Molecular and Cellular Biology, June 1998, p. 3384–3394 Vol. 18, No. 6
0270-7306/98/$04.00+0
Copyright © 1998, American Society for Microbiology

A Binding Site for Multiple Transcriptional Activators in the *fushi tarazu* Proximal Enhancer Is Essential for Gene Expression In Vivo†

WEI HAN,[1]‡ YAN YU,[1]§ KAI SU,[1] RONALD A. KOHANSKI,[2] AND LESLIE PICK[1]*

Brookdale Center for Developmental and Molecular Biology[1] and Department of Biochemistry,[2] Mt. Sinai School of Medicine, New York, New York 10029

Received 12 January 1998/Returned for modification 16 March 1998/Accepted 23 March 1998

The *Drosophila* homeobox gene *fushi tarazu* (*ftz*) is expressed in a highly dynamic striped pattern in early embryos. A key regulatory element that controls the *ftz* pattern is the *ftz* proximal enhancer, which mediates positive autoregulation via multiple binding sites for the Ftz protein. In addition, the enhancer is necessary for stripe establishment prior to the onset of autoregulation. We previously identified nine binding sites for multiple *Drosophila* nuclear proteins in a core 323-bp region of the enhancer. Three of these nine sites interact with the same cohort of nuclear proteins in vitro. We showed previously that the nuclear receptor Ftz-F1 interacts with this repeated module. Here we purified additional proteins interacting with this module from *Drosophila* nuclear extracts. Peptide sequences of the zinc finger protein Ttk and the transcription factor Adf-1 were obtained. While Ttk is thought to be a repressor of *ftz* stripes, we have shown that both Adf-1 and Ftz-F1 activate transcription in a binding site-dependent fashion. These two proteins are expressed ubiquitously at the time *ftz* is expressed in stripes, suggesting that either may activate striped expression alone or in combination with the Ftz protein. The roles of the nine nuclear factor binding sites were tested in vivo, by site-directed mutagenesis of individual and multiple sites. The three Ftz-F1–Adf-1–Ttk binding sites were found to be functionally redundant and essential for stripe expression in transgenic embryos. Thus, a biochemical analysis identified *cis*-acting regulatory modules that are required for gene expression in vivo. The finding of repeated binding sites for multiple nuclear proteins underscores the high degree of redundancy built into embryonic gene regulatory networks.

Embryonic development is controlled by the differential expression of regulatory genes. This process can be readily analyzed in the fruit fly, *Drosophila melanogaster*, since the basic body plan of the fly is established within the first few hours of embryogenesis by interacting sets of regulatory genes that are expressed in tightly regulated spatial and temporal patterns (1, 42, 44). One example of this is the homeobox gene *fushi tarazu* (*ftz*), which is expressed in a seven-stripe pattern in cellularizing embryos. The expression of *ftz* in these seven stripes directs the development of alternate body segments. In *ftz* mutant embryos, these segments are missing; the remaining portions of the embryo are fused, resulting in approximately half-sized embryos that die before hatching (20, 36, 38, 61). Similarly, ectopic expression of *ftz* throughout the embryo is lethal, demonstrating that *ftz* expression must be restricted to its seven-stripe domain for development to proceed normally (57).

Multiple regulatory elements dispersed over a distance of at least 10 kb interact to generate the *ftz* seven-stripe expression pattern (26, 27). One key regulatory element is the *ftz* upstream element (−3.6 to −6 kb), which contains two independently acting enhancers, the distal and proximal enhancers, that play roles in both stripe establishment and stripe maintenance (47, 64). These enhancers act in an orientation-independent fashion to direct *ftz*-like striped expression patterns in transgenic embryos via a heterologous promoter. Thus, the *ftz* enhancers serve as models for studying the control of cell-type-specific gene expression in an in vivo situation. One function of the upstream element enhancers is to mediate positive autoregulation that is thought to be necessary for maintenance of *ftz* stripes (26, 47, 52). This autoregulatory function is revealed when the enhancers are fused upstream of a minimal promoter and a *lacZ* reporter gene: expression of the reporter genes is dependent upon the presence of wild-type Ftz protein in transgenic embryos. Ftz, like other homeodomain-containing proteins, binds to DNA in a sequence-specific fashion and is thought to control embryonic development by regulating the expression of downstream target genes (11, 14, 46, 47). Strong evidence that the upstream element enhancers are direct targets of Ftz in vivo has been provided elsewhere (52). Ftz protein binds to approximately 20 sites in the upstream element. While deletion of individual Ftz sites had no effect on *ftz-lacZ* fusion gene expression, deletion of all Ftz binding sites abolished *lacZ* fusion gene expression in transgenic animals. When some of these binding sites were changed to those recognized by another homeodomain protein, Bicoid (Bcd), gene expression was not detected. However, when compensating changes in the Ftz homeodomain that would allow it to bind to Bcd sites were made, striped expression was restored. These studies demonstrate that Ftz protein can activate transcription through the upstream element and provide the strongest evidence to date that a native target site of a *Drosophila* homeodomain protein has been identified. However, Ftz protein alone is not sufficient to direct striped expression and Ftz

* Corresponding author. Mailing address: Brookdale Center for Developmental and Molecular Biology, Mt. Sinai School of Medicine, Box 1126, One Gustave L. Levy Place, New York, NY 10029. Phone: (212) 241-1966. Fax: (212) 860-9279. E-mail: pick@msvax.mssm.edu.

† Publication no. 213 from the Brookdale Center for Molecular Biology.

‡ Present address: Memorial Sloan-Kettering Cancer Center, New York, NY 10021.

§ Present address: Cornell University Medical College, New York, NY 10021.

Downloaded from http://mcb.asm.org/ on November 20, 2019 by guest

binding sites alone do not generate stripes in transgenic embryos (43, 47). Other factors, which may interact with *cis*-regulatory sequences within the enhancers and/or with Ftz protein itself, are required to modulate the activity of Ftz in vivo.

Previously, we defined a core 323-bp region of the *ftz* proximal enhancer that directs *ftz*-like stripe expression in transgenic embryos (21). We refer to this region as the 323-bp *ftz* proximal enhancer (323 fPE), since the full-length proximal enhancer from which it was derived directs striped expression in an orientation-independent fashion via a heterologous promoter in transgenic embryos, suggesting that it acts as a classical enhancer (47). The 323 fPE contains five binding sites for Ftz protein that were identified with purified Ftz homeodomain (47). In addition, nine binding sites for at least 10 nuclear proteins in *Drosophila* embryo extracts were identified within the 323 fPE (21) (Fig. 1A). Some of these proteins bound to multiple sites in the enhancer, while others bound to a single site. For example, Tramtrack (Ttk) protein (23, 24) binds to at least five sites in the 323 fPE, while a unique protein complex was generated with binding site 3 (BS3) in the enhancer (see Fig. 1 and reference 21) for numbering of these sites). In addition, a number of binding sites interacted with multiple proteins, forming discrete complexes in gel retardation assays; for example, at least three complexes were generated with BS6 (21) (see below). These three complexes were also generated with BS8 and BS9, suggesting that a minimum of three proteins interact repeatedly with these three binding sites.

As described above, the same cohort of nuclear factors interacts with three binding sites in the 323 fPE. To identify the proteins in these complexes, large amounts of *Drosophila* embryo nuclear extract were subjected to DNA affinity chromatography. The transcription factors Adf-1 (13) and Ttk (24) were each identified by N-terminal sequencing. The peptide sequence of a third, as-yet-uncharacterized enhancer binding protein (protein N), was also obtained. In addition, the nuclear hormone receptor protein Ftz-F1 (60) was previously shown to bind to this site. Adf-1 and Ftz-F1 each activate transcription via this binding site, suggesting potentially redundant roles in activating the expression of the *ftz* gene. To assess the function of the nuclear factor binding sites in vivo, expression of proximal enhancer-*lacZ* fusion genes carrying point mutations in individual binding sites was monitored in transgenic embryos. This mutagenesis demonstrated that the three Ftz-F1–Adf-1–protein N–Ttk binding sites are functionally redundant in vivo and are absolutely required for expression of reporter genes in stripes. These observations reveal a surprising degree of redundancy at the level of regulation of pair-rule gene stripe expression in vivo.

MATERIALS AND METHODS

Purification of *ftz* proximal enhancer binding proteins. Embryo nuclear extracts were prepared as described elsewhere (21, 54) with modifications suggested by the work of Wamplar et al. (62). In addition, we observed that DNA binding activity was recovered after embryos were stored in 50% glycerol at −20°C for up to several months. Oregon R embryos (0 to 12 h) were collected, dechorionated, and washed. All remaining steps were carried out at 4°C, and all buffers contained 50 mg of soybean trypsin inhibitor per ml, 1 mM benzamidine, 1 mM dithiothreitol (DTT), 2 U of aprotinin per ml, 1 mg of antipain per ml, and 1 mg of bacitracin per ml. Embryos were homogenized by one pass through a Yamato LH-21 homogenizer at 1,500 rpm in homogenization buffer (0.35 M sucrose, 15 mM HEPES [pH 7.6], 0.1 M KCl, 2.5 mM $MgCl_2$, 1 mM EDTA, 10% glycerol) at a concentration of 4 ml of buffer/g of embryo. The homogenate was filtered through three layers of nylon cloth (6.3-mm mesh), and 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 [pH 7.6], 0.1 M KCl, 3 mM $MgCl_2$, 0.1 mM EDTA, 10% glycerol) at a concentration of 1 ml of buffer/g of embryos in 70Ti tubes. A 1/10 volume of a saturated ammonium sulfate solution, pH 7.9, was added, and the tubes were rotated for 30 min. Following centrifugation 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 centrifugation for 20 min at 10,000 rpm. The pellet was resuspended in sample buffer (homogenization buffer without sucrose) at a concentration of 1 ml of buffer/10 g of embryo and dialyzed against the sample buffer without KCl for approximately 1 h. Samples were stored at −70°C. Protein concentration was determined by the method of Bradford (4) with a Bio-Rad protein assay solution. The protein concentration of the extracts was typically 20 to 30 mg/ml. A total of 1,900 mg of protein was obtained from pooled nuclear extracts derived from 1,300 g of embryos.

