The *Drosophila* Even-Skipped Promoter Is Transcribed in a Stage-Specific Manner In Vitro and Contains Multiple, Overlapping Factor-Binding Sites

DOUGLAS READ, TAKASHI NISHIGAKI†, AND JAMES L. MANLEY*

Department of Biological Sciences, Columbia University, New York, New York 10027

Received 16 March 1990/Accepted 15 May 1990

To investigate the factors contributing to regulation of expression of the *Drosophila* segmentation gene even-skipped (*eve*), we have analyzed both the in vitro transcription and *eve*-promoter-binding proteins in embryo extracts. We show that the *eve* promoter is accurately and efficiently expressed in nuclear extracts derived from *Drosophila* embryos and that transcription is more efficient in extracts prepared from embryos at early stages of development than in those from older embryos, broadly reproducing the temporal pattern of expression observed in vivo. This stage-specific expression is dependent on sequences upstream of the *eve* transcription start site which contain multiple binding sites for at least two distinct proteins present in embryo nuclei. One of these proteins, the binding sites for which correspond to the sequences required for stage-specific expression, appears to be the previously described GAGA factor. Although the binding activity of the GAGA factor remains constant, the level of the binding activity of the other protein, which we have called the TCCT factor, changes during the course of embryogenesis. Activity is first detected 3 to 5 h after fertilization and decreases during later stages of embryogenesis. We discuss the possibility that the TCCT factor plays a role in the maintenance or refinement of the *eve* expression pattern.

As a member of the pair-rule class of segmentation genes, even-skipped (*eve*) plays a critical role in the process of *Drosophila* embryogenesis. While some mutant *eve* alleles result in the typical pair-rule phenotype in which elements of alternate segments are missing, the null phenotype is a complete absence of evident segmentation (29, 30). By gastrulation, *eve* is expressed in a series of 14 alternating major and minor stripes. The positions of the seven major stripes correspond to the regions most severely disrupted in *eve* mutant embryos. Later in embryogenesis, the striped pattern decays, and *eve* expression is confined to a subset of cells in the central nervous system and in the anal plate (12, 21, 26). Analysis of the *eve* expression pattern in embryos with mutated segmentation genes has revealed that initial establishment of the seven major *eve* stripes requires the function of five of the known gap genes. Later refinement and maintenance of the major *eve* stripes is dependent on two other primary pair-rule genes, hairy and runt, and several lines of evidence indicate an *eve* autoregulatory function (12-14, 18, 20). The *eve* protein contains a homeo domain (12, 26) and has been shown to function as a negative regulator of transcription in *Drosophila* tissue culture cells (19, 25) and in vitro (5).

The expression of *eve* is thought to be controlled largely at the level of transcription initiation, and recent P-element transformation studies have led to the identification of specific regions within the *eve* promoter that are required for different aspects of *eve* expression (18, 20). Distinct and widely separated promoter elements have been shown to be required for the early expression of individual *eve* stripes, while a distal 200-base-pair (bp) sequence fused to the *hsp70* basal promoter is sufficient to direct late expression corresponding to all seven major *eve* stripes (20). At least some of

the sequence elements involved in both the early and late aspects of the *eve* expression are functionally dependent on the region immediately upstream of the transcription start site (18, 20). An understanding of how specific factors bound to various promoter elements interact with the basic transcriptional machinery to generate the complex *eve* expression pattern will require the development of cell-free systems that faithfully reproduce the in vivo states of *eve* promoter activity.

As a first step toward this goal, we have begun an in vitro analysis of factors interacting with the *eve* promoter and their roles in the initiation of *eve* transcription. Similar studies of other *Drosophila* genes with nuclear extracts from either cultured cells or embryos have led to the identification of a number of novel transcription factors, in addition to *Drosophila* homologs of mammalian transcription factors (3, 4, 7, 8, 11, 16, 17, 22, 23, 31-35, 39, 44, 45; see reference 6 for a review). In particular, nuclear extracts prepared from embryos have been used to show that the sequence immediately upstream of the engrailed (*en*) transcription start sites contains multiple protein binding sites. By oligonucleotide competition assays, it was shown that most of these sites are bound by the same protein, and that binding by this so-called GAGA factor is necessary for efficient *en* transcription in vitro and in tissue culture cells (39). Subsequently, it was shown that the Ultrabithorax (*Ubx*) promoter also contains multiple GAGA factor-binding sites and that the purified protein is able to specifically activate *Ubx* transcription in vitro (4). It has also been reported that the GAGA factor interacts weakly with the Antennapedia P2 (*Antp* P2) promoter (33) and that sequences important for transcription of the ecdysone-inducible *E74* promoter contain GAGA factor-binding sites (40). A protein with the same DNA-binding specificity and apparent molecular weight as the GAGA factor has recently been shown to bind to the *Drosophila* *hsp26* and *hsp70* promoters as well as to the promoter region between the divergently transcribed histone H3 and H4

* Corresponding author.

† Present address: Bio-Science Research Laboratories, Sankyo Co. Ltd., Tokyo, Japan.

genes (17). In this report, we present evidence that the GAGA factor binds multiple sites within the *eve* promoter and show that sequences containing these sites are important for efficient *eve* transcription, especially in extracts prepared from embryos at early stages of development.

We also demonstrate that three additional high-affinity binding sites in the *eve* promoter are bound by a developmental stage-specific factor that recognizes a sequence related to the consensus recognition sequence for the transcription factor Elf-1/NTF-1. Elf-1/NTF-1 consists of a set of polypeptides that bind important control elements in the *Ubx*, dopadecarboxylase (*Ddc*), and fushi tarazu (*ftz*) genes (7, 11). While three of the binding sites we have identified in the *eve* promoter contain a sequence that is related to the Elf-1/NTF-1 consensus sequence, we present evidence that these sites are bound by a protein, which we have called the TCCT factor, that is distinct from Elf-1/NTF-1.

MATERIALS AND METHODS

Plasmid constructions. The *eve* -574 5' deletion mutant was constructed by subcloning a 1.6-kilobase (kb) *Ssp*I-*Bam*HI *eve* genomic fragment that includes sequences between positions -574 and +903 with respect to the transcription start site (12, 20) into *Hinc*II-*Bam*HI-digested pUC19. BAL 31 digestion was initiated at the *Hind*III site in the polylinker of the -574 deletion construct, and following the addition of *Hind*III linkers and *Hind*III-*Bam*HI digestion, appropriately sized fragments were isolated from an agarose gel and reinserted into *Hind*III-*Bam*HI-digested pUC19. The BAL 31 deletion endpoints were determined by the dideoxynucleotide-sequencing method (36). The actin 5C-*eve* construct was a gift of K. Han (19), the d11/d13 mutant was provided by K. Harding and T. Hoey, and the p*Ubx* HIII Δ 5'-201 was provided by M. Biggin (4).

