MOLECULAR AND CELLULAR BIOLOGY, Dec. 1986, p. 4548–4557
0270-7306/86/124548-10$02.00/0

Vol. 6, No. 12

Delimiting Regulatory Sequences of the *Drosophila melanogaster* *Ddc* Gene

J. HIRSH,[1]* B. A. MORGAN,[1,2] AND S. B. SCHOLNICK[1]†

Department of Biological Chemistry,[1] and Program in Cell and Developmental Biology,[2] Harvard Medical School, Boston, Massachusetts 02115

Received 30 January 1986/Accepted 10 September 1986

We delimited sequences necessary for in vivo expression of the *Drosophila melanogaster* dopa decarboxylase gene *Ddc*. The expression of in vitro-altered genes was assayed following germ line integration via P-element vectors. Sequences between −209 and −24 were necessary for normally regulated expression, although genes lacking these sequences could be expressed at 10 to 50% of wild-type levels at specific developmental times. These genes showed components of normal developmental expression, which suggests that they retain some regulatory elements. All *Ddc* genes lacking the normal immediate 5′-flanking sequences were grossly deficient in larval central nervous system expression. Thus, this upstream region must contain at least one element necessary for this expression. A mutated *Ddc* gene without a normal TATA boxlike sequence used the normal RNA start points, indicating that this sequences is not required for start point specificity.

The mechanisms by which metazoan organisms express given genes in particular tissues and cell types and at specific times during development remain largely unresolved. Detailed knowledge of these basic mechanisms will contribute to our overall understanding of the developmental processes that shape an organism. A first step in elucidating these mechanisms is to define the *cis*-acting genetic elements necessary for proper regulation.

A method has been developed to stably reintroduce genes into the germ line of the fruit fly *Drosophila melanogaster* by using a naturally occurring *Drosophila* transposable element vector (30, 35). Several *Drosophila* genes reintegrated in this manner have shown nearly normal developmental expression, with generally small effects of chromosomal position on gene expression (11, 20, 29, 33, 35). Thus, a means now exists for the biological assay of *cis*-acting genetic elements in *D. melanogaster*.

The *Ddc* gene encodes the enzyme dopa decarboxylase, a central enzyme in catecholamine metabolism. *Ddc* is expressed both in the hypoderm and in the central nervous system. The gene must therefore contain an element(s) specifying expression in two different tissues and at multiple times during development. In the hypoderm metabolites produced by the *Ddc* gene product are used in pathways leading to pigmentation and cross-linking of the cuticle (for a review, see reference 5). The temporal course of *Ddc* enzyme activity during development primarily reflects this use. Peaks of activity are found at each stage in which extensive cuticular pigmentation and hardening occurs (17, 21). In the central nervous system, *Ddc* enzyme activity leads to the synthesis of serotonin as well as dopamine (18, 38a) and is required for normal learning behavior in *D. melanogaster* (37). During at least certain developmental times the temporal profile of *Ddc* expression in the central nervous system differs from that in the hypoderm (14; unpublished data).

Results of previous work demonstrate that all sequences necessary for normal developmental expression of *Ddc* lie within a 7.5-kilobase (kb) *Pst*I fragment containing the 4-kb *Ddc* structural gene, 2.8 kb of 5′-flanking sequences, and 0.7 kb of 3′-flanking sequences (20, 33). The expression of *Ddc* is likely to be regulated at multiple steps, since posttranscriptional mechanisms participate in regulating developmental expression (2; B. A. Morgan, W. A. Johnson, and J. Hirsh, EMBO J., in press). We have begun to delineate sequences that are essential for the tissue-specific and temporally correct expression of *Ddc* by assaying the developmental expression of altered *Ddc* genes that are stably integrated in the *D. melanogaster* germ line.