Approximately 1,584 mg of nuclear extract was diluted to 10 mg/ml with the 0.2 M NaCl HEG buffer (25 mM HEPES [pH 7.6], 0.1 mM EDTA, 10% glycerol, and protease inhibitors used as in the homogenization buffer). The diluted nuclear extract was loaded onto a 140-ml heparin-Sepharose CL-6B column which was washed with 0.2 M HEG and step eluted sequentially with 0.4 and 1.0 M HEG. The active 0.4 M fractions and 1.0 M fractions were pooled separately and precipitated by the addition of solid ammonium sulfate to 40% saturation. The 0.4 M heparin fraction (H0.4) contained 303 mg of protein, and the 1.0 M heparin fraction (H1.0) contained 99 mg of protein. The two protein fractions were frozen immediately at −80°C.

Each fraction was passed over a BS4 DNA affinity column before three rounds of DNA affinity chromatography (34) with BS6. This step was designed to remove proteins that interact specifically with BS4 as well as nonspecific proteins that might interact with any oligonucleotide. Columns were prepared by coupling to CNBr-Sepharose with binding sites flanked by *Hin*dIII sites to allow for concatamerization. Binding site sequences are shown in Fig. 5B. Heparin fractions were individually diluted to 0.2 M NaCl with no-salt HEG. Nonspecific DNA competitors, poly(dI-dC) (50 mg/mg of protein) and single-stranded oligonucleotides (BS7 upper strand, 10 mg/mg of protein), were added, and each sample was incubated on ice for 10 min. Following centrifugation at 10,000 × *g* for 10 min, the supernatant was loaded onto the BS4 column. The pass-through from the BS4 column was directly loaded onto the BS6 column. Following washes with 0.1 M HEG, activity was eluted from the BS6 column with 0.8 M HEG. The eluate fraction was diluted to 0.2 M NaCl, and nonspecific competitors, poly(dI-dC) and the single-strand oligonucleotides used above, were added. The mixture was treated as described above and loaded onto the BS6 affinity column a second time. Following washing and elution as described above, a third round of affinity chromatography was carried out.

Purification was monitored in gel retardation assays as described previously (21). Silver staining of sodium dodecyl sulfate (SDS)-polyacrylamide gels followed standard protocols. For microsequencing of the purified polypeptides, the BS6 third eluate (approximately 8 ml) was precipitated with 4 volumes of cold acetone (−20°C) for 18 h at −80°C. After centrifugation at 12,000 × *g* for 30 min, the pellet was resuspended in SDS sample buffer (6 mM EDTA, 62.5 mM Tris-HCl [pH 6.8], 6% SDS, 20% glycerol, 1% [wt/vol] bromphenol blue). The sample was electrophoresed through an SDS–10% polyacrylamide gel. After the gel was stained with Coomassie brilliant blue R-250 in dH_2O, protein bands were excised. The gel slice was loaded into the wells of a second SDS–15% polyacrylamide gel. The wells were then filled with V8 protease (0.1 mg) in SDS sample buffer. The protein in each gel slice was partially digested during electrophoresis. Following electrotransfer to polyvinylidene difluoride, the membrane was stained with Coomassie brilliant blue R-250 and individual bands were excised. Peptide sequence was then obtained from these proteolytic fragments with a Porton gas-phase sequencer with on-line phenylthiohydantoin (PTH) amino acid detection and quantitation, with the manufacturer's programs.

Characterization of DNA binding activity by denaturation and renaturation. The renaturation experiment was performed essentially as described elsewhere (5, 13). Eluate (0.2 ml) from the second round of DNA affinity chromatography was precipitated with 0.8 ml of cold acetone. The proteins were separated by SDS–10% polyacrylamide gel electrophoresis (PAGE). Prestained protein molecular weight markers were run on the gels. The gel was cut into four regions. Each gel slice was crushed with a Teflon pestle in an Eppendorf tube containing 0.25 ml of elution buffer (0.1% SDS, 50 mM Tris-HCl [pH 7.6], 0.1 mM EDTA, 5 mM DTT, 0.1 mg of bovine serum albumin per ml, 0.15 M NaCl) and incubated for 1 h at room temperature. After the supernatant was recovered by a brief spin for 1 min in an Eppendorf centrifuge, the gel was washed with 0.1 ml of elution buffer. Recovered supernatant (0.3 ml) was precipitated with 1.2 ml of acetone. The pellet was resuspended in 10 μl of 6 M guanidine-HCl made in dilution buffer (25 mM HEPES [pH 7.6], 0.1 M NaCl, 0.5 mM DTT, 10% glycerol, 0.1 mg of bovine serum albumin per ml) and incubated for 30 min at room temperature. Dilution buffer (0.5 ml) was added and incubated for 50 min at room temperature. The renatured protein was concentrated in a Centricon-30 filter (Amicon) from 0.5 ml to 50 μl. Finally, 50 μl of renatured protein was brought to 100 μl by adding 50 μl of dilution buffer. Eight microliters was used in one reaction mixture for the gel retardation assay.

Identification and purification of Adf-1 expressed in *Escherichia coli* with a pET-3a system. *Adf-1* cDNA in plasmid expression vector pET-3a (Adf-1/pET-3a [12]) was transformed into a host *E. coli* strain, BL21(DE3)plysS (Novagen). Five milliliters of overnight culture of the transformed cells was added to 500 ml of Luria broth and grown for 2 h. The expression of Adf-1 was induced with 0.2 mM isopropyl thiogalactopyranoside (IPTG) at 37°C for 3 h. The cells were collected by centrifugation and resuspended in 20 ml of lysis buffer (20 mM

Downloaded from http://mcb.asm.org/ on November 20, 2019 by guest

HEPES [pH 7.6], 0.1 M NaCl, 1 mM DTT, 0.1 mM phenylmethylsulfonyl fluoride). They were disrupted by sonication six times for 30 s repeatedly with a 5-mm probe at a 3.5 output, on ice. Proteins (180 mg) in the supernatant were loaded onto 1 ml of a BS6-Sepharose DNA affinity column equilibrated with 0.1 M NaCl lysis buffer. The column was washed with 10 ml of lysis buffer and eluted with 8 ml of 0.8 M NaCl lysis buffer. Eight milliliters of the eluate was concentrated and desalted to 1 ml (0.3 mg of protein) with a Centriprep-10 concentrator (Amicon). The purified proteins were analyzed by SDS–12% PAGE and confirmed by Western blotting with a polyclonal anti-Adf-1 antibody (from R. Tjian).

Transcriptional activation function of Adf-1 and Ftz-F1. To express Adf-1 in yeast (*Saccharomyces cerevisiae*), a full-length *Adf-1* cDNA from plasmid pBSADFa was subcloned into yeast expression vector pADNS, which contains an Adh (*alcohol dehydrogenase* gene) promoter and terminator and a *LEU2* selective marker (10). The *Adf-1* cDNA fragment was released from the vector pBSADFa by digestion with *Bgl*II and *Not*I and inserted into the *Not*I site of pADNS. The *Not*I sites of the fragment and vector were first ligated. The *Bgl*II site of the fragment and the other *Not*I site of the vector were then filled in with the DNA polymerase large fragment (Klenow).

A *Hin*dIII/*Not*I fragment containing a *ftz-f1* cDNA was isolated from a pBluescript SK vector (kindly provided by J. P. Gergen) and inserted into the *Hin*dIII/*Not*I sites of pADNS. The BS6 reporter genes were constructed by insertion of an oligonucleotide containing three tandem copies of either the wild-type or the mutant binding site (Fig. 5B) flanked by *Not*I sites into pSH201 (kindly provided by S. Hanes). This plasmid contains a basal CYC1 promoter upstream of the *lacZ* gene and a *URA3* selective marker. Reporter genes were transformed into yeast strain W3031A by standard methods. Either Ftz-F1/pADNS or Adf-1/pADNS was then transformed into the reporter strains. β-Galactosidase activity was measured for three individual transformants by a liquid colorimetric assay (2). Briefly, the yeast cell extracts were prepared by glass bead disruption, and *o*-nitrophenyl-β-D-galactopyranoside (ONPG; Sigma) solution was added. After incubation in 30°C for 2 h, β-galactosidase units were calculated as optical density at 420 nm (OD_{420}/OD_{595}) where OD_{595} was measured with Bio-Rad reagent to determine protein concentration (Bio-Rad Laboratories). The mean units and standard deviations of the β-galactosidase activity for each construct were calculated from the results of three individual transformants.

Site-specific mutagenesis. Site-specific mutagenesis was performed by overlap extension PCRs with *Taq* DNA polymerase (Promega) (29, 56). The DNA template used was a *ftz* proximal enhancer fragment covering positions 2168 to 2574 (*Dra*I-*Xba*I fragment) subcloned in Bluescript KS$^+$. Fixed primers were from the polylinker of the KS vector with *Xba*I recognition sites at the 5′ end. Sequences were 5′-TCTAGAAACAGCTATGACCATG-3′ (upper strand) and 5′-TCTAG AGTAAAACGACGGCCAGT-3′ (lower strand). Nine pairs of mutant oligonucleotides were used as internal primers. Sequences are given in Fig. 5B. Products from PCRs were digested with *Xba*I and subcloned back into the KS vector. These were then used as templates for the next round of site-specific mutagenesis to generate multiple site-specific mutations (Fig. 5A). The sequences of the mutant DNA constructs in the Bluescript KS$^+$ vectors were confirmed by double-stranded DNA sequencing with sequencing-grade *Taq* DNA polymerase (Promega) (30, 51).