Extract preparation and fractionation. Embryo nuclear extracts were prepared by the method of Soeller et al. (39) as modified by Biggin and Tjian (4). Protein measurements were made by the method of Warburg and Christian (43) and were typically in the range of 12 to 15 mg/ml. In experiments in which either the transcriptional or DNA-binding activities of different extracts were compared, equal amounts of protein were included in the relevant reactions. To accumulate sufficient quantities of embryos for the preparation of staged extracts, appropriately aged embryos were occasionally stored on collection plates at 4°C for up to 2 days as described by Biggin and Tjian (4). Heparin-Sepharose (Pharmacia) chromatography was performed as described by Heberlein et al. (22). For Superose 6 (Pharmacia) gel filtration chromatography, 300 mg of the H.4 fraction (22) of a 1- to 12-h embryo extract was first precipitated with ammonium sulfate. The pellet was suspended in 10 ml of 20 mM HEPES (N-2-hydroxyethylpiperazine-N'-2-ethanesulfonic acid) (pH 7.6)-0.2 M ammonium sulfate-10% glycerol-1 mM dithiothreitol-0.2 M phenylmethylsulfonyl fluoride and applied to a Superose 6 Prep Grade column (2.6 by 70 cm) equilibrated with the same buffer. The column was run at a flow rate of 35 ml/h, and 15-ml fractions were collected. Fractions were dialyzed against the same buffer containing 0.1 M KCl instead of ammonium sulfate.

In vitro transcription and primer extension. Transcription reactions were performed in a total volume of 20 μ l of reaction mixture containing 100 μ g of nuclear extract protein, 5 μ g of *eve* template DNA, 10 μ g of actin 5C-*eve* template DNA, 20 μ g of pUC19 vector carrier DNA, and 500 μ M each of ATP, CTP, GTP, and UTP. Incubations were

performed for 60 min at 25°C. Transcription was stopped by the addition of 300 μ l of 50 mM Tris hydrochloride (pH 7.6)-0.15 M NaCl-0.5% sodium dodecyl sulfate-10 mg of yeast tRNA-50 μ g of proteinase K (Boehringer Mannheim Biochemicals). Following incubation for 30 min at 42°C samples were extracted and then were ethanol precipitated and washed with 80% ethanol. Primer extension was performed essentially as described by McKnight and Kingsbury (28). Dried RNA samples were resuspended in 20 μ l of 10 mM PIPES [piperazine-N,N'-bis(2-ethanesulfonic acid)] (pH 6.4)-400 mM NaCl containing 50 fmol of 5'-end-labeled 34-nucleotide primer. The primer used was complementary to the *eve* untranslated leader sequence from +57 to +90 and has the sequence 5'-gcgtctgtgattcaagtggcttattcagagg (12). Following 3 h of incubation at 60°C, 80 μ l of 50 mM Tris hydrochloride (pH 8.2)-5 mM MgCl₂-10 mM dithiothreitol-0.5 mM deoxyguanosine nucleoside triphosphates-25 μ l of dactinomycin per ml containing 10 U of avian myeloblastosis virus reverse transcriptase (Life Sciences, Inc.) was added, and the primer extension reaction was allowed to proceed for 1 h at 42°C. Following phenol-chloroform extraction and ethanol precipitation, samples were electrophoresed on 6% acrylamide-8 M urea sequencing gels.

DNase protection. Footprinting was performed essentially as described by Heberlein et al. (22). All DNA fragments were 5'-end labeled with [γ -³²P]ATP and polynucleotide kinase and were purified from acrylamide gels (27). Binding reaction mixtures consisted of 2 to 5 ng of labeled DNA fragment, 1 μ g of poly(dI · dC), and up to 40 μ g of the H.4 fraction protein (dialyzed to 0.1 M KCl) in a total volume of 40 μ l. The final buffer composition, including the contribution by the extract protein fraction, was 12.5 mM HEPES (pH 7.6)-50 mM KCl-6 mM MgCl₂-0.5 mM EDTA-5% glycerol. Incubations were performed for 20 to 30 min on ice; after incubation, an equal volume of 10 mM CaCl₂-5 mM MgCl₂ containing 10 to 25 ng of freshly diluted DNase I (Worthington Diagnostics) was added. DNase I digestion was continued for 5 min on ice before being stopped by the addition of 80 μ l of 50 mM Tris hydrochloride (pH 7.6)-0.15 M NaCl-0.5% sodium dodecyl sulfate-10 μ g of tRNA-25 μ g of proteinase K. Following 15 min of incubation at 42°C, samples were extracted and ethanol precipitated before electrophoresis on 7% acrylamide-8 M urea sequencing gels. For the oligonucleotide competition experiments, the double-stranded oligonucleotide and poly(dI · dC) were preincubated with the protein fraction for 20 min on ice before addition of the labeled DNA fragment and then incubated an additional 20 min. The 19-bp GAGA oligonucleotide has the same sequence as that used by Biggin and Tjian (4) to affinity purify the GAGA factor, except that *Hind*III overhanging ends were substituted for *Bam*HI ends. The sequences of the two strands are 5'-agcttaacagagagccaga and 5'-agcttctgctctctgtta. The 17-nucleotide TCCT oligonucleotide includes 12 bp in common with the 3' half of the *eve* site 8, and the two strands have the sequences 5'-ccggatggtcctgccga and 5'-agcttcggcaggaccat.

RESULTS

Transcription of *eve* in embryo nuclear extracts. To begin to analyze the complex set of interactions that are responsible for the regulated expression of the *eve* gene, we adopted the method of Soeller et al. (39) for the preparation of transcriptionally competent nuclear extracts from embryos. Extracts were initially prepared from unstaged 0- to 18-h embryos, and the template DNA was a cloned *Xba*I-*Bam*HI fragment

FIG. 1. Transcription of *eve* in embryo nuclear extracts. (a) *eve* transcription initiates accurately in vitro. Poly(A)⁺ RNA isolated from 2- to 3-h embryos and *eve* RNA transcribed in a 0- to 18-h embryo extract were analyzed by primer extension. The template DNA for in vitro transcription was an *eve* genomic fragment that extends from -6.4 to +0.9 kb relative to the transcription start site. The primer used was complementary to the *eve* leader sequence from nucleotide positions +90 to +57. Numbers to the left indicate sizes of marker DNA fragments in nucleotides. (b) Transcription of *eve* promoter 5' deletions in staged embryo extracts. In vitro transcription was performed as described in Materials and Methods. Quantitative primer extension analysis was used to compare the levels of transcription supported by seven deletion mutants in each of three staged embryo extracts. Endpoints of the 5' deletions relative to the transcription start site are indicated above each lane; the age of the embryo population from which each extract was prepared appears below each set of lanes. A truncated actin 5C promoter fused to the *eve* untranslated leader served as an internal control.

