolnik, and R. D. Kornberg. 1986. Synthetic peptides as nuclear localization signals. *Nature* **322:**641-644.

Goto, T., P. Murdonald, and T. Maniatis. 1989. Early and late periodic patterns of *even-skipped* expression are controlled by distinct regulatory elements that respond to different spatial cues. *Cell* **57:**413-422.

Gutjahr, T., E. Frei, and M. Noll. 1993. Complex regulation of early *paired* expression: initial activation by gap genes and pattern modulation by pair-rule genes. *Development* **117:**609-623.

Hafen, E., A. Kuroiwa, and W. J. Gehring. 1984. Spatial distribution of transcripts from the segmentation gene *fushi tarazu* during *Drosophila* embryonic development. *Cell* **37:**833-841.

Han, K., and M. S. Levine. 1989. Synergistic activation and repression of transcription by *Drosophila* homeobox proteins. *Cell* **56:**573-583.

Harding, K., T. Hoey, R. Warrior, and M. Levine. 1989. Autoregulatory and gap gene response elements of the *even-skipped* promoter of *Drosophila*. *EMBO J.* **8:**1205-1212.

Harrison, S. D., and A. Travers. 1988. Identification of the binding sites for potential regulation proteins in the upstream enhancer element of *Drosophila fushi tarazu* gene. *Nucleic Acids Res.* **16:**11403-11416.

Harrison, S. D., and A. Travers. 1990. The *tramtrack* gene in *Drosophila* encodes a zinc finger protein that interacts with the *ftz* transcriptional regulatory region and shows a novel embryonic expression pattern. *EMBO J.* **9:**207-216.

Heberlein, U., B. England, and R. Tjian. 1985. Characterization of *Drosophila* transcription factors that activate the tandem promoters of the alcohol dehydrogenase gene. *Cell* **41:**965-977.

Hendzickson, W., and R. Schleif. 1985. A dimmer of AraC protein contacts three adjacent major groove regions at the AraI DNA site. *Proc. Natl. Acad. Sci. U.S.A.* **82:**3129-3133.

Hiromi, Y., and W. J. Gehring. 1987. Regulation and function of the *Drosophila* segmentation gene *fushi tarazu. Cell* **50:**963-974.

Hiromi, Y., A. Kuroiwa, and W. J. Gehring. 1985. Control elements of the *Drosophila* segmentation gene *fushi tarazu. Cell* **43:**603-613.

Hoey, T., and M. Levine. 1988. Divergent homeobox proteins recognize similar DNA sequences in *Drosophila. Nature* **332:**858-861.

Hori, N., K. Murakawa, R. Matoba, A. Fukushima, K. Okubo, and K. Matsubara. 1993. A new human ribosomal protein sequence, homologue of rat L9. *Nucleic Acids Res.* **21:**4395.

Ingham, P., and P. Gergen. 1988. Interactions between the pair-rule genes *runt, hairy, even-skipped* and *fushi tarazu* and the establishment of periodic pattern in the *Drosophila* embryo. *Development* **104 supplement:**51-60.

Ingham, P. W. 1988. The molecular genetics of embryonic pattern formation in *Drosophila. Nature* **335:**25-34.

Innis, M. A., and D. H. Gelfand. 1990. Optimization of PCRs. In *PCR protocols: a guide to methods and applications* (ed. Innis, M.A., Gelfand, D.H., Sninsky, J.J., White, T.J.), pp3-12. Academic Press, Inc., San Diego, California.

Innis, M. B., D. H. Myambo, Gelfand, and M. A. Drow. 1988. DNA sequencing with thermus aquaticus DNA polymerase and direct sequencing of polymerase chain reaction-amplified DNA. *Proc, Natl. Acad. Sci. U.S.A.* **85:**9436-9440.

Jaynes, T. B., and P. H. O'Farrell. 1988. Activation and repression of transcription by homeodomain-containing proteins that bind a common site. *Nature* **336:**744-749.

Jiang, J., and M. Levine. 1993. Binding affinities and cooperative interactions with bHLH activators delimit threshold responses to the dorsal gradient morphogen. *Cell* **72:**741-752.