P-element-mediated transformation and analysis of transgenic embryos. Enhancer fragments were inserted into the *Xba*I site of the P-element transformation vector HZ50PL, which contains a basal *hsp70* promoter and an *E. coli lacZ* reporter gene (26). Fusion genes were injected into *rosy*^{506} embryos with helper plasmid p25.1WC according to the method of Rubin and Spradling (50). 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 (19). Briefly, dechorionated 0- to 12-h embryos were fixed with formaldehyde. The fixed embryos were incubated with anti-β-galactosidase antibodies (Cappel). After washing, the embryos were incubated with anti-rabbit immunoglobulin G biotinylated antibodies. Avidin-biotin-peroxidase complex reagents (Vector Laboratories) were added, providing avidin-mediated biotinylated peroxidase coupling to the biotinylated antibodies. Staining was carried out in the presence of diaminobenzidine, $NiCl_2$, and H_2O_2. Expression was also monitored by X-Gal (5-bromo-4-chloro-3-indolyl-β-D-galactopyranoside) staining, by standard methods (data not shown). Embryos were photographed with Nomarski optics on a Zeiss Axiophot microscope with TMAX-100 film.

RESULTS

Purification and identification of *ftz* enhancer binding proteins. Our previous studies identified nine binding sites for *Drosophila* nuclear proteins within the 323 fPE that directs *lacZ* fusion gene expression in a seven-stripe pattern in transgenic embryos. These binding sites are referred to by binding site number (BS1 to -6 and -8 to -10 [see Fig. 1A]). These studies suggested that the same set of proteins interacts with BS6, -8, and -9, forming three discrete complexes with 0- to 9-h *Drosophila* embryo nuclear extracts, *ftz* enhancer binding complexes 8, 9, and 10 (fEBC8, -C9, and -C10 [see Fig. 1B and 3A]). To identify the protein components of these complexes, *Drosophila* nuclear extract was prepared from 0- to 16-h embryos. Binding activities precipitated between 10 and 30% ammonium sulfate saturation were separated on a heparin-Sepharose column, which was sequentially eluted with 0.4 and 1 M NaCl. After precipitation with ammonium sulfate, each fraction was subjected to nonspecific DNA affinity chromatography followed by three rounds of sequence-specific DNA affinity chromatography with an oligonucleotide corresponding to BS6 (O6 [Materials and Methods]). As shown in Fig. 1B, while fEBCs 8 and -C9 were highly enriched in the 1 M heparin-Sepharose and first affinity fractions, these complexes were detected only weakly after multiple rounds of affinity chromatography. However, two other complexes with higher mobilities were enriched in these fractions (X and Y [Fig. 1B]). The apparent change in migration of these complexes may be due to proteolysis, despite the inclusion of a large number of protease inhibitors in these fractions, or may result from separation of protein components of the complex during affinity chromatography. The binding activities in the final affinity fraction were specific; binding was competed by a 50-fold molar excess of BS6 itself (Fig. 1B, lane 6) but not by an unrelated sequence (BS10 [Fig. 1B, lane 7]). The protein components of the active fractions were identified by SDS-PAGE and silver staining (Fig. 1C). Three major polypeptides were present in the final affinity fraction with apparent molecular masses of 29 kDa (P1), 34 kDa (P2), and 53 kDa (P3).

To determine which of the proteins in the affinity fractions had DNA binding activity, a portion of the second affinity fraction was separated on an SDS-polyacrylamide gel, which was then cut into four slices as indicated on the right of Fig. 1C. Proteins were renatured (Materials and Methods) and used for DNA binding assays. As shown in Fig. 1D, proteins from region 1 (lane 1) and region 2 (lane 2) did not interact with BS6. Proteins from region 3 (lanes 3 and 4) and region 4 (lanes 5 and 6) interacted strongly with BS6 but only very weakly with an unrelated binding site from the proximal enhancer (BS2 [21]). The migration of these two complexes was significantly faster than that of any of the complexes formed with the native affinity fractions. Thus, it is likely that the complexes detected with native extract contain additional proteins that do not bind directly to DNA. These would have been separated from the DNA binding moieties on denaturing SDS gels.

The results presented above suggest that polypeptides P1 and/or P2 as well as P3 harbor DNA binding activities that are specific for BS6. To identify these proteins, partial peptide sequence was obtained following digestion with V8 protease (Materials and Methods). Peptide sequences for P1 and P3 did not match any known sequences in database searches. Analysis of P3 (protein N [below]) is in progress. For P1, a cDNA was isolated with degenerate oligonucleotides based upon the N-terminal peptide sequence. The 3′ half of this 930-bp cDNA is homologous to human, rat, pea, *Caenorhabditis elegans*, and yeast ribosomal protein L9 (9, 28, 32, 41, 58). While it is likely that this protein was isolated by DNA affinity chromatography because of nonspecific interactions with nucleic acids, preliminary studies suggest that this protein can bind to double-stranded DNA in a sequence-specific fashion (data not shown and reference 20a). For P2, the N-terminal sequence matched the N-terminal amino acids of the transcription factor Adf-1. An unambiguous sequence, M-D-K-L-D-A-N-L-E-Q-Q-X-D-P-N, was obtained, with an initial yield of 0.9 pmol. This sequence differs from the published conceptual translation of Adf-1 by one amino acid at position 14 of our sequence. This

Downloaded from http://mcb.asm.org/ on November 20, 2019 by guest

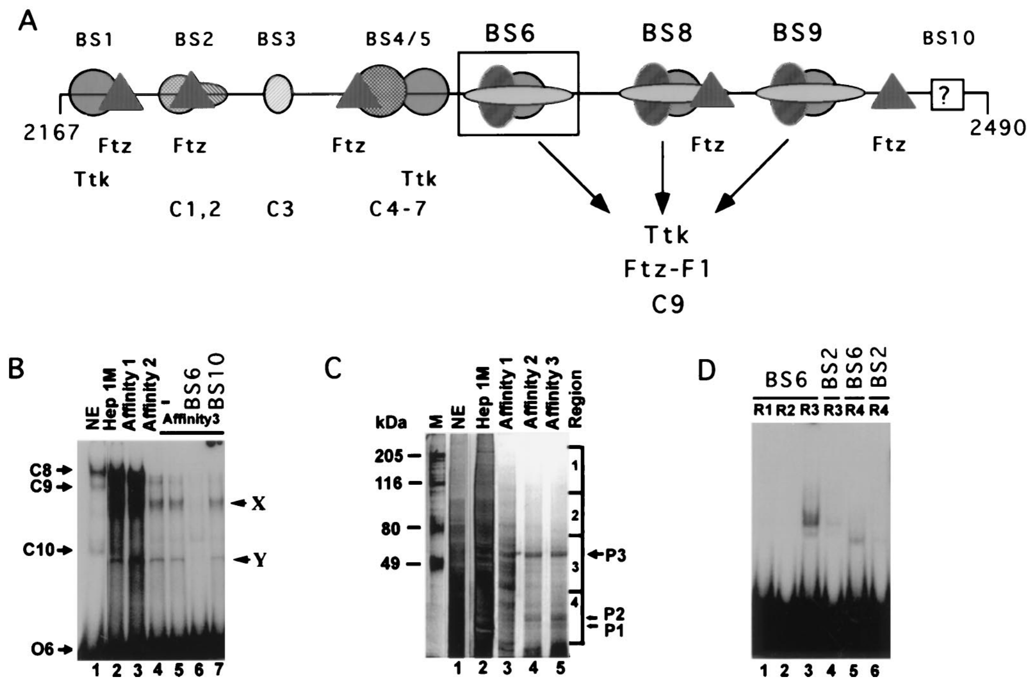

FIG. 1. Purification of *ftz* enhancer binding proteins. (A) A schematic representation of proteins interacting with the *ftz* proximal enhancer is shown, adapted from the work of Han et al. (21). Nine binding sites for *Drosophila* nuclear proteins are indicated above the line (BS1 to -6 and BS8 to -10). Binding sites for Ftz protein (47), Ttk, and Ftz-F1 are indicated below the line. BS6 was used for DNA affinity chromatography, as described below, and is boxed in the figure. Complexes observed in gel retardation assays with *Drosophila* embryo extracts (21) are indicated as follows: complexes C1 and C2 with BS2; complex C3 with BS3; complexes C4 to C7 with BS4; and complex C9 with BS6, -8, and -9. Note that the same pattern of complex formation was found with BS6, -8, and -9, suggesting that they interact with the same cohort of nuclear factors. (B) Gel retardation assays, carried out as described previously (21) with a ^{32}P-labeled oligonucleotide corresponding to BS6 (O6), were used to monitor DNA binding activity during the course of purification. Lane 1, 4 μg of nuclear extract and 2 μg of poly(dI-dC); lane 2, 2 μg of 1 M heparin fraction and 0.2 μg of poly(dI-dC); lane 3, 2 μl of BS6 first eluate and 5 ng of poly(dI-dC); lane 4, 2 μl of BS6 second eluate and 1 ng of poly(dI-dC); lanes 5 to 7, 2 μl of BS6 third eluate and no poly(dI-dC). To test for specificity, a 50-fold molar excess of BS6 (lane 6) or BS10 (lane 7) was included in binding reaction mixtures. (C) Proteins from each purification step were electrophoresed through an SDS–10% polyacrylamide gel. A photograph of the silver-stained gel is shown. NE, 10 μg of nuclear extract; Hep 1M, 10 μg of 1 M heparin fraction; Affinity 1, 20 μl; Affinity 2, 50 μl; Affinity 3, 100 μl; M, protein markers. 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 for the gel shown in panel D. (D) Renatured proteins (8 μl) from regions 1 (R1), 2 (R2), 3 (R3), and 4 (R4) were used in gel retardation assays with ^{32}P-labeled BS6 oligonucleotides or an unrelated site, BS2, as indicated above the lanes. Proteins from R3 and R4 interacted strongly with BS6 (lanes 3 and 5) and very weakly with BS2 (lanes 4 and 6). P1 and P2 were both present in R4. It is likely that only P1 activity was renatured, since P2 (Adf-1) did not interact at all with BS2 (see below), but P1 did interact with this binding site (data not shown). No DNA binding activities were detected with proteins from R1 and R2 (lanes 1 and 2).

is presumed to be due to a mistake in the DNA sequence, which reports a CTC triplet, encoding a leucine, rather than a CCC triplet, which would encode the proline residue that we obtained. Furthermore, the molecular mass of P2 and that of Adf-1 (12) appear to be the same, 34 kDa, as estimated by SDS-PAGE. This factor was previously purified on the basis of its binding to the *alcohol dehydrogenase* (*Adh*) promoter (12, 13) and will be discussed in more detail below.

