

## In Vivo Transcriptional Analysis of the TATA-Less Promoter of the *Drosophila melanogaster* vermilion Gene

YIH-WOEI C. FRIDELL† AND LILLIE L. SEARLES\*

Department of Biology, The University of North Carolina at Chapel Hill,  
Chapel Hill, North Carolina 27599-3280

Received 5 June 1992/Returned for modification 15 July 1992/Accepted 22 July 1992

**Transcriptional regulation of the TATA-less promoter of the *Drosophila melanogaster* vermilion (*v*) gene was investigated. Developmental Northern (RNA) blot analysis showed that *v* transcripts accumulate during late embryo, larval, and adult stages. Sequences that control expression in adults were delineated by analyzing a series of 5' and 3' deletion constructions after germ line transformation. These studies defined two regions, -300 to -600 and -60 to -160, relative to the major transcription start site, as important for maximal levels of expression. Analysis of transformants bearing *v-lacZ* promoter fusions showed that larval expression is fat body specific and that expression depends on sequences located between +19 and +36 downstream of transcription start site. This downstream element can be functionally replaced by a TATA box in vivo. Furthermore, when added to the wild-type *v* promoter, a TATA element augments the level of *v* transcription by three- to fivefold.**

A typical eukaryotic gene has a TATA element located about 30 bp upstream of the transcription start site. This sequence element plays a pivotal role in eukaryotic transcription initiation. The recognition of the TATA box by the TATA-binding protein (formerly TFIID) marks an early step in assembly of the transcription initiation complex. With the subsequent binding of other general transcription factors and RNA polymerase II, transcription initiates at the downstream start site (4, 6, 11, 19, 37).

Although much progress has been made toward understanding how transcription of TATA-containing genes is controlled, less is known about how initiation of transcription is directed on promoters lacking TATA elements. Studies have shown that the same factors necessary for initiation on TATA-containing promoters participate in transcription of TATA-less promoters (5, 21, 31). However, it is not clear how these factors assemble in the absence of the TATA box. Recent evidence suggests that TATA-binding protein-associated factors are involved in recognizing TATA-less promoter sequences (22). In *Drosophila melanogaster*, certain developmentally regulated, TATA-lacking genes, such as engrailed, Antennapedia, Ultrabithorax, *E74*, and retrotransposons gypsy and *mdg 1*, contain sequences within the transcription unit that are essential for transcription (1, 2, 12, 18, 32, 35). For these genes, sequences extending to +60 relative to the transcription start site have been shown to be important for proper transcription in vitro. For the Antennapedia P2 and the *Ubx* promoters, potential factor-binding sites have been identified downstream of the transcription start sites, indicating the involvement of specific protein-DNA interactions.

Since most of the information about TATA-less promoters has been obtained from in vitro studies, we have chosen the *D. melanogaster* TATA-less vermilion (*v*) gene as a model to examine how transcription in vivo is controlled in the absence of a TATA box. The *v* gene product, tryptophan

oxygenase, catalyzes the first step in brown eye pigment synthesis, converting tryptophan to *N*-formylkynurenine (20 and references therein). The *v* gene has been independently cloned by P-element tagging and by screening a *D. melanogaster* genomic library with a rat tryptophan oxygenase cDNA clone (28, 38). The intron-exon structure of the gene, which encodes a 1.4-kb transcript, has been determined. The major transcription initiation site has been mapped by both primer extension and S1 nuclease protection experiments (27).

In this report, we examine the temporal and spatial expression of the *v* gene and delineate sequences that control transcription initiation. Not only do we find that two regions of DNA upstream of the start site are required, we also show that sequences in the transcribed leader region are essential for *v* transcription initiation in vivo. Furthermore, we demonstrate that the element downstream of the transcription start site can be functionally replaced by a TATA box upstream of that site.

### MATERIALS AND METHODS

***Drosophila* strains and germ line transformation procedure.** Fly stocks were reared at 25°C on standard food. Oregon R flies were used as a wild-type strain. Homozygous *v*<sup>36f</sup> *ry*<sup>506</sup> flies were constructed and used as a parental line for injecting the series of *v* promoter deletion constructions. The *v*<sup>36f</sup> allele has a roo/B104 insertion in intron 4 of the *v* gene; as a result, a very low level of nonfunctional, truncated *v* transcript accumulates in this mutant (27). A homozygous *ry*<sup>506</sup> strain was used for injection of *v-lacZ* fusion constructions. Germ line transformants were generated as described by Karess (13). G1 transformants were selected on the basis of *ry* eye color rescue. To make homozygous transformant lines, we used standard second and third chromosome balancer chromosomes.

**DNA manipulation.** A series of 5' and 3' *v* deletion constructions was made by using either convenient restriction sites (14) or the polymerase chain reaction (PCR) technique (26). DNA was amplified in a DNA Thermal Cycler (Perkin-Elmer Cetus Corp.) according to the proce-

\* Corresponding author.

† Present address: 237 Lineberger Comprehensive Cancer Center, University of North Carolina at Chapel Hill, Chapel Hill, NC 27599-7295.

dures provided by the manufacturer. Desired  $\nu$  fragments were subcloned into the polylinker of the pBluescript ks(-) vector (Stratagene) to obtain *NotI*- and *SalI*-flanking restriction sites. The *NotI*-*SalI* fragments were then transferred into the *D. melanogaster* transformation vector pDM30 (16).