MATERIALS AND METHODS

Nucleic acid manipulations. Recombinant DNA methods were carried out, in general, by the methods of Maniatis et al. (19). Restriction enzymes were from New England BioLabs, Inc. (Beverly, Mass.) or Boehringer Mannheim Biochemicals (Indianapolis, Ind.) under the conditions recommended by the manufacturers. Avian myelobastosis virus reverse transcriptase was purchased from Life Sciences, Inc. (St. Petersburg, Fla.). Northern analysis of RNA was carried out as described by O'Tousa et al. (25). The oligonucleotide primer complementary to *Ddc* RNA was synthesized by Alex Nussbaum.

Construction of 5′ end-deleted *Ddc* genes. The starting material for all 5′-end-deleted *Ddc* genes was the plasmid subclone p1099, consisting of the 7.5-kb *Pst*I fragment encompassing *Ddc* in the vector pUC8 (39). The 5′-end-deleted genes were constructed by cleaving at appropriately positioned restriction sites. To construct the $Ddc^{5'\Delta-800}$ and $Ddc^{5'\Delta-384}$ genes, p1099 DNA was cleaved partially with either *Hin*dIII or *Eco*RI, respectively. To construct the $Ddc^{5'\Delta-209}$ and $Ddc^{5'\Delta-25}$ genes, the DNA was digested to completion with either *Eco*RV or *Nco*I. All DNAs were then cleaved at the *Pst*I site 3′ from *Ddc*, and the relevant *Ddc* fragments were isolated. *Pst*I linkers (New England Biolabs) were then ligated to the ends, and the relevant fragments were recloned into pUC8.

The $Ddc^{\Delta(-357,-34)}$ gene was constructed by deleting a segment between two *Aha*III (*Dra*I) cleavage sites. DNA from a $Ddc^{5'\Delta-800}$ gene cloned in pUC8 was partially cut with *Aha*III, and the full-length minus 300-bp band region was

* Corresponding author.
† Present address: Carnegie Mellon University, Department of Biological Sciences, Pittsburgh, PA 15213.

isolated. This material was then circularized with DNA ligase. The desired deletion removed a unique *Eco*RV site at −208, such that it was possible to enrich for the deleted molecules by cleaving with *Eco*RV following ligation but before transformation into *Escherichia coli*.

The altered *Ddc* genes were inserted into derivatives of the vector pSXΔ1, a gift from Jim Posakony. pSXΔ1 is a deletion derivative of the intact P-element π25.1 (23) containing a polylinker from the vector πvx (34) in the P-element *Eco*RI site. We destroyed one of the two *Pst*I sites within pSXΔI and inserted a 4.8-kb *Eco*RI fragment containing the alcohol dehydrogenase structural gene *Adh* to yield vectors containing *Adh* in either orientation. The deleted *Ddc* genes were cloned into the unique *Pst*I site of these vectors. The structures of the resulting vectors, containing the in vitro-altered *Ddc* genes and Adh^+ genes to allow selection of transformants, are shown in Fig. 1.

Primer extension mapping of RNAs. All primer extension mapping was performed by using hybridization to detect the cDNA products. Unlabeled cDNAs were synthesized from synthetic oligonucleotide primers. The products were electrophoretically separated on polyacrylamide sequencing gels and then eletrophoretically transferred to GeneScreen filters (New England Nuclear Corp., Boston, Mass.) by the method of Church and Gilbert (8). Following UV cross-linking of the products to the filter (8), the extension products were detected by hybridization with appropriate probes. This method allows detection of low-abundance extension products from primers which are themselves too short to hybridize specifically to a unique sequence in a complex RNA population. It also allows the use of high deoxyribonucleotide concentrations during the initial extension reaction, thereby promoting the synthesis of full-length extension products. Finally, it allows the specific detection of transcripts from hybrid genes in a background where the normal gene transcripts are present.