Proteins present in the 0.4 M heparin-Sepharose fraction were analyzed as described above for the 1 M fraction (data not shown). The DNA binding activity necessary for formation of fEBC10 was retained in this fraction. We showed previously that fEBC10 contained Ttk protein (21, 23, 24). Consistent with this, an exact match to the amino acid sequence of Ttk protein from positions 334 to 348 was obtained by peptide sequencing, with an initial yield of 2.9 pmol. The purified Ttk protein appeared to be truncated, since its apparent molecular mass on SDS-polyacrylamide gels was 45 kDa, while the full-length protein has a predicted molecular mass of 69 kDa.

In sum, four proteins appear to bind specifically to a repeated nuclear factor binding site—BS6, BS8, or BS9—in the 323 fPE (Fig. 1A). This repeated module is now referred to as the FANT binding site, representing the binding of Ftz-F1, Adf-1, protein N (as yet unidentified; P3 above), and Ttk.

Recombinant Adf-1 interacts with multiple sites in the proximal enhancer. To examine the interaction of Adf-1 with the *ftz* proximal enhancer, recombinant Adf-1 was expressed in *E. coli* and partially purified by BS6 affinity chromatography (Materials and Methods). DNA binding activity of the partially purified recombinant Adf-1 protein was tested in gel retardation assays with ^{32}P-labeled BS6 (Fig. 2). Adf-1 generated a single complex in gel retardation assays (Fig. 2A). The shifted band was abolished by addition of anti-Adf-1 antibody to the reaction mixtures (lane 2) but not by the addition of preimmune serum (lane 3).

To test the DNA binding specificity of Adf-1, varying amounts of unlabeled wild-type BS6 or an oligonucleotide with point mutations in the binding site (mutant site 6 [M6] [see Fig. 5B]) were added to binding reaction mixtures. The Adf-1–BS6 complex (lane 1) was inhibited by the addition of unlabeled BS6 at a 50- or 100-fold molar excess and was abolished at a 250-fold molar excess (lanes 2 to 4). M6, tested at the same concentrations (lanes 5 to 7), did not inhibit complex formation. No DNA binding activity was detected with bacterial protein isolated from cells which did not express recombinant Adf-1 (lane 8). These results indicate that Adf-1 binds specifically to site 6 in the 323 fPE.

To determine whether Adf-1 can interact with any of the

Downloaded from http://mcb.asm.org/ on November 20, 2019 by guest

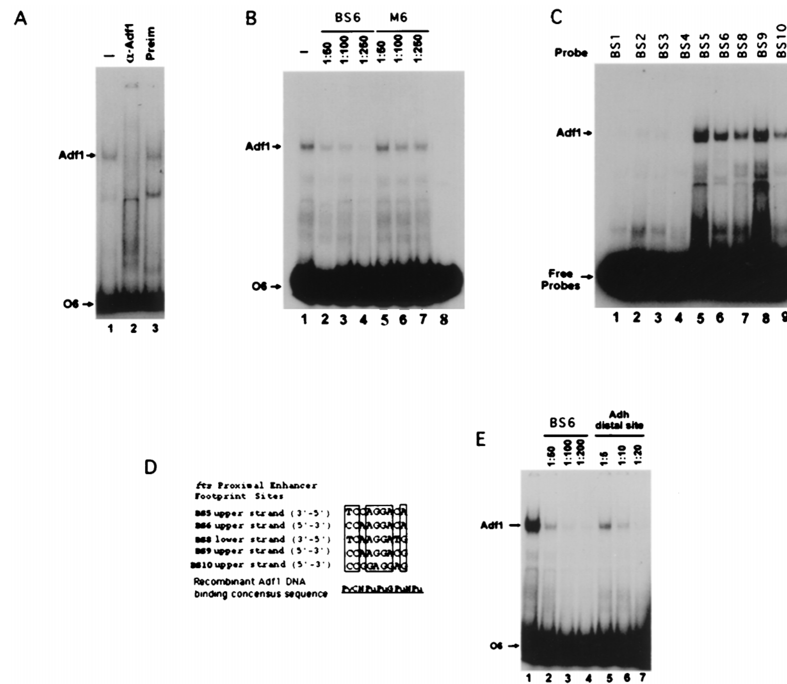

FIG. 2. Recombinant Adf-1 protein interacts with multiple sites in the *ftz* proximal enhancer. (A) Recombinant Adf-1 (approximately 0.3 μg) was used in a gel retardation assay with BS6 (O6). An autoradiogram is shown. The Adf-1–O6 complex in lane 1 was abolished by addition of 1 μl of anti-Adf-1 antibody (α-Adf-1 [lane 2]) but was not affected by addition of 1 μl of preimmune serum (Preim [lane 3]). The band below Adf-1 in all three lanes was a nonspecific species. Free probe is indicated (O6). (B) The specificity of Adf-1 binding was tested in gel retardation assays by competition with unlabeled wild-type BS6 and mutant BS6 (M6). Recombinant Adf-1 (approximately 0.3 μg) was incubated with ^{32}P-labeled probe (BS6) in the presence of increasing amounts of unlabeled BS6 and M6, as indicated in the figure. An autoradiogram is shown. The major Adf-1–BS6 complex (lane 1) indicated by the arrow was abolished by addition of a 250-fold molar excess of unlabeled BS6 (lanes 2 to 4) but not by equivalent amounts of M6 (lanes 5 to 7). As a control, proteins purified with BS6 affinity chromatography from *E. coli* transformed with the pET-3a vector only were tested; no interaction was detected (lane 8). (C) The recombinant Adf-1 (approximately 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, as indicated above the autoradiogram (Fig. 1A). The same amount of probe was used in each reaction mixture (approximately 10 fmol). Of the nine probes, five probes corresponding to BS5, BS6, BS8, BS9, and BS10 interacted with Adf-1 (arrow). (D) The recombinant Adf-1 DNA binding consensus sequence was deduced from the five footprint sites (BS5, BS6, BS8, BS9, and BS10). The strand and orientation of the sequences are indicated. Py, pyrimidine (T or C); Pu, purine (A or G); N, any nucleotide. (E) Binding of Adf-1 to the *ftz* proximal enhancer and that to the Adh distal promoter sites were compared. Recombinant Adf-1 (approximately 0.3 μg) was used in a gel retardation assay with ^{32}P-labeled BS6 probe. An Adf-1 binding site from the *Adh* distal promoter (Adh distal site) was synthesized according to the reported sequence 5′-AGCTGCTGC TGCATCCGTCGACGTCGACTGCACTCGCCCC-3′ (13). Different amounts of unlabeled BS6 and Adh distal site were added to the binding reaction mixtures as competitors. The major Adf-1–O6 complex (lane 1) was inhibited by a titration of unlabeled BS6 at 50-, 100-, and 200-fold molar excesses (lanes 2, 3, and 4). It was inhibited to the same extent by a titration of the Adh distal site at 5-, 10-, and 20-fold molar excesses (lanes 5, 6, and 7).

other binding sites identified in the proximal enhancer, gel retardation assays were carried out with nine different ^{32}P-labeled oligonucleotides corresponding to the nine binding sites identified with *Drosophila* nuclear extracts. Five of the nine oligonucleotides interacted with Adf-1 (Fig. 2C): BS5, -6, -8, -9, and -10 interacted with Adf-1 (lanes 5 to 9). Four sites, 1, 2, 3, and 4, did not interact detectably with the protein (lanes 1 to 4). The consensus DNA binding sequence deduced from the five Adf-1 binding sites is (C/T)CN(A/G)(G/A)G(A/G)N(G/A) (Fig. 2D). This consensus sequence cannot be found in the four binding sites which did 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. (13). This observation suggested that Adf-1 might bind to the two sequences with different affinities. To test this idea, an oligonucleotide corresponding to an Adf-1 binding site in the *Adh* distal promoter (termed Adh distal site [13]) was synthesized. Different amounts of unlabeled BS6 and Adh distal site were added to DNA binding reaction mixtures containing Adf-1 and ^{32}P-labeled BS6 (Fig. 2E). The Adh distal site (lanes 5 to 7) was 10-fold more effective than BS6 itself (lanes 2 to 4) in the inhibition of Adf-1–BS6 complex formation (lane 1, control). The results indicate that Adf-1 binds to the Adh distal site with 10 times higher an affinity than it does to binding sites in the *ftz* proximal enhancer.

Native Adf-1 interacts with BS6. As described above, BS6 formed three complexes with *Drosophila* embryo nuclear extract. To determine which of these complexes contains Adf-1, gel retardation assays were performed with an anti-Adf-1 antibody. As shown in Fig. 3A, three fEBCs were detected (lane 1), as observed previously. Addition of the anti-Adf-1 antibody to binding reaction mixtures specifically abolished the formation of fEBC8 (lane 2). Addition of the preimmune serum to binding reaction mixtures did not affect the formation of any of the fEBCs (lane 3). This suggests that fEBC8 contains Adf-1 protein. Formation of complex fEBC8 was detected with nuclear extracts with BS6, -8, and -9 but not with any other sites in the proximal enhancer. Competition analysis with oligonucleotides corresponding to each of the nine nuclear extract binding sites in the 323 fPE demonstrated that this complex was specific and formed only with these three binding sites (21). This would suggest that native Adf-1 protein interacts with only these three sites, a subset of the binding sites detected with recombinant protein (Fig. 2C). However, it was also possible that Adf-1 could form complexes with other binding sites that have different specificities and different mobilities