To make the pTUF1.1  $\nu$ -*lacZ* fusion construction, a 1.1-kb *EcoRI*-*XhoI* promoter fragment was subcloned into the pBluescript ks(-) vector to obtain flanking *KpnI* and *BamHI* sites. This  $\nu$  *KpnI*-*BamHI* fragment was subcloned into the pDM66A vector, which contains a *lacZ* cassette (36). The resulting  $\nu$ -*lacZ* fusion gene was then subcloned into the *NotI* sites of the *D. melanogaster* transformation vector pDM30 (16). Construction pTUF1.1L was made as follows. A 1.1-kb *EcoRI*-*ApaI*  $\nu$  promoter fragment, containing all the  $\nu$  leader sequences and upstream sequences to -1114, was amplified by the PCR technique. An *ApaI* restriction site was introduced into this fragment by attaching an *ApaI* site to the 5' end of the downstream primer. The PCR-amplified fragment was subcloned into the plasmid pBluescript ks(-) to obtain *KpnI*- and *BamHI*-flanking restriction sites. The same procedure described above was used to subclone this fragment into pDM30. This cloning procedure introduced an extra G · C base pair at the 3' end of the  $\nu$  leader. Since the leader linker mutations were derived by the same procedure, this extra G · C base pair is present in these constructions as well. For the TATA box-containing constructions, the template-strand oligonucleotide 5'-GGAGATGAGATTTATACCCAGCG-3' and the nontemplate-strand oligonucleotide 5'-CGCTGGGGTATAAATCTCATCTCC-3' were used in an overlapping PCR (10) to introduce a 5'-TATAAAT-3' sequence into the  $\nu$  promoter region with the 3' T occupying position -25.

Linker substitution oligonucleotides were designed to replace wild-type  $\nu$  leader sequences with new sequences, each nucleotide being changed to a randomly selected, different nucleotide. The substitutions were introduced by the overlapping PCR technique (10). For pTUF1.1L1, the template-strand oligonucleotide 5'-ATGGGCCCAACTACTGAAGCAGATCAGATC-3' was used to replace the wild-type sequence between +47 and +57 immediately upstream of the translation start codon. The template-strand oligonucleotide 5'-TTCCTGAAAGACTTACGGGATGCTGGC-3' and the nontemplate-strand oligonucleotide 5'-CGTAAGTCCTCAGTGAAATCTGATCTGAGCTG-3' were used in an overlapping PCR to introduce a mutation from +19 to +36 in construction pTUF1.1L2. To ensure that the desired changes were introduced, all the PCR-amplified products were sequenced with Sequenase (US Biochemicals). The subcloning strategy was the same as that used in constructing the wild-type leader fusions (pTUF1.1L).

**mRNA isolation and Northern (RNA) analysis.** For the developmental Northern analysis, RNA was prepared from fresh embryos, larvae, pupae, and 2- to 3-day-old adults. The RNA isolation procedure and Northern analysis were performed as described by Searles et al. (27). For transcriptional analysis of transformants, total RNA was extracted from flies following a procedure described by Puissant and Houdebine (23) with the following modifications. Two- to 3-day-old flies were collected on dry ice and stored at -80°C. Frozen flies (~0.5 g) were homogenized in 5 ml of ice-cold grinding buffer. Subsequently, 100 to 120  $\mu$ g of total RNA was dissolved in 3 ml of binding buffer (0.5 M NaCl, 10 mM Tris [pH 7.5], 1 mM EDTA) and passed through a small column containing 20 mg of oligo(dT)-cellulose. The eluate was passed twice again through the column. After the column was washed with 4 ml of binding buffer, poly(A)<sup>+</sup>

RNA was eluted with 450  $\mu$ l of elution buffer (10 mM Tris [pH 7.5], 1 mM EDTA). Ethanol-precipitated samples were loaded on 1.5% formaldehyde-agarose gels (14). The fractionated RNA was blotted onto a nylon membrane (34). Prehybridization, hybridization, and posthybridization washes were performed as described by Zinn et al. (39). Riboprobes of  $\nu$  and *rp49* were synthesized as described by Melton et al. (15) except that  $\nu$  riboprobe synthesis reactions contained 200  $\mu$ Ci of [<sup>32</sup>P]CTP (800 Ci/mmol) to increase the specific activity. Northern filters were exposed to X-ray film at room temperature without intensifying screens for 2 to 3 days. Densitometry was performed with a densitometer (model 300; Molecular Dynamics).

**Genomic DNA isolation and Southern analysis.** Genomic DNA from each transformant line was prepared as described by Bingham et al. (3). Approximately 0.5 to 1 g of frozen flies was used in each preparation. Blotting, hybridization, and subsequent washes were performed as described by Southern (33). Linear plasmid DNA was radiolabeled with <sup>32</sup>P by the random priming procedure (7).

**$\beta$ -Galactosidase assays.**  $\beta$ -Galactosidase staining was performed by a procedure described by Glaser et al. (9) with the following modifications. After dissection, tissues were incubated in fixing solution for 3 to 5 min at room temperature. Treated tissues were then stained and incubated in the dark at room temperature for 3 h to overnight. For photography, tissues were immersed in glycerol, mounted on slides, and photographed with a Zeiss microscope.