To synthesize the unlabeled cDNAs, 5 pmol of a 15mer primer and 5 μg of poly(A)$^+$ RNA were dissolved in 20 μl of 50% formamide–40 mM PIPES [piperazine-*N*, *N*′-bis(2-ethanesulfonic acid); pH 6.4]–400 mM NaCl–1 mM EDTA. The mixture was heated to 65°C for 1 min and then at 37°C for 1 h. The RNA was ethanol precipitated and suspended in 20 μl of 50 mM Tris (pH 8.3)–10 mM $MgCl_2$–70 mM KCl–10 mM dithiothreitol–500 μM each of the four unlabeled deoxynucleoside triphosphates. To this was added 7.5 to 15 U of reverse transcriptase, and the reaction was incubated for 15 min at 20°C, for 15 min at 30°C, and for 1 h at 37°C. The reaction was then ethanol precipitated, heat denatured, and run on a 6% polyacrylamide-urea sequencing gel. The nucleic acid was electrophoretically transferred in 0.5× TBE (1× TBE is 0.05 M Tris-borate plus 0.01 M EDTA, pH 8.3) to GeneScreen (New England Nuclear), UV cross-linked, and hybridized by using the conditions given by Church and Gilbert (8).

To avoid bubbles during the electrophoretic transfer, we used a custom-built electrophoresis tank (Fred Gilchrist, Department of Biological Chemistry Shop, Harvard Medical School) that allowed the gel-filter sandwich to be assembled while submerged in buffer and while the apparatus was horizontal, yet it could be run in the normal vertical manner.

The oligonucleotide primer used for determining the 5′ end of *Ddc* transcripts consisted of the sequence 5′-TTGCCCT CCTGGTTT. This sequence is complementary to *Ddc* transcripts, starting at a position about 100 bases downstream from the RNA start point (33).

All nucleic acid sequences reported here were determined by dideoxynucleotide sequencing (31). Both strands of the relevant regions were sequenced in all instances.

Integration of *Ddc* P-element vectors into *D. melanogaster* and assays of integrated genes. Methods for integration of *Ddc* P-element vectors into *D. melanogaster* and assays of integrated genes were as described previously (30, 33), except as indicated. Injection needles were formed from 1.2-mm microstar capillaries (1812; Radnoti Glass, Arcadia, Calif.) with an electrode puller (Narishige). Needles were sharpened by chipping against the rough edge of a broken cover slip.

The *Ddc-Adh* vectors were injected at a concentration of 350 μg/ml, while the helper P element, either π25.1 (23) or "wings-clipped" π25.7 (16), was at 75 μg/ml. A host strain of genotype Ddc^{ts2} Adh^{fn23} was constructed. This host strain allowed for the selection of transformed flies with ethanol (40) and measurement of *Ddc* enzyme activity from the reintegrated genes, since the Ddc^{ts2} alleles were expressed at only several percent of the wild-type levels (33, 42). This strain was of adequate viability when grown at or below 20°C. Better growth was obtained with instant fly media (Carolina Biologicals) than with media containing higher concentrations of yeasts.

Ethanol-resistant flies were mated and grown into stocks. The stocks were characterized by Southern blot analysis of genomic DNA to check for gross sequence rearrangements, to confirm the identity of the in vitro sequence alteration, and as an initial indication of copy number (33). Homozygous stocks were obtained either by repeated intermatings or by crosses to appropriately marked balancer chromosome strains. The stocks were also characterized by in situ hybridization to polytene salivary chromosomes with *Ddc* gene probes, as described previously (15).

Ddc enzyme assays were done as described previously (33), except that the enzyme assay buffer consisted of 100 mM sodium phosphate (pH 6.8)–0.5 mM levodopa–0.1 mM pyridoxal 5′ phosphate, and 1 mM β-mercaptoethanol–1.0 mM EDTA was used to stabilize the levodopa (24). One specific activity unit of enzyme led to the formation of 1 nmol of dopamine per 20 min of incubation at 30°C per mg of protein (32). Assays of hypodermal tissue specificity used the rubber roller assay described by Scholnick et al. (33). Assays were performed on groups of five individual flies or on single flies as indicated.