of *eve* genomic DNA that includes sequence from -6.3 kb to +903 bp relative to the transcription start site, +1 (12, 20). Figure 1a shows the results of a primer extension analysis of the 5' ends of *eve* RNA synthesized in an extract prepared from 0- to 18-h embryos and, for comparison, an identical analysis of poly(A)⁺ RNA isolated from 2- to 3-h embryos. The oligonucleotide primer used in this experiment extends from +57 to +90, and the primer extension products of both RNAs were of the expected size of 90 nucleotides (12). Transcription of the *eve* promoter was shown to be sensitive to low levels of α -amanitin, indicating that transcription was RNA polymerase II dependent (results not shown). We also determined that transcription initiated accurately from a number of other promoters, including *en*, *zerknüllt*, *fushi tarazu (ftz)*, *Ubx*, *hunchback*, *hsp70*, *actin 5C*, and the *copa* long terminal repeat (results not shown). Our standard transcription conditions (see Materials and Methods) reflect our findings that the inclusion of nonspecific DNA, such as pUC19, allowed efficient transcription of less than 1 μ g of *eve* template DNA per ml in a reaction mixture consisting of 40% nuclear extract. When the total DNA concentration in the reaction was maintained at 35 μ g/ml, the rate of transcription was approximately linearly dependent on template concentration up to about 25 μ g/ml. Interestingly, we found that while supercoiled *eve* templates were efficiently transcribed in both early- and late-stage extracts, transcription of linear templates was inefficient in early extracts as compared with late extracts. This differential activity was not correlated with any obvious differences in either the topology or the stability of linear templates in the two types of extract (results not shown). Although we do not understand the basis for this phenomenon, plasmid DNA preparations used in the experiments reported here were checked to ensure that >95% was in supercoiled form.

To locate sequences within the *eve* promoter that are important for transcription, a series of 5' deletion mutants was constructed with BAL 31 exonuclease. These mutants were used as templates for in vitro transcription in nuclear extracts prepared from *Drosophila* embryos at specific stages of development. A quantitative primer extension assay of such an experiment comparing the transcription of seven 5' deletion mutants in extracts from 1- to 5-h, 5- to 10-h, and 10- to 15-h embryos is shown in Fig. 1b. A template containing the *Drosophila* actin 5C promoter fused to the *eve* untranslated leader and most of the *eve* coding region served as an internal standard (19). Similar results were obtained in experiments using the *Drosophila* *hsp70* or adenovirus major late promoters as internal standards, ruling out the possibility that the actin promoter influenced *eve* expression (results not shown).

Progressive deletion of sequences between bp -179 and -72 resulted in stepwise reductions in the rate of transcription. Specifically, the -139 deletion mutant supported a level of transcription one-half as high as that supported by the -179 mutant. The mutant with an endpoint at -109 similarly displayed a reduced level of transcription relative to the -139 deletion mutant, as did the mutant with an endpoint at -72 with respect to the -109 deletion mutant. Deletion to -21 removes a TATA box consensus sequence and essentially abolished transcription. In none of the extracts did deletion of sequences between -6.3 kb and bp -179 have a significant effect on transcription (results not shown). This probably reflects a general distance effect limitation of the in vitro transcription assay, as has been observed by others (see, for example, references 37 and 42). In addition, deletion of the sequence between -39 and -72 had no effect, which argues against the possibility that the reduced transcription of the upstream deletion mutants

FIG. 2. Protein binding sites in the region surrounding the *eve* transcription start site. DNase I protection was used to detect DNA-binding activities present in the H.4 fraction of a 5- to 10-h embryo extract. Each DNA fragment was 5'-end labeled at *Hind*III sites introduced during the construction of 5' deletion mutants. The nucleotide position of the labeled end of each fragment is indicated by an asterisk. Unlabeled ends of the fragments at positions -296, +7, and +174 correspond to natural restriction sites. Lanes marked 0 indicate DNase digestion in the absence of protein; arrows represent increasing amounts of H.4 protein present in the binding reaction mixtures (3, 10, and 30 µg). Lanes marked M indicate purine cleavage sequence ladders of each fragment. Each footprint is numbered as referred to in the text and in subsequent figures.

resulted from bringing inhibitory vector sequences closer to the *eve* promoter. Thus, in addition to the TATA box, the *eve* promoter contains at least three sequence elements within 200 bp upstream of the transcription start site that are required for optimal transcription *in vitro*.

It is also apparent from Fig. 1b that while the actin 5C promoter and the -39 and -72 deletion constructs were transcribed at similar rates in all three extracts, templates containing additional upstream sequence were transcribed more efficiently in the 1- to 5-h and 5- to 10-h extracts than in the 10- to 15-h extract. This high level of transcription in early extracts conforms to expectations based on *eve* expression *in vivo*, which peaks between 2 and 10 h postfertilization (12, 26). The greater dependence on sequence elements between -72 and -179 for this efficient transcription in early extracts led us to suspect that these elements may represent binding sites for factors involved in the developmental regulation of *eve* expression.

Multiple protein-binding sites in the *eve* promoter. To determine whether the sequences defined as important for transcription by the *in vitro* assay do in fact represent binding sites for factors present in the embryo extracts, DNase protection assays were performed on fragments spanning the promoter region. Figure 2 shows the results of experiments in which increasing amounts of a 0.1 to 0.4 M KCl heparin-Sepharose fraction (referred to as H.4 by Heberlein et al. [22]) of a 5- to 10-h extract were incubated with

each of four different end-labeled promoter fragments. The 5- to 10-h extract fraction was used in initial experiments, since this extract supported high levels of *eve* transcription. Incubation mixtures included excess poly(dI · dC) to reduce nonspecific binding. Together, these four overlapping probes include the entire *eve* sequence between position -574 and +175. At least 12 sites were protected from DNase digestion within this region. These include a 14- to 15-bp footprint centered over the transcription start site and a footprint in the untranslated leader at +45. Although we have never observed protection of the TATA box itself (-23 to -28), a region immediately upstream, from -32 to -46, was consistently protected. Nine additional footprints were detected between this site and position -400. In contrast to this dense clustering of binding sites between -400 and +10, no significant DNase protection was detected upstream of -400 or downstream of +60. The *eve* promoter sequence and a summary of these results are presented in Figure 3a.