Jiang, T., T. Hoey, and M. Levine. 1991. Autoregulation of a segmentation gene in *Drosophila* combination interaction of the *eve-skipped* homeobox protein with a distal enhancer element. *Genes & Dev.* **5:**265-277.

Johnston, D. S., and C. Nusslein-Volhard. 1992. The origin of pattern and polarity in the Drosophila embryo. *Cell* **68:**201-219.

Jones, D. G., U. Reusser, and G. H. Braus. 1991. Cloning and characterization of a yeast homolog of the mammalian ribosomal protein L9. *Nucleic Acids Res.* **19:**5785.

Kadonaga, J. T., and R. Tjian. 1986. Affinity purification of sequence-specific DNA binding proteins. *Proc. Natl. Acad. Sci. U.S.A.* **83:**5889-5893.

Kishore, H. M., and D. M. Shah. 1988. Amino acid biosynthesis inhibitors as herbicides. *Annu. Rev. Biochem.* **57:**627-663.

Krause, H. M., R. Klemenz, and W. J. Gehring. 1988. Expression, modification and localization of the *fushi tarazu* protein in *Drosophila* embryos. *Genes & Dev.* **2:**1021-1036.

Kuroiwa, A., E. Hafen, and W. J. Gehring. 1984. Cloning and transcriptional analysis of the segmentation gene *fushi tarazu* of *Drosophila*. *Cell* **37:**825-831.

Laemmli, U. K. 1970. Cleavage of structural proteins during the assembly of the head of *bacteriophage T4*. *Nature* **227:**680-685.

Lanford, R. E., and J. S. Butel. 1984. Construction and characterization of an SV40 mutant defective in nuclear transport of T antigen. *Cell* **37:**801-813.

Lavorgna, G., H. Leda, J. Clos, and C. Wu. 1991. FTZ-F1, a steroid hormone receptor-like protein implicated in the activation of *fushi tarazu*. *Science* **252:**848-851.

Lawrence, P., P. Johnston, P. Mardonald, and G. Struhl. 1987. Borders of parasegments in *Drosophila* embryos are delimited by the *fushi tarazu* and *even-skipped* genes. *Nature* **328:**440-442.

Levine, M., and T. Hoey. 1988. Homeobox proteins as sequence-specific transcription factors. *Cell* **55:**537-540.

Lewin, B. 1980. Gene expression. In *Eukaryotic chromosomes* (ed. B. Lewin), Vol.2, 2nd ed., pp. 479-502. Wiley, New York.

Lewis, E. B. 1978. A gene complex controlling segmentation in *Drosophila*. *Nature* **276:**565-570.

Maniatis, T., E. F. Fritsch, and S. Sambrook. 1982. Molecular cloning: A laboratory manual. Cold Spring Harbor Laboratory Press, New York.

Martinez-Arias, A., and P. A. Lawrence. 1985. Parasegments and compartments in the *Drosophila* embryos. *Dev. Biol.* **99:**261-264.

Maxam, A. M., and W. Gilbert. 1977. A new method for sequencing DNA. *Proc. Natl. Acad. Sci. U.S.A.* **74:**560-564.

Maxam, A. M., and W. Gilbert. 1980. Sequencing end-labeled DNA with base specific chemical cleavages. *Meth. Enzym.* **65:**409-560.

McCombie, W. R., J. M. Kelley, J. C. Venter, and C. A. Fields. 1993. Caenorhabditis elegans cDNA clone CEESK81 similar to ribosomal protein L9. Direct submission to Gene Bank. Unpublished.

McGinnis, W., M. S. Levine, E. Hafen, A. Kuroiwa, and W. J. Gehring. 1984. A conserved DNA sequence in homeotic genes of the *Drosophila melanogaster Antennapedia* and *Bithorax* complexes. *Nature* **308:**428-433.

Mitchell, P. J., and R. Tjian. 1989. Transcriptional regulation in mammalian cells by sequence specific DNA binding proteins. *Science* **245:**371-378.

Mlodzik, M., and W. J. Gehring. 1987. Expression of the *caudal* gene in the germ line of *Drosophila*: formation of an RNA and protein gradient during early embryogenesis. *Cell* **48:**465-478.

Murre, C., P. S. McCaw, and D. Baltimore. 1989. A new DNA binding and dimerization motif in immunoglobulin enhancer binding, daughterless, myoD, and myc proteins. *Cell* **56:**777-783.