RESULTS

Defining the sequences necessary for temporally regulated *Ddc* expression. To define the 5′-flanking sequences necessary for normal regulation, we constructed deleted genes retaining 800, 383, 208, or 24 bp of upstream-flanking DNA and one internally deleted gene, removing upstream sequences from −357 through −34 relative to the RNA start point (Fig. 1). This latter gene retains a TATA boxlike sequence at −33 (see Fig. 3) and *Ddc* upstream sequences from the deletion endpoint to −800. The deleted *Ddc* genes were inserted into Adh^+ P-element vectors and reintegrated into a host strain carrying mutant alleles of both the *Adh* and *Ddc* genes. Transformants were identified using an ethanol selection for Adh^+ function (40), and homozygous lines were established. A summary of the transformed lines discussed in this report is shown in Fig. 1.

Figure 2A shows developmental profiles of *Ddc* enzyme activity in a wild-type Canton-S strain and in two lines containing $Ddc^{5'\Delta\text{-}209}$ genes. Both of the $Ddc^{5'\Delta\text{-}209}$ lines shown in Fig. 2A and, similarly, all lines containing less

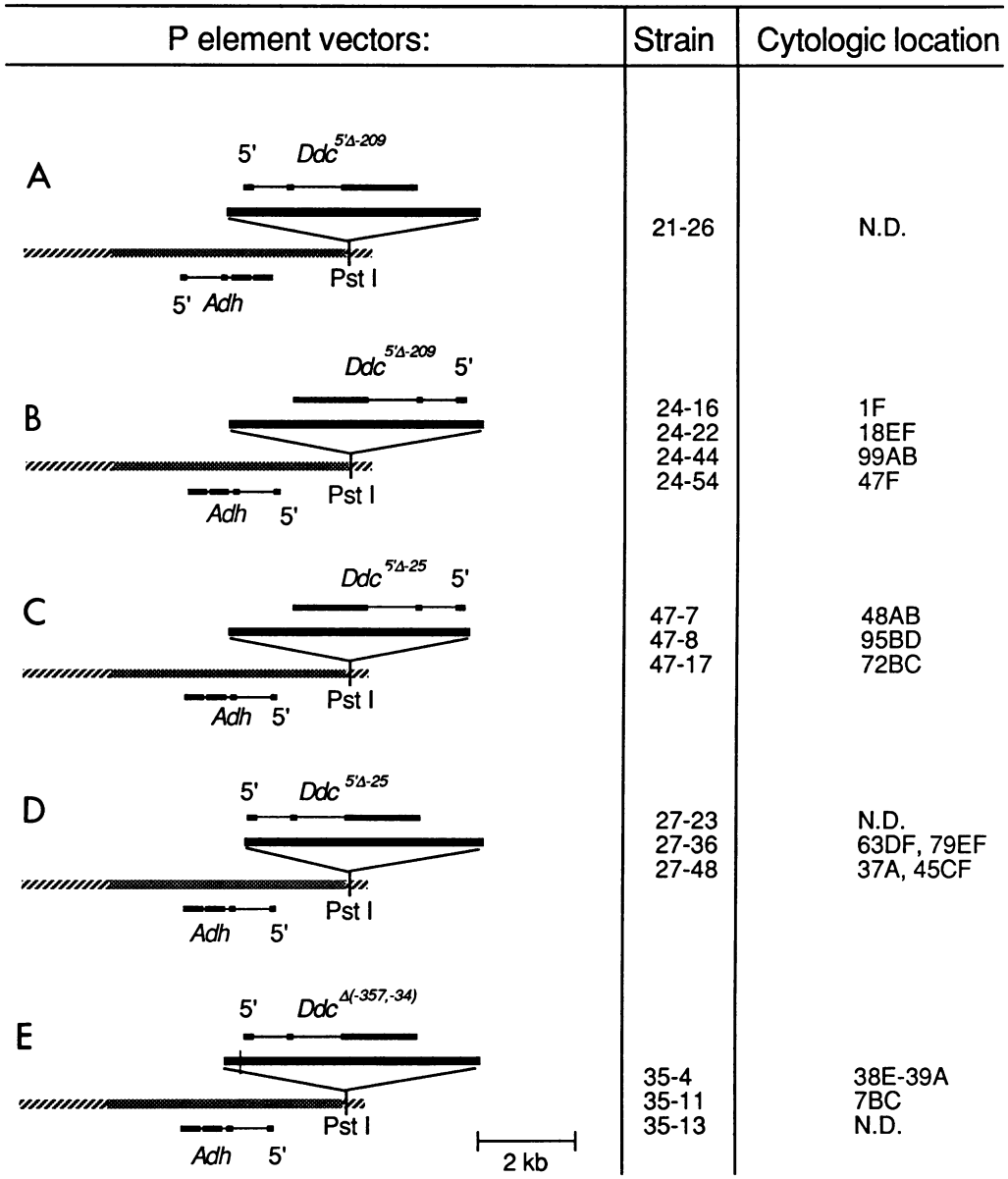

FIG. 1. P-element vectors, strains containing the inserted genes, and cytologic locations of inserts. Shown are the P-element vector sequences that integrate into *D. melanogaster* chromosomal DNA. The plasmids containing the vectors contain in addition a small amount of DNA from the *D. melanogaster* white locus and plasmid pUC8 sequences. The origins of and symbols for the DNA segments are as follows: P-element sequences, striped; *Adh*, stippled; *Ddc*, solid. The exons of the *Adh* (3, 13) and *Ddc* genes (2) are sketched above and below the drawings of the constructs. The