Two distinct binding activities protect *eve* promoter sequences. Footprints 1, 2, 3, 4, 5, 6, 9, and 10 each contain one or more copies of a sequence similar to the GAGA factor consensus sequence, 5'-GAGAG (4, 39), and footprints 7 and 8 each contain at least one copy of the sequence 5'-GGTCCTGC (TCCT sequence). In addition to the TCCT sequence, site 8 contains a copy of the related sequence 5'-GGTCCTGG, and site 11 includes another related sequence, 5'-GCTCCTGC. Thus, on the basis of their se-

FIG. 3. Nucleotide sequence of the *eve* promoter region and summary of protein-binding sites. (a) The sequence of the *eve* promoter from positions -400 to $+102$ is presented, with factor-binding sites indicated by brackets. Dotted lines at footprints 9 and 11 indicate that the footprint endpoints could not be precisely determined. GAGA factor consensus sequences are underlined, and TCCT factor consensus sequences appear in stippled boxes. The TATA box and translation initiation codon are shown in unstippled boxes, and the transcription start site is indicated by an arrow. The d11 and d13 deletions referred to in the text and in Fig. 5 are marked by dashed lines below the sequences deleted. (b) Diagrammatic representation of factor-binding sites. Factors are represented by ovals, and overlapping ovals indicate the binding of multiple factors at a particular site. Question marks (?) next to the symbols at sites 1 and 11 indicate that while the footprints include, respectively, GAGA factor and TCCT factor consensus sequences, footprint competition experiments were not performed on DNA fragments containing these sites.

quence content, it seemed possible that most of the footprints within the *eve* promoter represent binding by only two different proteins.

To test this hypothesis, DNase-protection assays in the presence of competitor DNAs were performed. Figure 4 shows the results of an experiment in which synthetic oligonucleotide competitors were preincubated with the H.4 fraction before incubation with a labeled DNA fragment and subsequent DNase I digestion. The 19-bp GAGA oligonucleotide contains the same sequence, 5'-GAGAG, used by Biggin and Tjian (4) to affinity purify the GAGA factor; the 17-bp TCCT oligonucleotide contains the sequence 5'-GGTCCTGC, which occurs within sites 7 and 8. Sites 4, 5, 6, 9, and 10 were effectively competed for by 800 ng of the GAGA oligonucleotide, indicating that protection at these sites was most likely due to GAGA factor. In contrast, protection at sites 7 and 8 was unaffected by the presence of GAGA oligonucleotide. Other experiments in which the samples were not run so far on the gel revealed that protection at sites 2 and 3 was also inhibited by the GAGA

oligonucleotide but not by the TCCT oligonucleotide. Similar results were obtained when a fragment of the *en* promoter that contains multiple GAGA factor-binding sites (39) was used as a competitor (results not shown). Figure 4 also shows that sites 7 and 8 were competed for by 200 ng of the TCCT oligonucleotide, while sites 4, 5, 6, 9, and 10 were unaffected.

We conclude from these results that the majority of DNase-protected sites in the *eve* promoter represent binding sites for two distinct nuclear proteins. At least seven of these, sites 2, 3, 4, 5, 6, 9, and 10, are probably bound by the previously described GAGA factor. Although site 1 was not tested, we believe that it also represents a binding site for GAGA factor. Another protein, which we provisionally refer to as the TCCT factor, is responsible for protection at sites 7 and 8 and possibly at site 11. Figure 3b summarizes these results.

Protection at some GAGA sites is due to binding by more than one protein molecule. Footprints 4, 6, and 9 each contain at least two copies of a sequence similar to the GAGA factor

FIG. 4. DNase protection in the presence of competitor oligonucleotides. Double-stranded oligonucleotides containing either the sequence 5'-GAGAG (GAGA) or the sequence 5'-GGTCCTGC (TCCT) were preincubated with the H.4 fraction of a 0- to 12-h extract prior to addition of the labeled DNA fragment. See Materials and Methods for the entire sequence of the oligonucleotides. The amount of oligonucleotide included in the preincubation reaction mixture is indicated in nanograms. Lanes marked 0 represent DNase digestion in the absence of protein; the arrow represents increasing amounts of H.4 protein present in the binding reactions (2.5, 10, and 40 μ g), and the bars (—) indicate 40 μ g of H.4 protein. Footprints 7 and 8 contain the TCCT sequence and are marked by heavy brackets; footprints 4, 5, 6, 9, and 10 each contain at least one copy of a sequence similar to 5'-GAGAG and are marked by light brackets. The DNA fragment was 5'-end labeled at the *Xho*I site at position +7 (*) and extends to -394. The lane marked M represents the purine cleavage ladder of the DNA fragment.

consensus sequence. In addition, the extent of the protected sequence at these sites is correspondingly greater than at sites 1, 2, 3, 5, and 10. These observations suggest the possibility that footprints 4, 6, and 9 represent binding sites for multiple GAGA factor proteins. To test this idea, a DNase protection experiment was performed with a DNA fragment containing two internal deletions that remove portions of sites 4 and 6. The results of this experiment are shown in Fig. 5. The 13-bp deletion from -115 to -127 overlaps the 5' end of site 4 and removes the two 5'-most bp of the putative GAGA factor recognition sequence 5'-CGCTC. The 11-bp deletion from -167 to -177 removes an entire GAGA factor recognition sequence (5'-CGCTC) in the 5' portion of site 6. The effect of both of these deletions was to reduce the size of the respective footprints, as would be

FIG. 5. Two sites are bound by multiple GAGA factor proteins. The panel on the left shows DNase protection of a wild-type (wt) *eve* promoter fragment; the panel on the right shows protection of a promoter fragment containing two internal deletions, d11 and d13. d11 removes 11 bp that include the 5'-GAGCG (noncoding strand sequence) in site 6, and d13 deletes 13 bp, including the two 3'-most bp in the sequence 5'-GAGCG (noncoding strand) within site 4. See Fig. 3 for boundaries of the d11 and d13 deletions. Both fragments were 5'-end labeled at position -394 and extend to +7. Arrows indicate increasing amounts of H.4 fraction from a 10- to 15-h extract (8, 16, and 32 μ g, from left to right). Hatched ovals indicate putative GAGA factor recognition sequences.

expected if each deletion destroyed the recognition site for a single protein factor. The extent of protection at site 4 was reduced from 25 to 18 bp and was centered over the remaining GAGA recognition sequences. Interestingly, the degree of protection at the truncated site 4 appeared to be greater than at the corresponding wild-type site. Footprint 6 was reduced in size and in several experiments appeared to have fused with the adjacent footprint 5. These observations confirm the supposition that footprints 4 and 6 are composed of multiple protein binding sites and suggest the possibility of interaction between protein molecules bound at adjacent sites. In addition, they support the argument that the GAGA factor is responsible for protection at most of the footprints in the *eve* promoter. Considering its size and sequence, it seems likely that site 9 also binds more than one protein.