Nulsslein-Volhard, C., and E. Wieschaus. 1980. Mutations affecting segment number and polarity in *Drosophila*. *Nature* **287:**795-801.

Nusslein-Volhard, C. 1991. Determination of the embryonic axes of *Drosophila*. *Development Supplement* **1:**1-10.

Nusslein-Volhard, C., H. G. Frohnhofer, and R. Lehmann. 1987. Determination of anterioposterior polarity in *Drosophila*. *Science* **238:**1675-1681.

Nusslein-Volhard, C., E. Wieschaus, and H. Kluding. 1984. Mutations affecting the pattern of the larval cuticle in *Drosophila melanogaster* I. Zygotic loci on the second chromosome. *Wilhelm Ronx's Arch. Dev. Biol.* **193:**267-282.

Ohno, C. K., and M. Petkovich. 1992. *FTZ-F1β* a novel member of the *Drosophila* nuclear receptor family. *Mech. Dev.* **40:**13-24.

Perkins, K. K., A. Admon, N. Patel, and R. Tjian. 1990. The *Drosophila* Fos-related AP-1 protein is a developmentally regulated transcription factor. *Genes & Dev.* **4:**822-834.

Perkins, K. K., G. M. Dailey, and R. Tjian. 1988. In vitro analysis of the *Antennapedia* P2 promoter: identification of a new *Drosophila* transcription factor. *Genes & Dev.* **2:**1615-1626.

Pick, L., A. Schier, M. T. Schmidt-Glenewinkel, and W. J. Gehring. 1990. Analysis of the *ftz* upstream element: Germ layer-specific enhancers are independently autoregulated. *Genes & Dev.* **4:**1224-1239.

Read, D., M. Levine, and J. L. Manley. 1992. Ectopic expression of the *Drosophila tramtrack* gene results in multiple embryonic defects, including repression of *even-skipped* and *fushi tarazu*. *Mech. Dev.* **38:**183-196.

Read, D., and J. L. Manley. 1992. Alternatively spliced transcripts of the *Drosophila tramtrack* gene encode zinc finger proteins with distinct DNA binding specificities. *EMBO J.* **11:**1035-1044.

Read, D., T. Nishigaki, and J. L. Manley. 1990. The *Drosophila even-skipped* promoter contains multiple, overlapping factor binding sites and is transcribed in a stage-specific manner *in vitro*. *Mol. Cell. Biol.* **10:**4334-4344.

Read, D. B. 1992. Expression and function of the *tramtrack* gene of *Drosophila*. PhD thesis. Columbia University, New York.

Regulski, M., S. Dessain, N. McGinnis, and W. McGinnis. 1991. High-affinity binding sites for the *Deformed* protein are required for the function of an autoregulatory enhancer of the *Deformed* gene. *Genes & Dev.* **5:**278-286.

Rubin, G. M., and A. C. Spradling. 1982. Genetic transformation of *Drosophila* with transposable element vectors. *Science* **218:**348-353.

Sanger, F., S. Nicklen, and A. R. Coulson. 1977. DNA sequencing with chain-termination inhibitors. *Proc, Natl. Acad. Sci. U.S.A.* **74:**5463-5467.

Schier, A. F., and W. J. Gehring. 1992. Direct homeodomain-DNA interaction in the autoregulation of the *fushi tarazu* gene. *Nature* **356:**804-807.

Scott, M. P., J. W. Tamkun, and G. W. Hartzell. 1989. The structure and function of the homeodomain. *BBA Rev Cancer* **989:**25-48.

Scott, M. P., and A. J. Weiner. 1984. Structural relationships among genes that control development: sequence homology between the *Antennapedia*, *Ultrabithorax* and *fushi tarazu*. *Proc. Natl. Acad. Sci. U.S.A.* **81:**4115-4119.

Simcox, A., and J. H. Sams. 1983. When does determination occur in *Drosophila* embryos? *Dev. Biol.* **97:**212-221.

Small, S., R. Kraut, T. Hoey, R. Warrior, and M. Levine. 1991. Transcriptional regulation of a pair-rule stripe in *Drosophila*. *Genes & Dev.* **5:**827-839.