The level of  $\beta$ -galactosidase activity produced from each  $\nu$ -*lacZ* fusion construction was quantitated by a spectrophotometric assay, according to the procedure of Simon and Lis (29). Ten third-instar crawling larvae from each transformant line were homogenized in 500  $\mu$ l of ice-cold assay buffer for 30 s at 4°C. To each sample, an additional 500  $\mu$ l of ice-cold assay buffer was added, and the mixture was briefly and gently vortexed. Samples were centrifuged at 4°C for 5 min. For each assay, 100  $\mu$ l of the supernatant was added to 1 ml of a solution of 0.06% chlorophenol red- $\beta$ -D-galactopyranoside substrate (CPRG; Boehringer Mannheim) in assay buffer. Samples were incubated at 37°C in the dark. The substrate conversion was measured at 574 nm at 0.5, 1, 2, and 3 h after extract addition. In this period, the rate of color development was linear. The  $\beta$ -galactosidase activity was defined as absorbance units per hour per microgram of protein. In larvae, endogenous  $\beta$ -galactosidase activity is found in the gut. To correct for this endogenous activity, extracts from *ry*<sup>506</sup> homozygous larvae (injection stock) were included in each experiment, and this background reading was subtracted from readings obtained for each transformant line. The protein concentration of the extracts was determined by the Bio-Rad protein assay.

## RESULTS

**Upstream DNA sequences modulate the level of  $\nu$  transcription.** Northern blot analysis was used to examine accumulation of the 1.4-kb  $\nu$  transcript during development. As shown in Fig. 1, the  $\nu$  gene is first turned on in 12- to 24-h-old embryos. The level of  $\nu$  mRNA remains relatively constant through larval development, although a higher level of  $\nu$  expression is observed in early third-instar larvae than in late third-instar larvae (data not shown). A dramatic decline in  $\nu$  mRNA accumulation is seen in the pupal stage, in which a very low level of  $\nu$  transcript is detected. However, in 2- to 3-day-old adults, the level of accumulation is relatively high. Since eye pigment deposition occurs before eclosion (20),  $\nu$

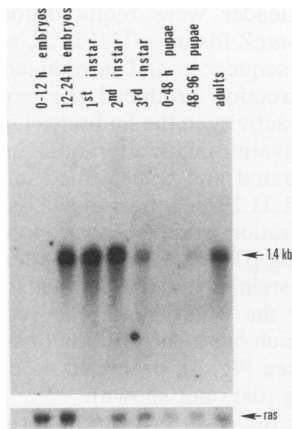


FIG. 1. Developmental profile of  $\nu$  RNA accumulation. RNA isolated from *D. melanogaster* at various developmental stages was analyzed on a Northern blot. Oregon R flies were allowed to lay eggs for various intervals at 25°C. After aging for appropriate times, RNA was prepared from samples that included the following ranges: early embryos, 0 to 12 h; late embryos, 12 to 24 h; first-instar larvae, 24 to 48 h; second-instar larvae, 48 to 72 h; third-instar larvae, 72 to 120 h; 0 to 48-h pupae, 5 to 7 days; 48 to 96-h pupae, 7 to 9 days; adults, 2 to 3 days after eclosion. Each lane contains 20  $\mu$ g of poly(A)<sup>+</sup> RNA. The blot was probed with a clone of the 1.9-kb *Xho*I-*Hind*III  $\nu$  fragment. As an internal control for loading, the filter was re-probed with a clone of the *D. melanogaster ras* gene, which is expressed at approximately the same level throughout development.

expression during the larval stage is responsible for controlling the eye color phenotype of adults.

To examine the transcriptional regulation of the  $\nu$  TATA-less promoter, we generated a series of wild-type  $\nu$  gene

derivatives and introduced them into the *D. melanogaster* genome by P-element-mediated germ line transformation (25). Each construction was tested for its ability to rescue the  $\nu$  mutant eye color phenotype and to produce RNA in adults (Fig. 2). A 5.4-kb *Hind*III fragment, including 3.2 kb upstream of the major transcription start site and 300 bp downstream of the polyadenylation signal, expressed wild-type levels of  $\nu$  transcript. Constructions with either 1.1 kb or 600 bp of upstream sequences expressed slightly lower levels of  $\nu$  transcript. When 300 or 160 bp of upstream sequences were included, the level of expression is sufficient to rescue the  $\nu$  mutant phenotype; however, adult transcript levels were reduced two- to threefold. On the other hand, constructions with 100 or 60 nucleotides upstream of the transcription start site are not capable of rescuing the mutant phenotype, and they produce low, but detectable, amounts of RNA. Thus, whereas sequences upstream of -1114 have at most a slight effect on  $\nu$  expression, sequences in the region between -600 and -300 and between -160 and -90 are largely responsible for regulating the level of transcription. Similar expression levels were observed with constructions including either 300 or 90 bp of downstream of the polyadenylation site (Fig. 2); this indicates that sequences in the 3' region of the gene do not appear to be involved in  $\nu$  regulation.

#### Internal regulatory element resides in the $\nu$ leader region.

Previous studies suggested that the  $\nu$  gene is expressed in larval fat body (24). To examine tissue-specific expression of the  $\nu$  gene, we used the bacterial *lacZ* gene as a reporter to monitor the  $\nu$  promoter activity. A  $\nu$ -*lacZ* chimeric construction, pTUF1.1, was made by fusing a  $\nu$  fragment containing 1.1 kb upstream of the major  $\nu$  transcription start site and 28 bp of  $\nu$  leader to a *lacZ* reporter cassette. This cassette