TCCT-binding activity changes during development. It has previously been shown that GAGA factor is present in Kc tissue culture cells and also that the level of DNA-binding activity does not change appreciably during embryogenesis (4). If the TCCT factor is a transcription factor involved in the developmental regulation of *eve*, it would be important to know whether the level of the TCCT-binding activity varies during the course of embryogenesis. To address this issue, we compared the levels of TCCT-binding activity present in

FIG. 6. TCCT factor-binding activity is developmentally regulated. The DNA fragment in panel a was 5'-end labeled at a synthetic *Hind*III site at position -394 and extends to the *Xho*I site at $+7$. The fragment in panel b was labeled at a synthetic *Hind*III site at position -340 and also extends to $+7$. Arrows indicate incubation with 4, 8, 16, and 32 μ g of the H.4 fraction of 1- to 5-h, 5- to 10-h, 10- to 15-h, 15- to 20-h, and 1- to 3-h extracts. Footprints 7 and 8, marked by thick brackets, are TCCT factor-binding sites. 0, No protein; M, purine sequence ladder.

extracts prepared from embryos in different developmental stages. In initial DNase protection experiments, the H.4 fractions of 1- to 5-h, 5- to 10-h, 10- to 15-h, and 15- to 20-h extracts resulted in essentially identical patterns of protection within the region from -200 to -40 ; no additional footprints were detected, and the degree of protection at GAGA factor-binding sites 2 through 6 was nearly the same for all four extract fractions (data not shown). Figure 6a presents the results of an experiment in which the H.4 fraction of each of the four extracts was incubated with a DNA fragment containing adjacent high-affinity sites for the GAGA- and TCCT-binding activities. This fragment, which extends from -394 to $+7$ with respect to the transcription start site, was used in order to facilitate comparisons by allowing the use of the GAGA-binding site as a reference. Site 9 contains a 13-bp string of alternating CT dinucleotides which apparently represents overlapping GAGA sites. Immediately downstream is the high-affinity site 8, which contains two copies of the 8-bp TCCT consensus sequence. For each extract, the amount of protein included in the binding reactions was increased in twofold steps from 4 to 32 μ g. As expected, the GAGA-binding activities were similar in all four extract fractions. However, the ratio of TCCT-binding activity to GAGA-binding activity was lower in the 1- to 5-h and 15- to 20-h fractions than in the 5- to 10-h and 10- to 15-h extract fractions. This difference in activity was somewhat greater for the 1- to 5-h than for the 15- to 20-h

fraction. Figure 6b shows that the TCCT-binding activity was nearly undetectable in the H.4 fraction of a 1- to 3-h extract, even though the level of GAGA factor-binding activity remained high.

TCCT-binding activity is different from Elf-1/NTF-1 activity. Embryo extracts have been shown to contain a binding factor, Elf-1/NTF-1, that interacts with sequences present in the *Ddc* and *Ubx* promoters and within the fushi tarazu (*ftz*) neurogenic enhancer (4, 7, 8, 11). Since the 8-bp TCCT sequence is contained within the consensus recognition sequence that has been derived for Elf-1/NTF-1, it seemed possible that the TCCT factor corresponds to Elf-1/NTF-1. However, in initial experiments, we found that an oligonucleotide containing the sequence corresponding to *Ubx* FP4 failed to compete for binding to *eve* sites 7 and 8, even when present at a concentration 40 times greater than was required to compete for binding to FP4 (unpublished results). Furthermore, DNA-affinity-purified Elf-1/NTF-1 (provided by M. Biggin) was unable to protect *eve* sites 7 and 8 under conditions in which FP4 was fully protected (unpublished results). We then sought further, direct evidence that the H.4 fraction of the embryo extracts contains distinct factors that bind the *eve* and *Ubx* promoter sites. Figure 7 presents the results of DNase protection assays of fractions of the H.4 fraction of a 1- to 12-h embryo extract obtained by Superose 6 gel filtration chromatography in the presence of 0.2 M ammonium sulfate. Under these conditions, the factor that

FIG. 7. The *eve* TCCT sequence-binding activity is distinct from that of Elf-1/NTF-1. A 0.1 to 0.4 M KCl heparin-Sepharose fraction of a 1- to 12-h extract was applied to a Superose 6 gel filtration column, and fractions were assayed by DNase footprinting for binding to sites in the *eve* promoter (a) and in the *Ubx* promoter (b). The *eve* promoter fragment was 5' end labeled at a synthetic *Hind*III site at position -340 and extends to +7; the *Ubx* fragment was 5' end labeled at -201 and extends to +115. Numbers across the top of each autoradiogram designate adjacent fractions (15 ml), and 0 indicates no protein. Brackets mark the boundaries of footprints 7 and 8 in the *eve* promoter and FP4 in the *Ubx* promoter (39).

binds the sites in the *eve* promoter that contain the TCCT sequence eluted predominantly in fractions 5 and 6, while Elf-1/NTF-1 eluted in fractions 7 through 10.

DISCUSSION

As part of an effort to understand the mechanisms underlying the developmental regulation of *eve* expression, we have taken an in vitro approach to study factors involved in the initiation of *eve* transcription. In this report we have shown that *eve* is accurately and efficiently transcribed in nuclear extracts derived from *Drosophila* embryos and that the temporal pattern of *eve* transcription can be broadly reproduced with extracts prepared from embryos at different developmental stages. The transcription of *eve* in vitro is dependent on multiple sequence elements upstream of the mRNA start site. In addition to a TATA box consensus sequence, optimal transcription in early extracts requires at least three additional sequence elements located between positions -72 and -179. DNase protection experiments revealed multiple protein-binding sites in this region. Several lines of evidence indicate that a single protein is responsible for protection at these sites and that this protein corresponds to the previously described GAGA factor.