vertical line in the $Ddc^{\Delta(-357,-34)}$ *Ddc* gene represents the site of deletion. ND, Not determined. (A and B) Adh^{+} $Ddh^{5'\Delta-209}$ *Ddc* vectors. (C and D) Adh^{+} $Ddc^{5'\Delta-25}$ *Ddc* vectors. Note that the strains 27-36 and 27-48 each contain two inserted P elements. (E) Adh^{+} $Ddc^{\Delta(-357,-34)}$ *Ddc* vector.

severe deletions (data not shown) show a pattern of developmental expression similar to that of the wild type. The observed variations from wild-type expression, such as the overexpression of *Ddc* activity during embryogenesis in the $Ddc^{5'\Delta-209}$ strain 24-44, are presumably caused by effects of chromosomal position. It does appear, however, that these genes contain all sequences sufficient for normal temporal modulation of *Ddc* expression.

Figure 2B shows the developmental expression of three lines carrying $Ddc^{5'\Delta-25}$ genes in the same relative orientation within the P-element vector as the $Ddc^{5'\Delta-209}$ strains described above. These lines show *Ddc* expression that is regulated in a temporally abnormal manner at levels that are about 10% of normal at pupariation and adult eclosion. (Note the expanded scale relative to that in Fig. 2A.) Induction of *Ddc* enzyme activity during embryogenesis was retarded by 4 h, although the times of hatching in these strains were unaffected. *Ddc* enzyme levels were near normal during larval development, but there was no significant induction at pupariation. The small increase in enzyme activity at pupariation can be accounted for by expression from the Ddc^{ts2} genes of the host strain. Levels of *Ddc* enzyme activity fell during early pupal development and increased near the time of adult eclosion, although to reduced levels