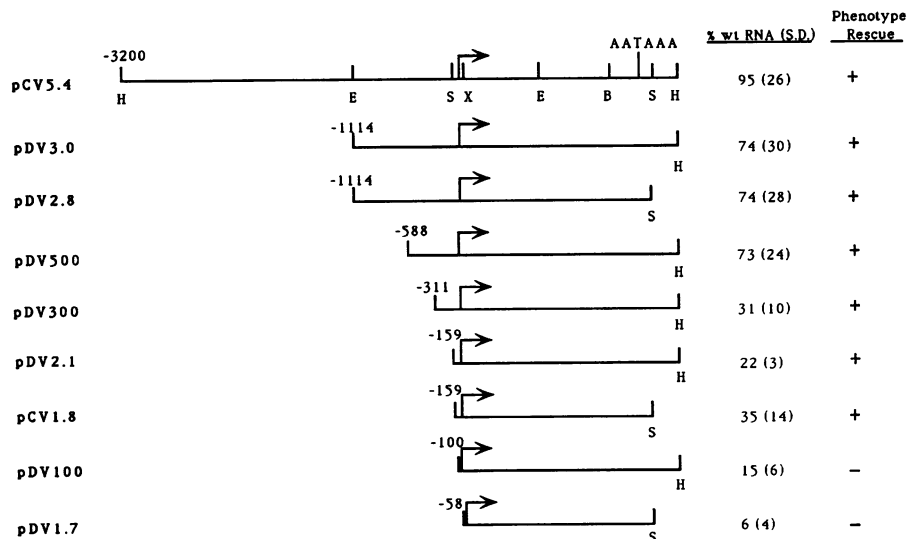


FIG. 2. Deletion constructions of the  $\nu$  promoter. The 5' and 3' deletion constructions were transformed into the *D. melanogaster* germ line. RNA levels were measured by Northern blot analysis of independently isolated transformants. For each deletion construction, with the exception of pDV500 and pDV300, for which two lines of each construction were examined, at least three independently isolated transformant lines were analyzed. Three or more independent RNA preparations were used to determine the average amount of  $\nu$  mRNA produced by each transformant line. To control for differences in loading, each Northern blot was also probed with the *D. melanogaster rp49* sequences (17). RNA levels are expressed as percentage of wild type. The eye color of newly eclosed adult flies was examined to determine phenotypic rescue. The numbers above the line indicate the 5' boundary (in base pairs) of each construction relative to the transcription start site, indicated by an arrow. The 3' *Hind*III and *Ssp*I sites are located approximately 300 and 90 bp, respectively, downstream of the polyadenylation signal. The position of the polyadenylation signal, AATAAAA, is shown. Abbreviations of restriction enzymes are as follows: B, *Bam*HI; E, *Eco*RI; H, *Hind*III; S, *Ssp*I; X, *Xho*I. S.D., standard deviation.

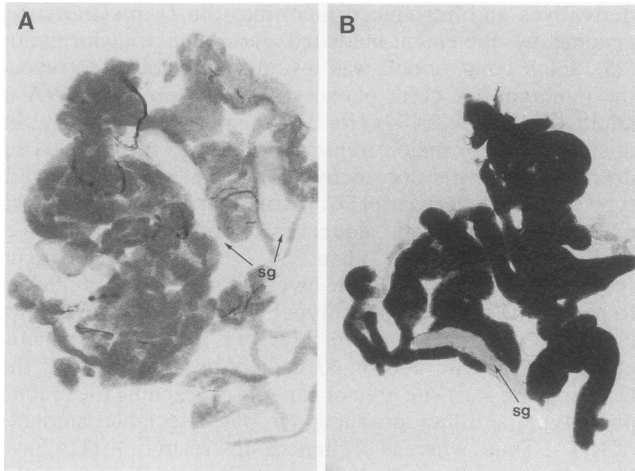


FIG. 3. Fat body-staining patterns generated by *v-lacZ* fusions in third-instar larvae. (A) Fat body tissue from larvae bearing the *v-lacZ* fusion construction containing 1.1 kb upstream of the transcription start site and 28 bp of the leader sequence. (B) Fat body tissue from larvae carrying a fusion construction identical to that described in panel A except that the entire leader region (extending to +57) was included. Sg, salivary gland.

contains the *D. melanogaster* alcohol dehydrogenase (*Adh*) leader and translational start codon, the bacterial *lacZ* structural gene, and the simian virus 40 polyadenylation signal (36). After stable germ line transformant lines bearing this construction were established, third-instar larvae were dissected and stained for  $\beta$ -galactosidase activity. Surprisingly, when several homozygous lines containing this construction were examined, no fat body expression was detected (Fig. 3A), and the only staining observed was endogenous gut staining (data not shown; see reference 9). Furthermore, no *lacZ* fusion transcripts were detected in adult transformant lines bearing the pTUF1.1 construction (data not shown). To test whether sequences further down-

stream in the leader were required for expression, we constructed a *v-lacZ* fusion, pTUF1.1L, which includes the entire *v* leader sequences and introduced it into the germ line; this construction produced the expected pattern of  $\beta$ -galactosidase activity in the fat body (Fig. 3B). Two linker substitution derivatives with alterations in the leader region were also generated and transformed into flies. The construction pTUF1.1L1 is identical to pTUF1.1L except that it contains an alteration from +47 to +57 (see Fig. 4). Transformants carrying pTUF1.1L1 exhibited a fat body-specific  $\beta$ -galactosidase staining pattern identical to that produced by pTUF1.1L. On the other hand, the second construction pTUF1.1L2, which has a substitution of sequences between +19 and +36 (see Fig. 4), produced no detectable  $\beta$ -galactosidase staining (data not shown).