The GAGA factor was initially identified as a protein that binds to sequences in the *en* promoter that are required for optimal transcription both in vitro and in transfected Schneider 2 cells (39). Subsequently, the factor was affinity purified from embryo extracts and Kc cells, and the purified protein

was shown to stimulate *Ubx* transcription in vitro in a binding site-dependent manner (4). GAGA factor-binding sites have also been identified within sequences important for transcription of the ecdysone-inducible *E74* gene (40). The promoters of the *hsp26* and *hsp70* genes and the intergenic region between the histone H3 and H4 genes also contain binding sites for a protein that apparently corresponds to the GAGA factor (17). Eight of the footprints we have observed in the *eve* promoter appear to result from binding by the GAGA factor as well. While a consensus recognition sequence has been derived (4, 39), the GAGA factor seems to be promiscuous in its binding specificity; the *en*, *Ubx*, and *E74* promoters each contain GAGA-dependent binding sites which consist of imperfect copies of the pentanucleotide consensus sequence, 5'-GAGAG. Four of the *eve* promoter footprints contain perfect copies of the pentanucleotide sequence, and four others contain degenerate copies that are similar or identical to ones that occur in the *en*, *Ubx*, and *E74* promoters. We have shown that a synthetic oligonucleotide containing the same sequence, 5'-GAGAGAG, which was used to affinity purify the GAGA-factor (4), was able to compete for binding to the *eve* promoter, as was a fragment of the *en* promoter that contains seven GAGA factor-dependent footprints (39). Furthermore, an internal deletion that removes only 2 bp of one of the imperfect GAGA factor recognition sequences prevented protein binding at that site. These results strongly suggest that it is the GAGA factor that is responsible for most of the DNase-

protected sites in the *eve* promoter. We note however, that three of these apparent GAGA-dependent footprints contain potential zeste-binding sites (3), although the results of our competition experiments indicate that either zeste does not bind with high affinity to these sites or our extracts do not contain high levels of zeste-binding activity.

The arrangement of GAGA factor-binding sites in the *eve* promoter resembles their arrangement in the *en*, *Ubx*, and *E74* promoters. In all four promoters, the multiple GAGA-binding sites surround or overlap binding sites for other proteins, and they occur in both orientations with respect to the transcription start sites. In the *en* promoter, seven GAGA factor-dependent footprints occur between positions -35 and -350 (39). In the *Ubx* promoter, four footprints are found between -30 and -200 (4), and in *E74*, six sites occur between -40 and -220 (40). We have detected eight apparent GAGA factor-dependent footprints between -350 and +10 in the *eve* promoter. Two or more of these result from binding by more than one GAGA factor protein. The extent of several of the GAGA factor-dependent footprints in the *en*, *Ubx*, and *E74* promoters and the presence of multiple copies of the consensus sequence within them suggest the possibility that these footprints result from binding by more than one GAGA factor protein as well. It is notable that the GAGA factor functions as an activator of transcription both in *eve*, which contains a TATA box, and in *en*, which lacks a TATA box consensus sequence (39). It is possible that the GAGA factor is functionally equivalent to mammalian transcription factors such as Sp1, which can increase the activity of a promoter in a manner that is dependent on the number of binding sites but independent of their orientation within the promoter and which operate both within TATA box containing promoters and within promoters lacking a TATA box (see, for example, references 15 and 42). In agreement with the results of Biggin and Tjian (4), we found that the GAGA factor was present throughout embryogenesis. Since it is also present in Kc cells (4) and apparently in Schneider 2 cells (39), it seems unlikely that the GAGA factor is a highly specific transcription factor. However, it may be significant that GAGA factor-binding sites have to date been reported to occur only in the promoters of inducible or developmentally regulated genes (4, 17, 33, 39, 40).

Using extracts prepared from developmentally staged embryos, we have broadly reproduced the temporal pattern of *eve* expression observed in vivo. Recapitulations of the developmentally regulated transcription of the alcohol dehydrogenase (*Adh*) and *Ubx* genes in embryo nuclear extracts have been reported previously (4, 23). In the case of the *Adh* promoter, stage-specific transcription was correlated with binding by the transcription factor Adf1. The transcription of *Ubx*, however, was not clearly related to changes in the binding of any single factor within the region shown to be responsible for stage-specific transcription. Within the region we have shown to be required for the stage-specific transcription of *eve*, only GAGA factor-binding sites were detected by using the H.4 fraction of embryo nuclear extracts. Since the GAGA factor-binding activity does not change appreciably through embryogenesis (reference 4 and results reported herein), changes in the concentration or DNA-binding activity of this protein cannot explain the temporal pattern of transcription we have observed. It is possible that the activity of the GAGA factor as a transcriptional activator is regulated by posttranslational modification such that it is highest between 1 and 10 h postfertilization. Another possibility is that a factor which does not interact

directly with the *eve* promoter, or which interacts only weakly, varies in its activity.

We have also detected a protein-binding site in the *eve* untranslated leader. Functionally relevant protein-binding sites occur in the leader sequences of the *Ubx* (4) and *Antp* P2 (33) genes, and the importance of downstream sequences for transcription of *en* (39) and *E74* (40) has been reported. The *Ubx* leader contains an AT-rich sequence which constitutes a binding site for a number of homeobox proteins (1, 4, 5). However, the protected sequence in the *eve* leader bears no obvious similarity to this sequence or to sequences in the *Antp* P2 or *en* leaders. We have not investigated the possible functional significance of the *eve* leader site.

A factor present in late embryo extracts, which has been designated Elf-1 by Bray et al. (7) and NTF-1 by Dynlacht et al. (11), binds sites present in the *Ddc* and *Ubx* promoters and in the *ftz* neurogenic enhancer. While footprints 7 and 8 in the *eve* promoter contain a sequence which is part of the Elf-1/NTF-1 consensus recognition sequence, we have presented direct evidence that a factor other than Elf-1/NTF-1 binds sites 7 and 8. In other experiments, DNA-affinity-purified Elf-1/NTF-1 (provided by M. Biggin) failed to bind to *eve* sites 7 and 8 under conditions in which *Ubx* FP4 was fully protected, and an oligonucleotide which was able to compete for Elf-1/NTF-1 binding to FP4 failed to compete for binding at *eve* sites 7 and 8 (unpublished results).

Figure 8 compares the sequences protected by Elf-1/NTF-1 with sequences in the *eve* promoter that are protected by the TCCT factor. Also included in Fig. 8 are sequences protected by embryo extracts which are present in the *ftz* promoter (41) and within a region of *eve* upstream sequence which has been shown to have autoregulatory function (18, 20; see below). As can be seen in Fig. 8, while the 8-bp TCCT sequence is contained within the 16-bp Elf-1/NTF-1 consensus recognition sequence, among the binding sites in the *eve* promoter, there is little sequence similarity outside the TCCT sequence. In the case of *Ddc* element I, nucleotides just 5' to the TCCT sequence were shown to be required for Elf-1/NTF-1 binding in vitro and element I function in vivo (7, 8). In addition, while the footprint in element I is not centered over the TCCT sequence on both DNA strands (8), footprints 7 and 8 include the entire TCCT sequence on both strands. Altogether, these observations indicate that while the TCCT sequence is probably a significant determinant of binding specificity at the *eve* sites, Elf-1/NTF-1 seems to recognize sequences outside the TCCT sequence. It may be that the TCCT sequence represents a core recognition sequence for several proteins present in embryo extracts, each of which makes different contacts with nucleotide residues in the region that includes the TCCT sequence. An example of a similar situation is found in HeLa cells, where multiple proteins that bind sequences containing the CCAAT box have been identified (see, for example, references 9 and 10). One of these, CP1, makes contacts almost exclusively within the CCAAT sequence, whereas CP2 makes significant contacts outside the CCAAT box as well (9).