Downloaded from http://mcb.asm.org/ on November 20, 2019 by guest

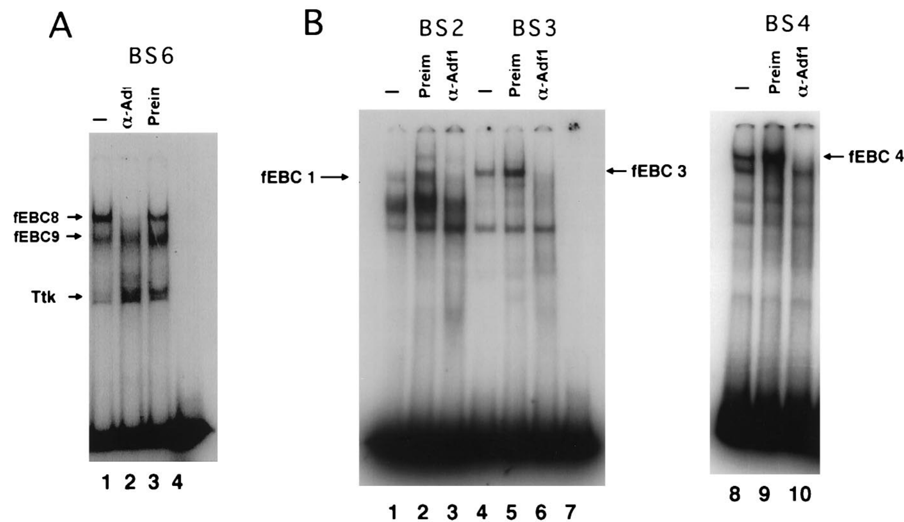

FIG. 3. Native Adf-1 is present in several fEBCs. (A) DNA binding reactions were carried out with 0- to 12-h embryo nuclear extract and ^{32}P-labeled BS6. fEBC8, -C9, and -C10 were formed with nuclear extract, as indicated elsewhere (Fig. 1) (21). fEBC8 was inhibited by addition of anti-Adf-1 antibody (α-Adf-1) to the binding reaction mixture (lane 2). It was not affected by addition of preimmune serum (Preim [lane 3]). fEBC9 and Ttk were not affected by addition of either antibody or preimmune serum. Lane 4, no protein. (B) Binding reactions were carried out with 0- to 12-h embryo nuclear extract and ^{32}P-labeled BS2 (lanes 1 to 3), BS3 (lanes 4 to 6), or BS4 (lanes 8 to 10). Lane 7, no protein. Preimmune serum (lanes 2, 5, and 9) or anti-Adf-1 antibody (lanes 3, 6, and 10) was included in binding reaction mixtures as indicated. The addition of anti-Adf-1 antibody specifically inhibited the formation of one fEBC in each case, as indicated by the arrows.

in gel retardation assays because of interactions with other proteins in the nuclear extract.

To test this, anti-Adf-1 antibody was added to binding reactions carried out with three additional oligonucleotides from the 323 fPE (21). BS5, which interacted with recombinant Adf-1, was not examined in this analysis because it interacts only with Ttk in *Drosophila* nuclear extract and this interaction was unaffected by addition of anti-Adf-1 antibody (Fig. 3A). Site 10, which also showed an interaction with recombinant protein, was not included because this site does not form any specific complex with *Drosophila* nuclear extracts (21). As shown in Fig. 3B, binding reactions with ^{32}P-labeled BS2, -3, or -4 were carried out in the absence of antibody (lanes 1, 4, and 8), in the presence of preimmune serum (lanes 2, 5, and 9), or in the presence of anti-Adf-1 antibody (lanes 3, 6, and 10). Surprisingly, in each case, one complex formed with each oligonucleotide was abolished by addition of the antibody (see arrows). For each oligonucleotide, several complexes are formed with nuclear extract, as published previously (21). However, in each case, only one of these complexes was specifically inhibited by addition of the antibody. Thus, it appears that native Adf-1 protein participates in the formation of complexes with BS2, -3, -4, and -6. This result suggests that the specificity and/or affinity of Adf-1 protein for DNA is modulated in vivo by interactions with another nuclear protein(s).

Adf-1 and Ftz-F1 activate transcription via BS6. As summarized above, Adf-1 binds directly to BS6 in the 323 fPE. In a previous study, we showed that Ftz-F1 and Ttk also bind to this module (21), and here we showed that Ttk protein was also purified by BS6 DNA affinity chromatography. While the expression pattern of Ttk suggests that it might repress *ftz* transcription, both Adf-1 and Ftz-F1 are present in early embryos at the time that *ftz* is expressed in seven stripes and are thus potential activators of *ftz* transcription (see Discussion). To test whether these proteins have the potential to activate transcription via this binding site, assays were carried out in yeast cells (Fig. 4). Three copies of either wild-type BS6 or mutant M6 were placed upstream of a minimal promoter and the *E. coli lacZ* reporter gene. Ftz-F1 or Adf-1 was expressed in cells under the control of a strong *Adh* promoter. Gene expression was monitored by measuring β-galactosidase units with a liquid spectrophotometric assay. As shown in Fig. 4A, both Ftz-F1 and Adf-1 activated the transcription of the *lacZ* gene. Ftz-F1 increased the reporter gene expression by 107-fold, while Adf-1 activated transcription more weakly, increasing reporter gene expression by 26-fold. Transcriptional activation by each protein was binding site dependent, since neither protein activated reporter gene expression via the mutant binding sites (M6).

To test whether these proteins have the ability to activate transcription via the full-length 323 fPE and to ensure that activation was not influenced by the basal promoter used for the *lacZ* fusion genes described above, a second set of transformations was performed (Fig. 4B). The 323 fPE was placed upstream of the yeast *HIS3* gene, along with its basal promoter (63). After integration of this reporter gene into the genome, Adf-1 was expressed in yeast cells. Activation of reporter gene expression was measured by testing for cell growth on plates lacking histidine and containing increasing amounts of the competitive inhibitor of the *HIS3* gene product, 3-aminotriazole (3-AT) (35). Expression of Adf-1 supported growth in 2.5 mM 3-AT. To test for specificity, concatamerized homeodomain binding sites (11) were placed upstream of the same reporter gene. In contrast to Adf-1, which did not activate expression, the homeodomain protein Ftz activated expression of this reporter (Fig. 4B, right panel). We showed previously that Ftz-F1 strongly activated gene expression via the 323-bp core proximal enhancer, allowing for growth in up to 25 mM 3-AT (63). This difference in activation levels is consistent with that seen with the concatamerized BS6 sites upstream of the *lacZ* reporter gene. Thus, Ftz-F1 and Adf-1 are potential redundant regulators of *ftz* expression in vivo, each of which interacts with the FANT module to activate transcription, at least in a yeast system.

Functional analysis of nuclear protein binding sites in vivo. To test the role of the FANT module, as well as the other individual nuclear protein binding sites identified previously (Fig. 1A), in directing stripe expression in vivo, site-specific

Downloaded from http://mcb.asm.org/ on November 20, 2019 by guest

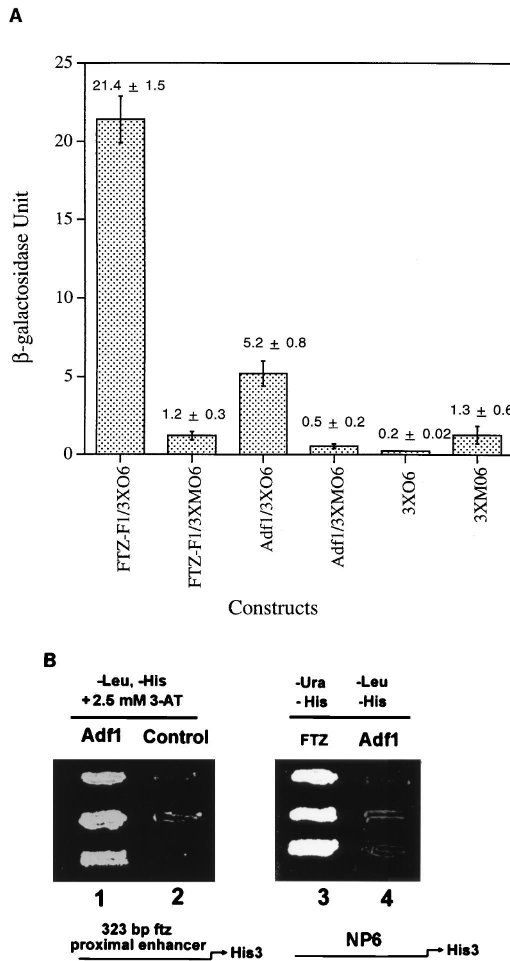

FIG. 4. Adf-1 and Ftz-F1 each activate transcription via the *ftz* proximal enhancer in a binding site-dependent fashion. (A) Adf-1 or Ftz-F1 was expressed in yeast cells along with *lacZ* reporter plasmids containing three copies of either wild-type BS6 (3X06) or mutant site 6 (3XMO6), as indicated. β-Galactosidase levels were determined as described in Materials and Methods for at least three independent transformants. (B) Adf-1 activates transcription via the *ftz* proximal enhancer. The *Adf-1* cDNA was expressed in a yeast strain carrying a 323 fPE-*HIS3* fusion gene as a reporter as indicated (lanes 1 and 2). Three Adf-1 transformants grew well in the absence of histidine and in the presence of 2.5 mM 3- AT (lane 1). Three colonies transformed with expression vector alone did not grow on the same plate (lane 2). To test specificity, a second reporter strain carrying six copies of a homeodomain binding site (NP6) upstream of the *HIS3* reporter was used (lanes 3 and 4). Three transformants expressing Ftz protein grew well in the absence of histidine (lane 3). Three Adf-1 transformants did not grow on this plate (lane 4). The Adf-1 expression plasmid carried a *LEU2* selective marker; the Ftz expression plasmid carried a *URA3* marker.

mutations were generated by overlap extension PCRs (Materials and Methods). The template for these mutations was the 406-bp fragment of the *ftz* proximal enhancer that directs *lacZ* fusion gene expression in a pattern that is identical to, although somewhat weaker than, that of the full-length proximal enhancer (21). As shown in Fig. 5A, for those sites previously shown to contact unique factors, mutations were generated in single binding sites; mutations in multiple binding sites were generated in the FANT module and then sequentially in BS4, BS1, and BS5, which also contact Adf-1 (BS4 [Fig. 3]) and/or Ttk (BS1 and BS5 [Fig. 1A]). Point mutations were generated based upon results of methylation interference assays carried out previously (21). Typically, three base changes were introduced into individual binding sites (Fig. 5B). Oligonucleotides