To quantitate the level of *v* promoter-driven expression from these four *v-lacZ* fusions, we assayed crude extracts prepared from early third-instar larvae carrying these constructions for  $\beta$ -galactosidase activity (Fig. 4). The level of expression from pTUF1.1L, which contains the entire leader region, was arbitrarily defined as 100%. With the partial leader construction, pTUF1.1, which includes sequences extending to +28 (Fig. 4), no  $\beta$ -galactosidase activity above background was detected. The linker substitution derivative pTUF1.1L1 (+47 to +57) produced nearly wild-type levels of  $\beta$ -galactosidase activity (Fig. 4). However, the other derivative, pTUF1.1L2 (+19 to +36), expressed only about 1% of the wild-type level of  $\beta$ -galactosidase activity (Fig. 4). Therefore, the analysis of transformants carrying either of two constructions, one that has a deletion of leader sequences downstream of +28 and a second which has a replacement of leader sequences between +19 and +36, defines a regulatory element downstream of the transcription initiation site that is required for *v* expression in vivo.

**Effects of a TATAAAT sequence on transcription from the *v* promoter.** To test whether a TATA box can functionally compensate for the absence of the downstream element, we used the PCR technique to introduce the sequence 5'-TATAAAT-3' at -25 in construction pTUF1.1, which ter-

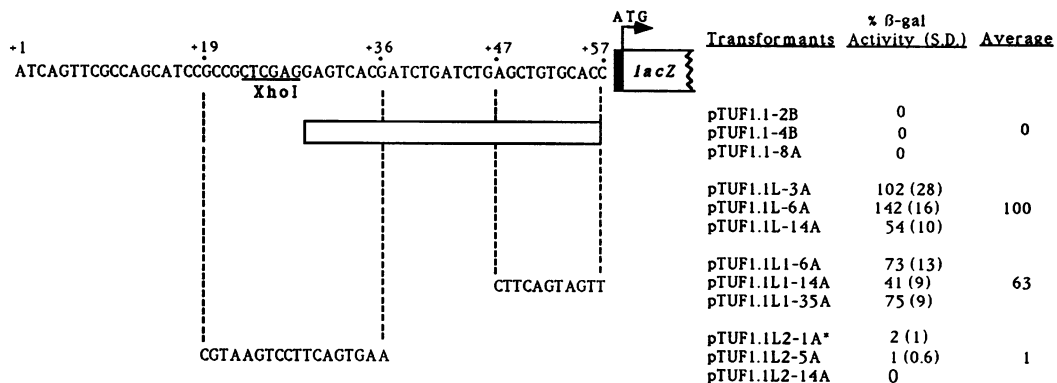


FIG. 4. Promoter activity of *v-lacZ* fusions carrying lesions in the leader region. The leader sequence of the nontemplate DNA strand is shown. The solid box indicates the *Adh* leader sequence. The open box beneath the sequence indicates the region deleted in pTUF1.1. The arrow indicates the translational start codon. Oligomer sequences used to substitute for the wild-type leader segments in constructions pTUF1.1L1 and pTUF1.1L2 are indicated. The broken lines indicate boundaries of the substitutions. For the construction pTUF1.1L, the leader sequence is wild type. To quantitate promoter activity, we performed  $\beta$ -galactosidase assays on crude extracts of third-instar larvae from each transformant line. The  $\beta$ -galactosidase activity from each construction is indicated relative to the wild-type leader construction (pTUF1.1L). For each construction, three independent lines were examined. The  $\beta$ -galactosidase activity of each line was assayed twice, each time in duplicate. The asterisk indicates a transformant which carried a male lethal mutation; thus, for this line,  $\beta$ -galactosidase activity was measured in heterozygous females, and the activity was doubled for comparison with that of the other homozygous lines. S.D., standard deviation.

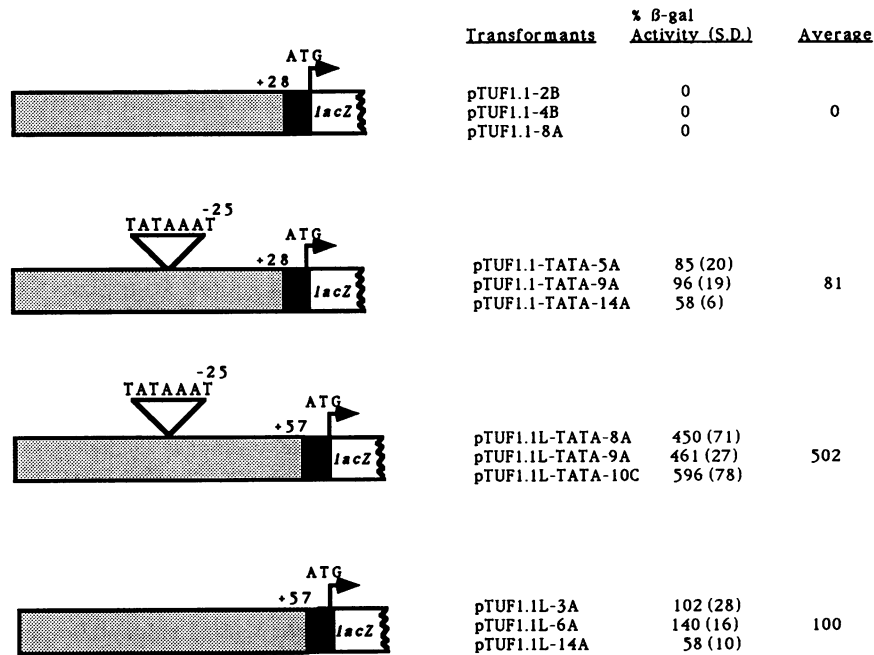


FIG. 5. Effects of a TATA box on the expression of *v-lacZ* fusions. The level of  $\beta$ -galactosidase ( $\beta$ -gal) activity in transformants bearing each construction is indicated relative to that of the wild-type leader construction (pTUF1.1L). For each construction, three independent lines were examined. The  $\beta$ -galactosidase assay was performed twice for each line, in duplicate. The arrow indicates the translation start codon. The inverted triangle indicates the insertion of a TATA box. The stippled box indicates the *v* sequences, which drive the same *Adh-lacZ* reporter cassette as described in the legend to Fig. 4. S.D., standard deviation.