In addition to the three sites in the promoter region, the 8-bp TCCT consensus sequence also occurs within a 200-bp fragment of *eve* upstream sequence that exhibits an autoregulatory function in gastrulating embryos (18, 20). The TCCT sequence within this fragment is protected from DNase digestion by the embryo extract H.4 fraction and by affinity-purified TCCT factor, but not by purified Elf-1/NTF-1 (T. Hoey and D. Read, unpublished results). The presence of TCCT factor-binding sites in the autoregulatory and promot-

eve				
autoreg.	gaagaa ggtcctgc ga	-5.25kb	a	
site 11	ccgctc gctcctgc gg	-384bp	s	
site 8	tgagga ggtcctgc cc	-271	a	
	cctcat ggtcctgc cg	-256	s	
site 7	gctccg ggtcctgc tc	-214	s	
Consensus:	ggtcctgc			
Ddc element I	tgaacc ggtcctgc gg	-66	s	
Ubx FP4	caatct ggttttga gc	-112	s	
ftz neur. enh.	tgattc ggtatcga aa		s	
	gggagc tggttttt cc		s	
	tcggat ggttttat at	-1.65kb	s	
	aataag ggtgtaga gt		a	
	ctttct ggttttgc ag		a	
Consensus:	$\frac{1}{2}$ naac $\frac{1}{2}$ ggt $\frac{1}{2}$ c $\frac{1}{2}$ tgagg			
ftz promoter	tgagct ggtcctgc gg	-3	a	
	gcggca gacccctgc at	-15	a	

FIG. 8. Related sequences in the *eve*, *Ddc*, *Ubx*, and *ftz* promoters and in the *ftz* neurogenic enhancer (neur. enh.). The TCCT-related sequences are set apart by spaces from flanking nucleotides within the binding sites shown. s and a at the right of the figure indicate that the sequence shown occurs on the sense or antisense strand, respectively. Numbers indicate the distance in base pairs from the transcription start site, except for the *eve* autoregulatory region site and the *ftz* neurogenic enhancer sites, where this distance is given in kilobases. Data for the following sequences may be found in the references indicated in parentheses: *Ddc* element I (7, 8, 11), *Ubx* FP4 (4, 11), *ftz* neurogenic enhancer (11), consensus (11), and *ftz* promoter (41).

er-proximal regions suggests the possibility of a common function in these elements.

It has been proposed that *eve* transcriptional regulation involves separate mechanisms for the initial establishment of *eve* stripes and for the refinement and maintenance of striped expression (13, 18, 20). The TCCT-binding activity is first expressed at detectable levels between 3 and 5 h and appears to be highest between 5 and 15 h postfertilization. Since striped *eve* expression is first detectable between 2.5 and 3 h postfertilization, it appears unlikely that the TCCT factor is involved in the establishment of *eve* stripes. If the TCCT factor has a role in the striped pattern of *eve* expression, the temporal profile of DNA-binding activity of the TCCT factor is more consistent with a role in the maintenance or refinement of *eve* stripes. It is also possible that the TCCT factor participates in neural- or anal plate-specific *eve* expression, which continues through later stages of embryogenesis (12, 21, 26).

Elements responsible for correct spatial and temporal expression of *eve* in vivo occur upstream of -400 and appear

to be included in sequences more than 8 kb upstream of the transcription start site (18, 20). However, at least some of the gap and pair-rule gene response elements identified in the *eve* promoter through P-element transformation studies seem to require the presence of sequences between bp -42 and -400 (18, 20; R. Warrior and M. Levine, personal communication). No independent function for the region downstream of -400 has been detected in P-element transformation assays. This situation is similar to that of *Ddc* promoter element I, which has been shown to be necessary but not sufficient for proper central nervous system expression (2, 24, 38). It will be of interest to understand the mechanisms involved in this kind of promoter arrangement and to know whether the promoter-proximal elements function by receiving or amplifying a state of promoter activity independently determined by factors binding at distal sites or whether they participate in a more direct way in programming promoter activity. The combination of in vitro studies and the powerful genetic techniques that can be applied in *Drosophila melanogaster* should eventually provide answers to such questions.

ACKNOWLEDGMENTS

We thank W. Soeller and M. Biggin for helpful advice in the early stages of this work and Y. Takagaki, P. Beachy, K. Han, T. Hoey, H. Doyle, K. Harding, and M. Levine for continuing advice and discussions. We also thank R. Warrior and M. Levine for sharing unpublished results. The authors are grateful to K. Han for the actin 5C/*eve* plasmid, T. Hoey for the *eve* genomic clone and sequence, T. Hoey and K. Harding for the *eve* d11/d13 mutant plasmid, and M. Biggin for providing the pUbx HIII Δ5'-201 plasmid and the affinity-purified NTF-1.

T. Nishigaki was supported by Sankyo Co. Ltd. This work was funded by Public Health Service grant GM 37971 from the National Institutes of Health.