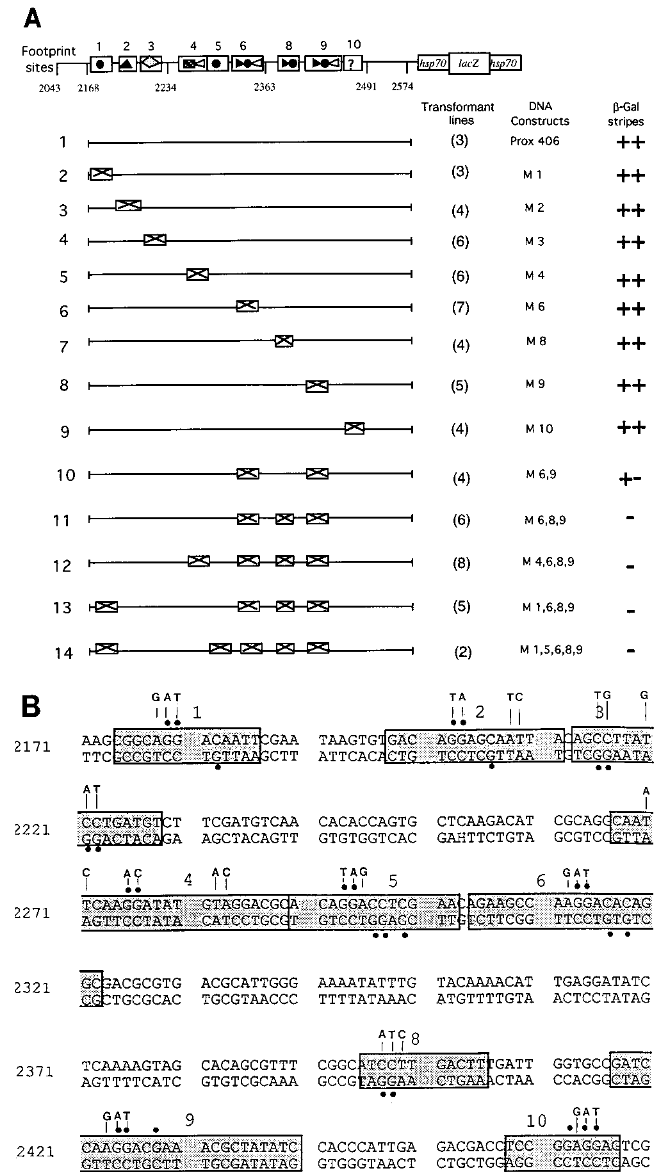

FIG. 5. Site-specific mutagenesis of nuclear factor binding sites in the *ftz* proximal enhancer. (A) A schematic of the *ftz* proximal enhancer *lacZ* fusion gene is shown at the top (positions 2043 to 2574 of the upstream element). *ftz* enhancer fragments were inserted into the P-element transformation vector HZ50PL, which contains a basal *hsp70* promoter upstream of the *E. coli lacZ* gene. Nine binding sites for *Drosophila* embryo nuclear factors identified previously are indicated by the boxes. Footprint sites are numbered 1 to 6 and 8 to 10, according to reference 21. Fusion gene 1 (Prox 406) directs *lacZ* fusion gene expression in a *ftz*-like seven-stripe pattern (21). This fragment was used as a template for introducing mutations into nuclear factor binding sites. These are indicated schematically below the binding site(s) for fusion genes 2 to 14. Fusion genes were inserted into the *Drosophila* germ line by P-element-mediated transformation, and expression was detected by immunohistochemical staining with anti-β-galactosidase antibodies or with X-Gal. The number of transformant lines examined and the name of each fusion gene are indicated on the right. A summary of fusion gene expression levels is indicated at the far right (Fig. 6). (B) The sequence of the core *ftz* proximal enhancer (positions 2171 to 2470) is indicated. Nuclear factor binding sites are boxed and numbered. Methylation interference analysis was used to identify bases in close contact with nuclear proteins (21); these are indicated by dots. These bases were the primary targets of mutagenesis. The sequences of mutations introduced into each footprint site are indicated for one strand above the enhancer sequence.

Downloaded from http://mcb.asm.org/ on November 20, 2019 by guest

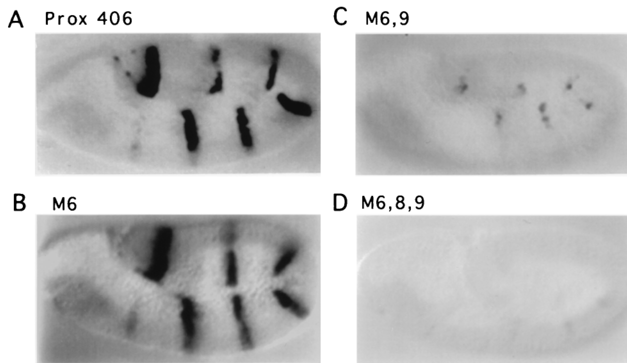

FIG. 6. Functionally redundant binding sites are necessary for *ftz-lacZ* fusion gene expression. The expression pattern of *ftz* proximal enhancer-*lacZ* fusion genes was monitored with anti-β-galactosidase antibodies in transgenic embryos following P-element-mediated transformation. Representative embryos are shown for four fusion genes. (A) Prox 406 is expressed in seven *ftz*-like stripes. (B) M6. Site-specific mutation of BS6 had no detectable effect on striped expression. (C) M6 and M9. Simultaneous mutation of sites 6 and 9 drastically decreased levels of expression in stripes. (D) M6, M8, and M9. Mutation of BS6, -8, and -9 virtually abolished *lacZ* fusion gene expression. Embryos were photographed at the end of germ band extension and are oriented as follows: anterior, left; dorsal side, up.

carrying these altered sequences were tested in gel retardation assays with *Drosophila* embryo extracts to confirm that the mutations abolished DNA binding in vitro (data not shown). In addition, the point mutations generated in the FANT sites abolish binding of bacterially expressed Ftz-F1, Adf-1, and Ttk to this site. Based on these experiments, wild-type or mutant fusion genes carrying the 323 fPE fragments upstream of a basal promoter and the *lacZ* reporter gene were integrated into the *Drosophila* genome by P-element-mediated transformation.

Mutations in any one of the nine binding sites in the 323 fPE had no detectable effect on *lacZ* fusion gene expression in the embryo, illustrated for one of the nine individual mutations in Fig. 6B (compare to the wild-type fusion gene [Fig. 6A]). A double mutation in two FANT sites, BS6 and BS9, drastically decreased levels of fusion gene expression. Examination of embryos derived from four independent lines showed that expression in two lines was extremely weak (Fig. 6C) while expression in the other two lines was undetectable (data not shown). An additional mutation in the third FANT site, BS8, virtually abolished expression in all lines examined (Fig. 6D). Fusion genes containing mutations in additional binding sites (sites 6, 9, 8, and 4; sites 6, 9, 8, and 1; or sites 6, 9, 8, 1, and 5) were also not detectably expressed in transgenic embryos (data not shown).

Since FANT and repeated Ftz sites are each necessary for gene expression, we have tested whether these sites are sufficient to generate a striped pattern with three copies of a neighboring Ftz-FANT site (BS8) upstream of a basal promoter in HZ50PL (64b). However, thus far, we have been unable to detect any expression of this reporter gene. Further analysis is in progress.

In summary, mutations in single binding sites had no detectable effect on gene expression in vivo. However, mutations in the repeated FANT module had a strong effect on *lacZ* fusion gene expression, demonstrating that these sites are functionally redundant but absolutely required for *lacZ* fusion gene expression. Each of the nuclear factors that interacts with this module is a candidate activator of *ftz* gene expression in vivo. Previous studies demonstrated that multiple Ftz protein binding sites in the enhancer are similarly repeated, redundant, and necessary for *ftz* enhancer-*lacZ* fusion gene expression in embryos (47, 52). Together, these results suggest that the *ftz* gene is controlled by multiple *cis*-acting modules that act redundantly to direct stripe expression.

DISCUSSION

The *ftz* proximal enhancer is a complex eukaryotic enhancer which directs reporter gene expression in a *ftz*-like seven-stripe pattern via a heterologous promoter in transgenic fly embryos (47, 52). One *trans*-acting factor that interacts directly with the enhancer in vitro and in vivo is Ftz protein itself, which maintains gene expression via positive autoregulation (26, 47). Ftz protein interacts with five sites in the core 323 fPE (Fig. 7). Simultaneous mutation of these Ftz binding sites abolishes *lacZ* fusion gene expression in transgenic embryos (52, 53). Thus, these Ftz binding sites represent one repeated *cis*-acting module necessary for gene expression in vivo.

While intact Ftz binding sites are necessary for autoregulation, these sites alone do not direct striped expression in vivo (43). In addition, the observation that the upstream distal and proximal enhancers mediate autoregulation differentially in the ectodermal and mesodermal primordia suggested the existence of cofactors that limit the domains of Ftz activity in vivo (47). Thus, while Ftz protein is necessary for proximal enhancer-directed autoregulation, it is not sufficient. In vitro DNA binding studies revealed a surprisingly diverse array of proteins that interacted with the *ftz* proximal enhancer (21). Nine sites for 10 factors were identified with *Drosophila* embryo nuclear extracts. These sites were clustered within a 323-bp region of the enhancer that was shown to direct *lacZ* fusion gene expression in an *ftz*-like seven-stripe pattern, as does the full-length proximal enhancer. Here we have shown that a second repeated module—nuclear factor BS6, -8, and -9—within the proximal enhancer is necessary for *ftz-lacZ* fusion gene expression in vivo (Fig. 7). As with the repeated Ftz binding sites, these binding sites are redundant—mutation of any single site had no effect on gene expression. However, simultaneous mutation of multiple sites drastically decreased *ftz-lacZ* fusion gene expression. This redundancy is further highlighted by the fact that only two of the three site 6-type binding sites are conserved in a *Drosophila hydei ftz* proximal enhancer fragment (40) that directs striped expression in transgenic *D. melanogaster* embryos (33). Thus, *ftz* gene expression is buffered against random mutations that affect interactions with a single binding site by the presence of additional sites within the same regulatory element that interact with the same protein(s).