minates at +28. The resulting construction, pTUF1.1-TATA, was used to generate transformants, and various tissues from third-instar larvae bearing this construction were stained for  $\beta$ -galactosidase activity. Fat body-specific staining was detected in all three independent lines tested (data not shown). To determine the level of  $\beta$ -galactosidase activity produced from the pTUF1.1-TATA construction, we assayed crude extracts of third-instar larvae carrying this construction for  $\beta$ -galactosidase activity. As shown in Fig. 5, 81% of the wild-type (pTUF1.1L) level of enzyme activity was generated by the pTUF1.1-TATA construction. This finding strongly argues that the TATA box can functionally replace the downstream sequences and that the downstream sequences are probably important for transcription initiation. Furthermore, these observations indicate that the upstream regulatory regions control the fat body specificity of expression because, as noted above, the wild-type pattern of expression was observed with the TATA-containing *lacZ* fusion construction that lacked the downstream element.

As a parallel approach to examine the effect of a TATA box on the *v* promoter, a TATA box was added to a *v-lacZ* leader fusion that has the entire leader segment. It was observed that the  $\beta$ -galactosidase staining of fat bodies dissected from transformants bearing this construction, pTUF1.1L-TATA, was stronger and detected earlier (~30 min compared to ~3 h) than that observed from construction pTUF1.1L (data not shown). This suggested that the addition of a TATA box to the pTUF1.1L construction resulted in a higher level of *v-lacZ* expression, and measurements of  $\beta$ -galactosidase activity verified this observation. As depicted in Fig. 5, a fivefold increase in  $\beta$ -galactosidase activity was seen in this TATA-containing construction compared with the pTUF1.1L construction.

A similar result was obtained when a TATA box was

incorporated into the authentic *v* gene. This construction, designed pDVTATA, contains the *v* transcription unit, 1.1 kb of sequences upstream of the major transcription start site, and 300 bp downstream of the polyadenylation signal. The identical construction without the TATA sequence (pDV3.0, Fig. 2) produces nearly the wild-type levels of *v* transcript. As shown in Fig. 6, levels of *v* expression from three independent transformant lines carrying this pDV TATA construction were, on average, threefold greater than *v* transcript accumulation observed in the wild type. Thus, the inclusion of a TATA box in the *v* promoter that normally

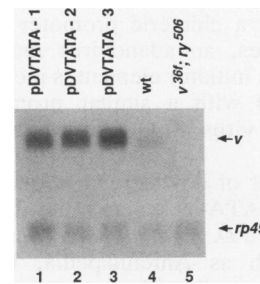


FIG. 6. Effects of adding a TATA box to the authentic *v* gene. Northern analysis was performed on RNA isolated from transformants carrying the pDVTATA construction, which has a TATA box added at -25 of the *v* gene. Each lane contains poly(A)<sup>+</sup> RNA isolated from 100  $\mu$ g of total adult RNA (see Materials and Methods). Lanes 1 to 3 contain poly(A)<sup>+</sup> RNA from three independent transformant lines bearing the pDVTATA construction. Lane 4, wild-type RNA. Lane 5, *v*<sup>36f</sup> *ry*<sup>506</sup> RNA (injection stock). The blots were probed with anti-sense RNA fragments derived from a 3' *v* cDNA clone and from an *rp49* genomic clone.

lacks a TATA box activates  $\nu$  transcription about threefold. S1 nuclease protection analysis was used to map the 5' end of pDVTATA-derived transcripts and revealed a protection pattern identical to the wild-type pattern (data not shown). This result indicates that the addition of an artificial TATA box at position -25 does not alter the transcription start site.

### DISCUSSION

The results presented in this report defined regions upstream and downstream of the transcription start site that control  $\nu$  transcription in vivo. While no attempt was made in this study to specifically delineate a tissue-specific regulatory element, the analysis of  $\nu$ -*lacZ* fusions indicates that expression in the fat body is controlled by upstream sequences. An upstream element in the *D. melanogaster Adh* gene has previously been shown to be responsible for larval fat body-specific expression of that gene (8). The 88-bp *Adh* fat body-specific element contains a 17-bp sequence that is very similar (15 of 17 match) to a sequence found in the 5'-flanking region of the  $\nu$  gene, between 825 and 841 bp upstream of the transcription start site (see reference 27). However, deletion analysis of the  $\nu$  gene indicates that sequences upstream of -600 are not critical for  $\nu$  expression as measured both by phenotypic rescue, which is a reflection of larval expression, and by adult RNA analysis.

Analyses of internal deletion and linker substitution derivatives indicate that sequences in the leader region between +19 to +36 are essential for  $\nu$  expression. The addition of a TATA box to a construction lacking the  $\nu$  internal element restores  $\nu$  expression to a level comparable to that of wild type. Furthermore, the addition of a synthetic TATA sequence to the  $\nu$  promoter results in a 3- to 5-fold elevation of  $\nu$  expression, whereas adding a TATA element to the leader deletion derivative results in a greater than 80-fold increase in  $\beta$ -galactosidase activity. These results suggest that the internal element is functionally equivalent to a TATA box and that it plays a role in transcription rather than RNA stability. The observation that transcription of a construction containing both the TATA box and internal element is higher than one having the internal element alone is similar to the results of Smale et al. (30, 31) on the murine terminal deoxynucleotidyl transferase gene (TdT). The TdT gene, which lacks a TATA box, has an 8-bp element spanning the transcription start site that is necessary and sufficient for basal levels of transcription. In their studies, the level of transcription from a chimeric promoter containing simian virus 40 GC boxes, an adenovirus major late promoter TATA box, and an initiator element is increased by three- to fivefold compared with a similar promoter construction containing simian virus 40 GC boxes and an initiator element.