Smith, A. P., M. Muller, M. Affolter, and W. Gehring. 1990. The interaction with DNA of wildtype and mutant *fushi tarazu* homeodomains. *EMBO J.* **9:**3967-3974.

Soeller, W. C., C. E. Oh, and T. B. Kornberg. 1993. Isolation of cDNAs encoding the *Drosophila* GAGA transcription factor. *Mol. Cell. Biol.* **13:**7961-7970.

Soeller, W. C., S. J. Pook, and T. Kornberg. 1988. *In vitro* transcription of the *Drosophila engrailed* gene. *Genes & Dev.* **2:**68-81.

Sprandling, A. C., and G. M. Rubin. 1982. Transposition of cloned P element into *Drosophila* germ line chromosomes. *Science* **218:**341-347.

Staudt, L. M., and M. J. Lenardo. 1991. Immunoglobulin gene transcription. *Annu. Rev. Immunol.* **9:**373-398.

Steffan, N. H., H. D. Hunt, and L. R. Pease. 1989. Site-directed mutagenesis by overlap extension using the polymerase chain reaction. *Gene* **77:**51-59.

Struhl, G. 1985. Near-reciprocal phenotypes caused by inactivation of indiscrimination expression of the *Drosophila* segmentation gene *ftz*. *Nature* **318:**677-680.

Suzuki, K., J. Olvera, and I. G. Wool. 1990. The primary structure of rat ribosomal protein L9. *Gene* **93:**297-300.

Tautz, D., and C. Pfeifle. 1989. A non-radioactive in situ hybridization method for the localization of specific RNAs in *Drosophila* embryos reveals translational control of the segmentation gene *hunchback*. *Chromosoma* **98:**81-85.

Technau, G. M. 1987. A single cell approach to problems of cell lineage and commitment during embryogenesis of *Drosophila melanogaster*. *Development* **100:**1-12.

Topol, T., C. R. Dearolf, K. Prakash, and C. S. Parker. 1991. Synthetic oligonucleotide recreate *Drosophila fushi tarazu* zebra-stripes expression. *Genes & Dev.* **5:**855-867.

Ueda, H., S. Sonoda, L. Brown, M. P. Scott, and C. Wu. 1990. A sequence-specific DNA-binding protein that activates *fushi tarazu* segmentation gene expression. *Genes & Dev.* **4:**624-635.

Vavra, S. H., and S. B. Carroll. 1989. The zygotic control of *Drosophila* pair-rule gene expression I. A search for new pair-rule regulatory loci. *Development* **107:**663-672.

Wakimoto, B. T., and T. C. Kaufmann. 1981. Analysis of larval segmentation in the lethal genotypes associated with the *Antennapedia* gene complex in *Drosophila melanogaster*. *Dev. Biol.* **81:**51-64.

Wamplar, S. L., C. M. Tyree, and J. T. Kandonaga. 1990. Fraction of the general RNA polymerase II transcription factors from *Drosophila* embryos. *J. Biol. Chem.* **265:**21223-21231.

Weir, M. P., and T. Kornberg. 1985. Patterns of *engrailed* and *fushi tarazu* transcripts reveal novel intermediate stages of *Drosophila* segmentation. *Nature* **318:**433-439.

White, B. A., G. M. Preston, T. C. Lufkin, and C. Bancraft. 1985. Detection of two chromatin proteins which bind specifically to the 5'-flanking region of the rat prolactin gene. *Mol. Cell. Biol.* **5:**2967-2974.

Xiong, W. C., and C. Montell. 1993. *tramtrack* is a transcriptional repressor required for cell fate determination in the *Drosophila* eye. *Genes & Dev.* **7:**1085-1096.

Yang-Yen, H. F., J. C. Chambard, Y. L. Sun, T. Smeal, I. J. Schmidt, J. Drouin, and M. Karin. 1990. Transcriptional interference between c-Jun and the glucocorticoid receptor: mutual inhibition of DNA binding due to direct protein-protein interaction. *cell* **62:**1205-1215.

Zeng, C., J. Pinsonneault, G. Gellon, N. McGinnis, and W. McGinnis. 1994. Deformed protein binding sites and cofactor binding sites are required for the function of a small segment-specific regulatory element in *Drosophila* embryos. *EMBO J.* **13**:2362-2377.