LITERATURE CITED

1. Beachy, P. A., M. A. Krasnow, E. R. Gravis, and D. S. Hogness. 1988. An *Ultrabithorax* protein binds sequences near its own and the antennapedia P1 promoters. *Cell* 55:1069-1081.
2. Beall, C. J., and J. Hirsh. 1987. Regulation of the *Drosophila* dopa decarboxylase gene in neuronal and glial cells. *Genes Dev.* 1:510-520.
3. Biggin, M. D., S. Bickel, M. Benson, V. Pirrota, and R. Tjian. 1988. Zeste encodes a sequence-specific transcription factor that activates the *Ultrabithorax* promoter in vitro. *Cell* 53:713-722.
4. Biggin, M. D., and R. Tjian. 1988. Transcription factors that activate the *Ultrabithorax* promoter in developmentally staged extracts. *Cell* 53:699-711.
5. Biggin, M. D., and R. Tjian. 1989. A purified *Drosophila* homeodomain protein represses transcription in vitro. *Cell* 58:433-440.
6. Biggin, M. D., and R. Tjian. 1989. Transcription factors and the control of *Drosophila* development. *Trends Genet.* 5:377-383.
7. 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.
8. Bray, S. J., W. A. Johnson, J. Hirsh, U. Heberlein, and R. Tjian. 1988. A cis-acting element and associated binding factor required for CNS expression of the *Drosophila melanogaster* dopa decarboxylase gene. *EMBO J.* 7:177-188.
9. Chodosh, L. A., A. S. Baldwin, R. W. Carthew, and P. A. Sharp. 1988. Human CCAAT-binding proteins have heterologous subunits. *Cell* 53:11-24.
10. Dorn, A., J. Bollekens, A. Staub, C. Benoist, and D. Mathis. 1987. A multiplicity of CCAAT box-binding proteins. *Cell* 50:863-872.
11. 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.
12. Frasch, M., T. Hoey, C. Rushlow, H. Doyle, and M. Levine. 1987. Characterization and localization of the *even-skipped* protein of *Drosophila*. *EMBO J.* 6:749-759.
13. Frasch, M., and M. Levine. 1987. Complementary patterns of *even-skipped* and *fushi tarazu* expression involve their differential regulation by a common set of segmentation genes in *Drosophila*. *Genes Dev.* 1:981-995.
14. Frasch, M., R. Warrior, J. Tugwood, and M. Levine. 1988. Molecular analysis of *even-skipped* mutants in *Drosophila* development. *Genes Dev.* 2:1824-1838.
15. Gidoni, D., J. T. Kadonaga, H. Barrera-Saldaña, K. Takahashi, P. Chambon, and R. Tjian. 1985. Bidirectional SV40 transcription mediated by tandem Sp1 binding interactions. *Science* 230:511-517.
16. Gilmour, D. S., T. J. Dietz, and S. C. R. Elgin. 1988. TATA box-dependent protein-DNA interactions are detected on heat shock and histone gene promoters in nuclear extracts derived from *Drosophila melanogaster* embryos. *Mol. Cell. Biol.* 8:3204-3214.
17. Gilmour, D. S., G. H. Thomas, and S. C. R. Elgin. 1989. *Drosophila* nuclear proteins bind to regions of alternating C and T residues in gene promoters. *Science* 245:1487-1490.
18. Goto, T., P. Macdonald, 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.
19. Han, K., M. S. Levine, and J. L. Manley. 1989. Synergistic activation and repression of transcription by *Drosophila* homeobox proteins. *Cell* 56:573-583.
20. 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.
21. Harding, K., C. Rushlow, H. Doyle, T. Hoey, and M. Levine. 1986. Cross-regulatory interactions among pair-rule genes in *Drosophila*. *Science* 233:953-959.
22. 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.
23. Heberlein, U., and R. Tjian. 1988. Temporal pattern of alcohol dehydrogenase gene transcription reproduced by *Drosophila* stage-specific embryonic extracts. *Nature (London)* 331:410-415.
24. Hirsh, J. 1989. Molecular genetics of dopa decarboxylase and biogenic amines in *Drosophila*. *Dev. Genet.* 10:232-238.
25. Jaynes, J. B., and P. H. O'Farrell. 1988. Activation and repression of transcription by homeodomain-containing proteins that bind a common site. *Nature (London)* 336:744-749.
26. Macdonald, P. M., P. Ingham, and G. Struhl. 1986. Isolation, structure, and expression of *even-skipped*: a second pair-rule gene of *Drosophila* containing a homeo box. *Cell* 47:721-734.
27. Maxam, A. M., and W. Gilbert. 1980. Sequencing end-labeled DNA with base-specific chemical cleavages. *Methods Enzymol.* 65:499-560.
28. McKnight, S. L., and R. Kingsbury. 1982. Transcriptional control signals of a eukaryotic protein-coding gene. *Science* 217:316-324.
29. Nüsslein-Volhard, C., H. Kluding, and G. Jurgens. 1985. Genes affecting the segmental subdivision of the *Drosophila* embryo. *Cold Spring Harbor Symp. Quant. Biol.* 50:145-154.
30. Nüsslein-Volhard, C., and E. Weischaus. 1980. Mutations affecting segment number and polarity in *Drosophila*. *Nature (London)* 287:795-801.
31. Parker, C., and J. Topol. 1984. A *Drosophila* RNA polymerase II transcription factor contains a promoter-region-specific DNA-binding activity. *Cell* 36:357-369.
32. Parker, C. S., and J. Topol. 1984. A *Drosophila* RNA polymerase II transcription factor binds to the regulatory site of an hsp70 gene. *Cell* 37:273-283.
33. 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.
34. Perkins, K. K., G. M. Dailey, and R. Tjian. 1988. Novel Jun- and Fos-related proteins in *Drosophila* are functionally homologous to enhancer factor AP-1. *EMBO J.* 7:4265-4273.
35. Price, D. H., A. E. Sluder, and A. L. Greenleaf. 1989. Dynamic interaction between a *Drosophila* transcription factor and RNA polymerase II. *Mol. Cell. Biol.* 9:1465-1475.
36. Sanger, F., S. Nicklen, and A. R. Coulson. 1977. DNA sequencing with chain-terminating inhibitors. *Proc. Natl. Acad. Sci. USA* 74:5463-5467.
37. Sassone-Corsi, P., J. P. Dougherty, B. Wasyluk, and P. Chambon. 1984. Stimulation of *in vitro* transcription from heterologous promoters by the simian virus 40 enhancer. *Proc. Natl. Acad. Sci. USA* 81:308-312.
38. Scholnick, S. B., S. J. Bray, B. A. Morgan, C. A. McCormick, and J. Hirsh. 1986. CNS and hypoderm regulatory elements of the *Drosophila melanogaster* dopa decarboxylase gene. *Science* 234:998-1002.
39. Soeller, W. C., S. J. Poole, and T. Kornberg. 1988. In vitro transcription of the *Drosophila* engrailed gene. *Genes Dev.* 2:68-81.
40. Thummel, C. 1989. The *Drosophila* *E74* promoter contains essential sequences downstream from the start site of transcription. *Genes Dev.* 3:782-792.
41. Topol, J., G. Wiederrecht, and C. S. Parker. 1987. Biochemical analysis of the *fushi tarazu* and *Ultrabithorax* promoter regions, p. 3-12. In W. F. Loomis (ed.), *Genetic regulation of development*. Alan R. Liss, Inc., New York.
42. Vigneron, M., H. A. Barrera-Saldaña, D. Baty, R. E. Everett, and P. Chambon. 1984. Effect of the 21-bp repeat upstream element on *in vitro* transcription from the early and late SV40 promoters. *EMBO J.* 3:2373-2382.
43. Warburg, O., and W. Christian. 1941. Isolierung und Kristallisation des Gärungsferments Enolase. *Biochem. Z.* 310:384-421.
44. Wu, C. 1984. Activating protein factor binds *in vitro* to upstream control sequences in heat shock gene chromatin. *Nature (London)* 311:81-84.
45. Wu, C., S. Wilson, B. Walker, I. Dawid, T. Paisley, V. Zimarino, and H. Ueda. 1987. Purification and properties of *Drosophila* heat shock activator protein. *Science* 238:1247-1253.