The involvement of downstream sequences in transcription of several TATA-less genes has been reported. As discussed earlier, in *D. melanogaster*, developmentally regulated genes such as *Antennapedia*, *Ultrabithorax*, *engrailed*, and *E74* as well as the retrotransposons *gypsy* and *mdg 1* have been shown to contain sequences extending to +60 that are necessary for transcription initiation. These results suggest that in the absence of a TATA box, specific *cis*-acting elements located downstream of the transcription start site serve as a recognition site for binding of one or more transcription factors, the binding of which may be essential for assembly of the general transcription initiation complex. In fact, in the *mdg 1* promoter, the binding of a protein factor to the downstream +30 region has been shown

to be required for transcription initiation in vivo (1). Taken together, these findings suggest that a subset of RNA polymerase II-transcribed promoters requires specific sequences other than the canonical TATA box. A comparison of promoter-region sequences of these genes revealed no striking homology. It was observed, however, that  $\nu$  sequences between +19 and +29 match sequences in the same region of the engrailed gene at 9 of 11 positions. The lack of an obvious consensus in the downstream regions among these TATA-less promoters suggests that several different factors interact with internal sequences of these genes.

The addition of a TATA box into the  $\nu$  gene did not alter the transcription initiation site. Perhaps that is because sequences other than the TATA box or the  $\nu$  downstream element specify the point of initiation. The sequences at the  $\nu$  transcription start site (ATCAGTTC) are quite similar to those found at other *D. melanogaster* genes (NTCAGTYN) (31). These sequences also resemble the initiator element of the murine TdT gene (CTCANTCT) (31). It has been demonstrated that the 8-bp initiator element spanning the TdT transcription start site can alone control basal levels of accurate transcription and respond to transcriptional activation by an upstream binding factor (30). Perhaps sequences spanning the  $\nu$  transcription start site are involved in specifying the site of initiation; however, these sequences are not apparently sufficient for low levels of expression because, as previously indicated, a  $\nu$ -*lacZ* construction containing sequences from -1114 to +28 failed to produce detectable levels of activity. In summary, based on the observations presented in this report, we propose a model for  $\nu$  gene regulation in which the internal element, located in the  $\nu$  transcribed region, is the binding site for a specific transcription factor, perhaps a TATA-binding protein-associated factor. Protein-DNA interactions involving this element may stabilize the association of TATA-binding protein with the promoter and recruit the general transcription machinery, the transcription start site being specified by sequences near the point of initiation.

### ACKNOWLEDGMENTS

We thank J. Harwell for assistance with the initial germ line transformation experiments. We thank B. Kay, G. Maroni, and M. Peifer for critically reading the manuscript, and S. Whitfield for the artwork.

This work was supported by grant DMB-9004708 from the National Science Foundation.

### REFERENCES

1. Arkhipova, I. R., and Y. V. Ilyin. 1991. Properties of promoter regions of *mdg 1* *Drosophila* retrotransposon indicate that it belongs to a specific class of promoters. *EMBO J.* **10**:1169-1177.
2. Biggin, M. D., and R. Tjian. 1988. Transcription factors that activate the Ultrabithorax promoter in developmentally staged extracts. *Cell* **53**:699-711.
3. Bingham, P. M., R. Lewis, and G. M. Rubin. 1981. Cloning of DNA sequences from the white locus of *D. melanogaster* by a novel and general method. *Cell* **25**:693-704.
4. Buratowski, S., S. Hahn, L. Guarente, and P. A. Sharp. 1989. Five intermediate complexes in transcription initiation by RNA polymerase II. *Cell* **56**:549-561.
5. Carcamo, J., S. Lobos, A. Merino, L. Buckbinder, R. Weinmann, V. Natarajan, and D. Reinberg. 1989. Factors involved in specific transcription by mammalian RNA polymerase II. *J. Biol. Chem.* **264**:7704-7714.
6. Dynlacht, B. D., T. Hoey, and R. Tjian. 1991. Isolation of coactivators associated with the TATA-binding protein that mediate transcriptional activation. *Cell* **66**:563-576.
7. Feinberg, A. P., and B. Vogelstein. 1983. A technique for