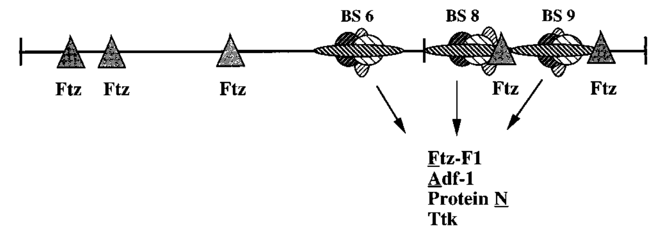

FIG. 7. Two essential *cis*-acting modules are located in a core 323 fPE. A schematic of the core *ftz* proximal enhancer (2167 to 2490) is shown. This regulatory element contains five binding sites for Ftz protein and three binding sites for a cohort of at least four nuclear factors, Ftz-F1, Adf-1, protein N, and Ttk (FANT). For each module, individual binding sites are redundant—mutation of any single Ftz or FANT site had no effect on *ftz-lacZ* fusion gene expression. However, each module is essential for gene expression since simultaneous mutations in all sites, for either module, abolished gene expression.

Downloaded from http://mcb.asm.org/ on November 20, 2019 by guest

More surprising than this redundancy at the level of binding sites is the apparent redundancy at the level of *trans*-acting factors that interact with the FANT module. We have found that at least four nuclear proteins interact with this module—Adf-1, Ftz-F1, Ttk, and an as-yet-unidentified protein (protein N) for which we obtained inconclusive N-terminal sequence information. This sort of redundancy at the level of *trans*-acting transcription factors has recently been shown for the apparently unrelated proteins Zeste and GAGA, which regulate expression of the *Ubx* gene (reference 37; see Discussion therein). For the *ftz* 323 fPE, two of the proteins identified to date—Adf-1 and Ftz-F1—are expressed in all somatic cells of preblastoderm and blastoderm-stage embryos, the time at which *ftz* stripes are activated (12, 63). Each of these proteins activates transcription via BS6, suggesting that each has the potential to positively regulate *ftz* stripe expression in vivo. Thus, the loss of expression observed when all three binding sites were mutated may be due to disruption of binding of Adf-1 or of Ftz-F1. Conversely, *ftz* enhancer-directed gene expression may survive the loss of either Adf-1 or Ftz-F1 proteins in the embryo, as long as the other is present. Analysis of mutations in the *ftz-f1* gene is consistent with this postulated redundancy since *ftz* stripes can still form in the absence of functional Ftz-F1 protein (18, 63). However, we have found that levels of *ftz* expression are reduced in *ftz-f1* mutant embryos and that the 323 fPE is not detectably expressed in the absence of maternal Ftz-F1 (57a). We are in the process of analyzing *adf-1* mutant embryos as well as *adf-1–ftz-f1* double mutant embryos to assess this in vivo. Finally, it is possible that Ftz-F1 and Adf-1 interact cooperatively to regulate *ftz* gene expression. As shown in Fig. 3A, anti-Adf-1 antibodies inhibited the formation of complex fEBC8 with nuclear extract. Previous studies showed that anti-Ftz-F1 antibodies inhibited the formation of the same complex (21), suggesting that it is composed of both proteins. Future studies will determine whether these two proteins independently regulate *ftz* expression via the FANT module or whether they interact in either a cooperative or a competitive fashion through these binding sites.

The three identified proteins interacting with the FANT module belong to three different classes of DNA binding proteins. Adf-1 was first identified as a transcription factor that interacted with the distal promoter of the *alcohol dehydrogenase* gene (*Adh*) (12, 13, 25). Other DNA binding sites for Adf-1 were found in the promoters of *Antennapedia* P1 and *dopa decarboxylase*. Adf-1 was shown to activate transcription in vitro and was postulated to be an activator of a larger number of genes during development. Adf-1 protein is not a member of any well-characterized class of transcription factors. Some homology was found to the DNA binding motif of Myb protein (12), and more recently, the *Drosophila stonewall* gene (8) has been identified as another potential member of this novel group of DNA binding proteins. In contrast to this, Ftz-F1 is a member of the larger class of nuclear hormone receptor superfamily proteins (39, 60) (reviewed in 59). Ftz-F1 has been shown to bind DNA as a monomer and to activate transcription in a number of systems (3, 45). Finally, Ttk is a zinc finger protein (24) that is unrelated to either Adf-1 or Ftz-F1.

It was proposed several years ago that Ttk acts as a repressor of *ftz* stripes since the protein is present before and after *ftz* is expressed in stripes but is not detected during the time that *ftz* is expressed in stripes (24). The proximal enhancer used in our studies contains multiple binding sites for Ttk. Therefore, it was our initial intention to test the role of Ttk as a repressor of *ftz* stripes by simultaneously mutating multiple Ttk binding sites. It was expected that fusion gene expression would initiate earlier and/or persist later in the absence of repression by Ttk. Fusion genes 12 and 13 (Fig. 5) carry mutations in four Ttk sites, while all five sites are mutated in fusion gene 14. However, three of the five Ttk binding sites overlap with binding sites for activator proteins that are necessary to activate expression of fusion genes (fusion gene 11 [Fig. 5]). Therefore, it was not possible to test whether Ttk represses through its proximal enhancer binding sites, since mutations result in loss of activation due to this overlap. Currently, the role of Ttk in regulating *ftz* is unclear. Mutation of Ttk binding sites in the zebra element resulted in premature activation of *ftz* gene expression, and ectopic expression of Ttk at later stages caused a decrease in *ftz* expression levels (6, 7, 49). However, given our observation that most Ttk binding sites also interact with other nuclear proteins, it is difficult to know whether these observations are a result of direct negative regulation of *ftz* by Ttk. Preliminary results from our laboratory suggest that, at least in yeast, Ttk can act as a transcriptional activator (64a), raising the possibility either that Ttk interacts with a corepressor to decrease *ftz* expression levels or that observed effects of Ttk overexpression in embryos are indirect. Analysis of *ftz* expression in embryos derived from germ line clones that lack both maternally deposited and zygotically expressed Ttk will be necessary to definitively resolve its role in vivo. Although we are attempting to generate such embryos, preliminary results from others suggest that this approach may not be feasible since their attempts to generate germ line clones with a number of *ttk* alleles have been unsuccessful (60a).

Autoregulatory elements similar to that described here for *ftz* have been characterized for both the pair-rule gene *even-skipped* (*eve*) and the homeotic gene *Deformed* (*Dfd*) (16, 22, 65)). *eve* and *ftz* are expressed in complementary striped patterns in the embryo (15). In each case, maintenance of stripes within defined domains is achieved by autoregulation utilizing multiple binding sites for either the Eve or the Ftz homeodomain protein. It is likely that the domains of action of other factors involved in autoregulation are limited by the preexisting striped pattern of the homeodomain protein itself. Thus, ubiquitously expressed factors such as Ftz-F1 and Adf-1, which appear to play roles in *ftz* autoregulation, would be available to positively regulate *eve* expression via its autoregulatory element in parallel with regulation of *ftz*. Multiple binding sites for nuclear factors were identified in the *eve* autoregulatory element by DNase I footprinting with *Drosophila* nuclear extracts (31). Some of these were thought to be binding sites for Ttk protein, which also interacts with other *eve* regulatory regions (48). However, the sequences of several of the binding sites in the *eve* autoregulatory element are similar to that of the FANT module, suggesting the possibility that Ftz-F1, Adf-1, and possibly Ttk coordinately regulate the expression of *eve* and *ftz* in stripes. For *Dfd*, which is expressed in a more restricted domain and slightly later than *ftz* and *eve*, an equally complex enhancer element that contains several essential regulatory sequences in addition to those that interact directly with Dfd protein has been identified (65). One factor that interacts with an essential region of the enhancer, DEAF-1, has recently been purified. This protein does not correspond to any of the *ftz* regulators identified to date (17). However, DEAF-1 is expressed throughout the embryo very early, prior to the onset of zygotic transcription (17), and so may regulate genes such as *eve* and *ftz* as well.

The *ftz* proximal enhancer is a complex eukaryotic enhancer which is composed of multiple modular elements that act in concert to direct a tightly controlled spatial and temporal pattern of gene expression. Similarly complex enhancers, with

Downloaded from http://mcb.asm.org/ on November 20, 2019 by guest

multiple binding sites for a number of proteins, were identified in classical studies of simian virus 40 and immunoglobulin gene expression (55, 66). The identification of binding sites in the *ftz* enhancer that are essential for gene expression in vivo, along with the isolation of a number of proteins that interact with these sites, provides the tools for functional analyses of the enhancer in an in vivo situation. These studies will provide insight into why eukaryotes have evolved such complex regulatory systems to direct the expression of single genes.

ACKNOWLEDGMENTS

We are grateful to Gene Coulter and Bob Tjian for providing Adf-1 cDNAs and antibodies. We thank Checco Ramirez for critical comments on the manuscript.

This work was supported by a grant from the National Institutes of Health (HD27937) to L.P.

ADDENDUM IN PROOF

While this paper was under review, Coulter et al. (Mol. Cell. Biol. **18**:2252–2261, 1998) reported that Adf-1 binds DNA with high affinity as a dimer. This likely explains the difference in affinity of Adf-1 for the *Adh* site (a dimer binding site) and the *ftz* site (a monomer binding site) shown in our Fig. 5E.

REFERENCES

1. **Akam, M.** 1987. The molecular basis for metameric pattern in the *Drosophila* embryo. Development **101:**1–22.
2. **Ausubel, F. M., R. Brent, R. E. Kingston, D. P. Moore, J. G. Seidman, J. A. Smith, and K. Struhl.** 1992. Current protocols in molecular biology. John Wiley & Sons, Inc., New York, N.Y.
3. **Ayer, S., N. Walker, M. Mosammaparast, J. P. Nelson, B. Z. Shilo, and C. Benyajati.** 1993. Activation and repression of *Drosophila* alcohol dehydrogenase distal transcription by two steroid hormone receptor superfamily members binding to a common response element. Nucleic Acids Res. **21:** 1619–1627.
4. **Bradford, M. M.** 1976. A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. Anal. Biochem. **72:**248–254.
5. **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.
6. **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.
7. **Brown, J. L., and C. Wu.** 1993. Repression of *Drosophila* pair-rule segmentation genes by ectopic expression of *tramtrack*. Development **117:**45–58.
8. **Clark, K. A., and D. M. McKearin.** 1996. The *Drosophila* stonewall gene encodes a putative transcription factor essential for germ cell development. Development **122:**937–950.
9. **Cohn, N. S.** 1992. *P. sativum* GA mRNA (clone F) similar to ribosomal protein L9. GenBank accession no. X65155.
10. **Colicelli, J., C. Birchmeier, T. Michaeli, K. O'Neill, M. Riggs, and M. Wigler.** 1989. Isolation and characterization of a mammalian gene encoding a high-affinity cAMP phosphodiesterase. Proc. Natl. Acad. Sci. USA **86:**3599–3603.
11. **Desplan, C., J. Theis, and P. H. O'Farrell.** 1985. The *Drosophila* developmental gene, *engrailed*, encodes sequence-specific DNA binding activity. Nature **318:**630–635.
12. **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. USA **89:**683–687.
13. **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.
14. **Florence, B., R. Handrow, and A. Laughon.** 1991. DNA-binding specificity of the fushi tarazu homeodomain. Mol. Cell. Biol. **11:**3613–3623.
15. **Frasch, M., R. Warrior, J. Tugwood, and M. Levine.** 1988. Molecular analysis of *even-skipped* mutants in *Drosophila* development. Genes Dev. **2:**1824–1838.
16. **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.
17. **Gross, C. T., and W. McGinnis.** 1996. DEAF-1, a novel protein that binds an essential region in a Deformed response element. EMBO J. **15:**1961–1970.
18. **Guichet, A., J. W. R. Copeland, M. Erdelyi, D. Hlousek, P. Zavorszky, J. Ho, S. Brown, A. Percival-Smith, H. M. Krause, and A. Ephrussi.** 1997. The nuclear receptor homologue Ftz-F1 and the homeodomain protein Ftz are mutually dependent cofactors. Nature **385:**548–552.
19. **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.
20. **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.

20a. **Han, W.** 1994. Ph.D. thesis. Mt. Sinai School of Medicine, New York, N.Y.

21. **Han, W., Y. Yu, N. Altan, and L. Pick.** 1993. Multiple proteins interact with the *fushi tarazu* proximal enhancer. Mol. Cell. Biol. **13:**5549–5559.
22. **Harding, K., T. Hoey, W. Warrior, and M. Levine.** 1989. Auto-regulatory and gap response elements of the *even-skipped* promoter of *Drosophila*. EMBO J. **8:**1205–1212.
23. **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.
24. **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.
25. **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.
26. **Hiromi, Y., and W. J. Gehring.** 1987. Regulation and function of the *Drosophila* segmentation gene *fushi tarazu*. Cell **50:**963–974.
27. **Hiromi, Y., A. Kuroiwa, and W. J. Gehring.** 1985. Control elements of the *Drosophila* segmentation gene *fushi tarazu*. Cell **43:**603–613.
28. **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.
29. **Innis, M. A., and D. H. Gelfand.** 1990. Optimization of PCRs in PCR protocols: a guide to methods and applications. Academic Press, Inc., San Diego, Calif.
30. **Innis, M. B., D. H. Myambo, D. H. 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. USA **85:** 9436–9440.
31. **Jiang, J., T. Hoey, and L. Levine.** 1991. Autoregulation of a segmentation gene in *Drosophila*: combinatorial interaction of the *even-skipped* homeo box protein with a distal enhancer element. Genes Dev. **5:**265–277.
32. **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.
33. **Jost, W., Y. Yu, L. Pick, A. Preiss, and D. Maier.** 1995. Structure and regulation of the *fushi tarazu* gene from *Drosophila hydei*. Roux's Arch. Dev. Biol. **205:**160–170.
34. **Kadonaga, J. T., and R. Tjian.** 1986. Affinity purification of sequence-specific DNA binding proteins. Proc. Natl. Acad. Sci. USA **83:**5889–5893.
35. **Kishore, G. M., and D. M. Shah.** 1988. Amino acid biosynthesis inhibitors as herbicides. Annu. Rev. Biochem **57:**627–663.
36. **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.
37. **Laney, J. D., and M. D. Biggin.** 1996. Redundant control of *Ultrabithorax* by *zeste* involves functional levels of zeste protein binding at the *Ultrabithorax* promoter. Development **122:**2303–2311.
38. **Laughon, A., and M. P. Scott.** 1984. Sequence of a *Drosophila* segmentation gene: protein structure homology with DNA-binding proteins. Nature **310:** 25–31.
39. **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.
40. **Maier, D., A. Press, and J. R. Powell.** 1990. Regulation of the segmentation gene *fushi tarazu* has been functionally conserved in *Drosophila*. EMBO J. **9:** 3957–3966.
41. **McCombie, W. R., J. M. Kelley, J. C. Venter, and C. A. Fields.** 1993. *Caenorhabditis elegans* cDNA clone CEESK81 similar to ribosomal protein L9. GenBank accession no. T00931.
42. **McGinnis, W., and R. Krumlauf.** 1992. Homeobox genes and axial patterning. Cell **68:**283–302.
43. **Nelson, H., and A. Laughon.** 1993. *Drosophila* glial architecture and development: analysis using a collection of new cell-specific markers. Roux's Arch. Dev. Biol. **202:**341–354.
44. **Nusslein-Volhard, C., and E. Wieschaus.** 1980. Mutations affecting segment number and polarity in *Drosophila*. Nature **287:**795–801.
45. **Ohno, C., H. Ueda, and M. Petkovich.** 1994. The *Drosophila* nuclear receptors FTZ-F1a and FTZ-F1b compete as monomers for binding to a site in the *fushi tarazu* gene. Mol. Cell. Biol. **14:**3166–3175.
46. **Percival-Smith, A., M. Muller, M. Affolter, and W. J. Gehring.** 1990. The interaction with DNA of wild-type and mutant *fushi tarazu* homeodomains. EMBO J. **9:**1209–1223.
47. **Pick, L., A. F. Schier, M. Affolter, T. Schmidt-Glenewinkel, and W. J. Gehring.** 1990. Analysis of the *ftz* upstream element: germ layer-specific en-

Downloaded from http://mcb.asm.org/ on November 20, 2019 by guest

hancers are independently autoregulated. Genes Dev. **4:**1224–1239.
48. **Read, D., T. Nishigaki, and J. T. Manley.** 1990. The *Drosophila even-skipped* promoter is transcribed in a stage-specific manner in vitro and contains multiple, overlapping factor-binding sites. Mol. Cell. Biol. **10:**4334–4344.
49. **Read, D. B.** 1992. Ph.D. thesis. Columbia University, New York, N.Y.
50. **Rubin, G. M., and A. C. Spradling.** 1982. Genetic transformation of Drosophila with transposable element vectors. Science **218:**348–353.
51. **Sanger, F., S. Nicklen, and A. R. Coulson.** 1977. DNA sequencing with chain-terminating inhibitors. Proc. Natl. Acad. Sci. USA **74:**5463–5467.
52. **Schier, A. F., and W. J. Gehring.** 1992. Direct homeodomain-DNA interaction in the autoregulation of the *fushi tarazu* gene. Nature **356:**804–807.
53. **Schier, A. F., and W. J. Gehring.** 1993. Analysis of a *fushi tarazu* autoregulatory element: multiple sequence elements contribute to enhancer activity. EMBO J. **12:**1111–1119.
54. **Soeller, W. C., S. J. Pook, and T. Kornberg.** 1988. In vitro transcription of the *Drosophila engrailed* gene. Genes Dev. **2:**68–81.
55. **Staudt, L. M., and M. J. Lenardo.** 1991. Immunoglobulin gene transcription. Annu. Rev. Immunol. **9:**373–398.
56. **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.
57. **Struhl, G.** 1985. Near-reciprocal phenotypes caused by inactivation or indiscriminate expression of the *Drosophila* segmentation gene *ftz*. Nature **318:** 677–680.
57a. **Su, K., M. Yussa, and L. Pick.** Unpublished data.
58. **Suzuki, K., J. Olvera, and I. G. Wool.** 1990. The primary structure of rat ribosomal protein L9. Gene **93:**297–300.
59. **Thummel, C. S.** 1995. From embryogenesis to metamorphosis: the regulation and function of *Drosophila* nuclear receptor superfamily members. Cell **83:** 871–877.
60. **Ueda, H., S. Sonoda, J. 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.
60a. **Vasisht, V., and S. Small.** Personal communication.
61. **Wakimoto, B. T., F. R. Turner, and T. C. Kaufman.** 1984. Defects in embryogenesis in mutants associated with the *Antennapedia* gene complex of *Drosophila melanogaster*. Dev. Biol. **102:**147–172.
62. **Wamplar, S. L., C. M. Tyree, and J. T. Kadonaga.** 1990. Fraction of the general RNA polymerase II transcription factors from *Drosophila* embryos. J. Biol. Chem. **265:**21223–21231.
63. **Yu, Y., W. Li, K. Su, W. Han, M. Yussa, N. Perrimon, and L. Pick.** 1997. The nuclear hormone receptor FTZ-F1 is a cofactor for the *Drosophila* homeodomain protein Ftz. Nature **385:**552–555.
64. **Yu, Y., and L. Pick.** 1995. Non-periodic cues generate seven *ftz* stripes in the *Drosophila* embryo. Mech. Dev. **50:**163–175.
64a. **Yu, Y., and L. Pick.** Unpublished data.
64b. **Yussa, M.** 1997. Ph.D. thesis. Universidade Federal do Rio de Janeiro, Rio de Janeiro, Brazil.
65. **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.
66. **Zenke, M., T. Grundstrom, H. Matthes, M. Wintzerith, C. Schatz, A. Wildeman, and P. Chambon.** 1986. Multiple sequence motifs are involved in SV40 enhancer function. EMBO J. **5:**387–397.

Downloaded from http://mcb.asm.org/ on November 20, 2019 by guest