- radiolabeling DNA restriction endonuclease fragments to high specific activity. *Anal. Biochem.* **132**:6–13.
8. Fischer, J. A., and T. Maniatis. 1988. *Drosophila Adh*: a promoter element expands the tissue specificity of an enhancer. *Cell* **53**:451–461.
  9. Glaser, R. L., M. F. Wolfner, and J. T. Lis. 1986. Spatial and temporal pattern of *hsp26* expression during normal development. *EMBO J.* **5**:747–754.
  10. Higuchi, R. 1990. Recombinant PCR, p. 177–183. *In* M. A. Innis, D. H. Gelfand, J. J. Sninsky, and T. J. White (ed.), PCR protocol, a guide to methods and applications, Academic Press, Inc., San Diego.
  11. Horikoshi, M., T. Yamamoto, Y. Ohkuma, P. A. Weil, and R. G. Roeder. 1990. Analysis of structure-function relationships of yeast TATA box binding factor TFIID. *Cell* **61**:1171–1178.
  12. Jarrell, K. A., and M. Meselsson. 1991. *Drosophila* retrotransposon promoter includes an essential sequence at the initiation site and requires a downstream sequence for full activity. *Proc. Natl. Acad. Sci. USA* **88**:102–104.
  13. Karess, R. E. 1985. P element mediated germ line transformation of *Drosophila*, p. 121–141. *In* D. M. Glover (ed.), DNA cloning, a practical approach, vol. 2. IRL Press, Oxford.
  14. Maniatis, T., E. F. Fritsch, and J. Sambrook. 1982. Molecular cloning: a laboratory manual. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
  15. Melton, D. A., P. A. Krieg, M. R. Rebagliati, T. Maniatis, K. Zinn, and M. R. Green. 1984. Efficient in vitro synthesis of biologically active RNA and RNA hybridization probes from plasmids containing a bacteriophage SP6 promoter. *Nucleic Acids Res.* **12**:7035–7056.
  16. Mismar, D., and G. M. Rubin. 1987. Analysis of the promoter of the *nina E* opsin gene in *Drosophila melanogaster*. *Genetics* **116**:565–578.
  17. O'Connell, L. P., and M. Rosbash. 1984. Sequence, structure, and codon preference of the *Drosophila ribosomal protein 49* gene. *Nucleic Acids Res.* **12**:5495–5513.
  18. Perkins, K. K., G. 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.
  19. Peterson, M. G., N. Tanese, B. F. Pugh, and R. Tjian. 1990. Functional domains and upstream activation properties of cloned human TATA binding protein. *Science* **248**:1625–1630.
  20. Phillips, J. P., and H. S. Forrester. 1980. Ommochromes and pteridines, p. 542–611. *In* M. Ashburner and T. R. F. Wright (ed.), The genetics and biology of *Drosophila*, vol. 2D. Academic Press, Inc. (London), Ltd., London.
  21. Pugh, B. F., and R. Tjian. 1990. Mechanism of transcriptional activation by Sp1: evidence for coactivators. *Cell* **61**:1187–1197.
  22. Pugh, B. F., and R. Tjian. 1991. Transcription from a TATA-less promoter requires a multisubunit TFIID complex. *Genes Dev.* **5**:1935–1945.
  23. Puissant, C., and L.-M. Houdebine. 1990. An improvement of the single-step method of RNA isolation by acid guanidinium thiocyanate-phenol-chloroform extraction. *BioTechniques* **8**:148–149.
  24. Rizki, T. M. 1978. Fat body, p. 541–621. *In* M. Ashburner and T. R. F. Wright (ed.), The genetics and biology of *Drosophila*, vol. 2B. Academic Press, Inc. (London), Ltd., London.
  25. Rubin, G. M., and A. C. Spradling. 1982. Genetic transformation of *Drosophila* with transposable element vectors. *Science* **218**:348–353.
  26. Scharf, S. J. 1990. Cloning with PCR, p. 84–91. *In* M. A. Innis, D. H. Gelfand, J. J. Sninsky, and T. J. White (ed.), PCR protocols, a guide to methods and applications. Academic Press, Inc., San Diego.
  27. Searles, L. L., R. S. Ruth, A.-M. Pret, R. A. Fridell, and A. J. Ali. 1990. Structure and transcription of the *Drosophila melanogaster* vermilion gene and several mutant alleles. *Mol. Cell. Biol.* **10**:1423–1431.
  28. Searles, L. L., and R. A. Voelker. 1986. Molecular characterization of the *Drosophila* vermilion locus and its suppressible alleles. *Proc. Natl. Acad. Sci. USA* **83**:404–408.
  29. Simon, J. A., and J. T. Lis. 1987. A germline transformation analysis reveals flexibility in the organization of heat shock consensus elements. *Nucleic Acids Res.* **15**:2971–2988.
  30. Smale, S. T., and D. Baltimore. 1989. The “initiator” as a transcription control element. *Cell* **57**:103–113.
  31. Smale, S. T., M. Schmidt, A. Berk, and D. Baltimore. 1990. Transcriptional activation by Sp1 as directed through TATA or initiator: specific requirement for mammalian transcription factor IID. *Proc. Natl. Acad. Sci. USA* **87**:4509–4513.
  32. Soeller, W. C., S. Poole, and T. Kornberg. 1987. In vitro transcription of the *Drosophila engrailed* gene. *Genes Dev.* **2**:68–81.
  33. Southern, E. M. 1975. Detection of specific sequences among DNA fragments separated by gel electrophoresis. *J. Mol. Biol.* **98**:503–517.
  34. Thomas, P. S. 1983. Hybridization of denatured RNA transferred or dotted to nitrocellulose paper. *Methods Enzymol.* **100**:255–266.
  35. Thummel, C. S. 1989. The *Drosophila E74* promoter contains essential sequences downstream from the start site of transcription. *Genes Dev.* **3**:782–792.
  36. Thummel, C. S., A. M. Boulet, and H. D. Lip. 1988. Vectors for *Drosophila* P-element-mediated transformation and tissue culture transfection. *Gene* **74**:445–456.
  37. Van Dyke, M. W., R. G. Roeder, and M. Sawadogo. 1988. Physical analysis of transcription preinitiation complex assembly on a class II gene promoter. *Science* **241**:1335–1338.
  38. Walker, A. R., A. J. Howells, and R. G. Tearle. 1986. Cloning and characterization of the *vermilion* gene of *Drosophila melanogaster*. *Mol. Gen. Genet.* **202**:102–107.
  39. Zinn, K., D. Dimaio, and T. Maniatis. 1983. Identification of two distinct regulatory regions adjacent to the human  $\beta$ -interferon gene. *Cell* **34**:865–879.