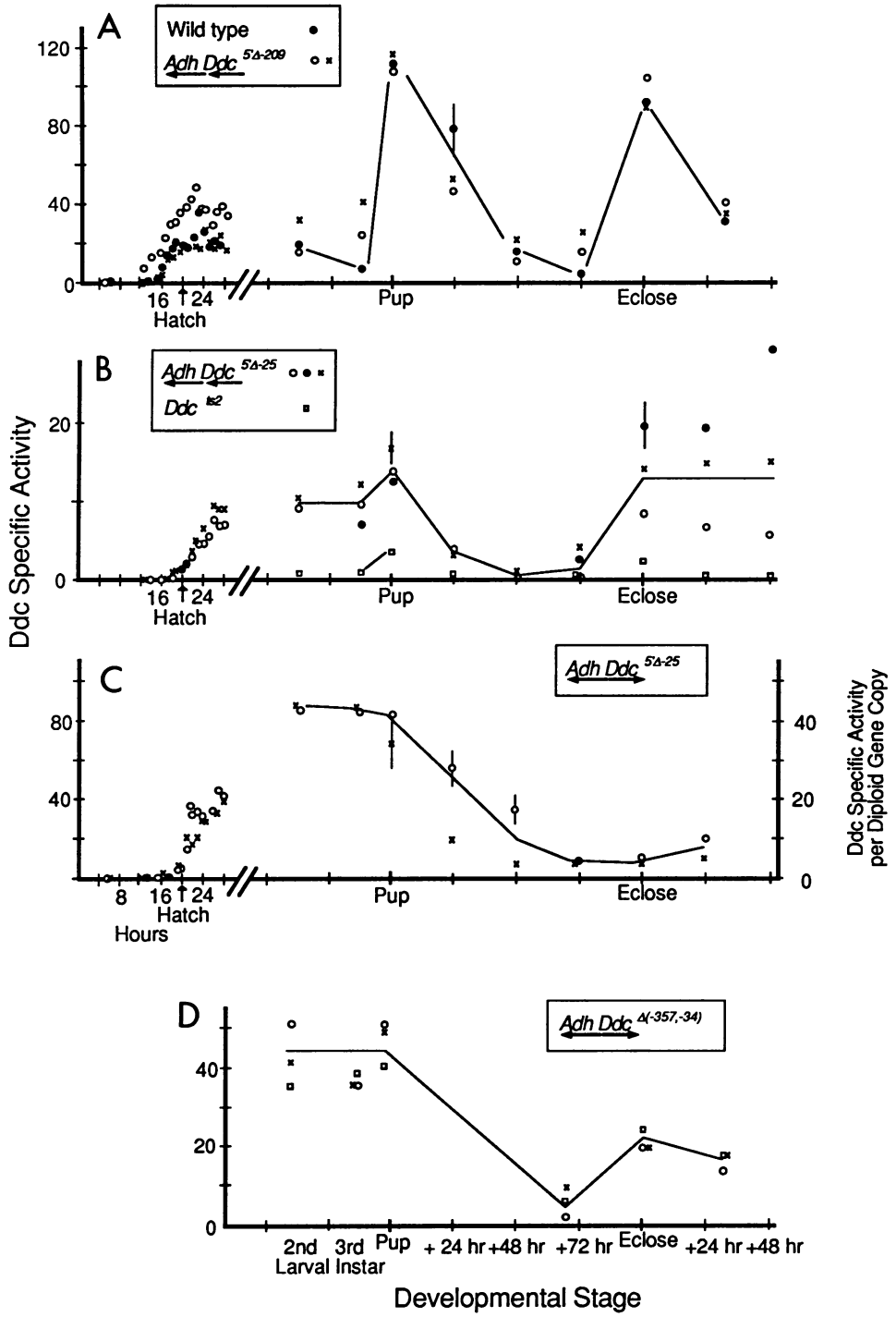

FIG. 2. Developmental profiles of *Ddc* enzyme activity. Data shown are for homozygous strains. Time points are as follows: time of embryonic hatching (Hatch); pupariation (Pup); adult eclosion (Eclose). Embryonic time points were taken from collections every 2 h; points are plotted at the mean age of each sample. All postembryonic data points were compiled from multiple samples. Error bars are shown when errors exceeded 10% of the mean. Data were derived from the following strains. (A) Canton-S (●); $Ddc^{5'\Delta-209}$ genes oriented with the 5′ end distal from *Adh* in the P-element vector: strain 24-44 (○); strain 24-54 (×). (B) $Ddc^{5'\Delta-25}$ strains with the *Ddc* gene oriented with the 5′ end distal to *Adh*: strain 47-7 (○) strain 47-8 (×); strain 47-17 (●); basal level of expression in the Ddc^{ts2} host strain (□). (C) $Ddc^{5'-\Delta-25}$ strains with the *Ddc* gene oriented with the 5′ end proximal to *Adh*: strain 27-48 (○); strain 27-48 (×). Each of these strains contains two reintegrated genes such that the levels of enzyme activity must be divided by 2 to allow direct comparison of expression per gene copy with the other strains in this figure. This normalized scale is shown at the right. (D) $Ddc^{\Delta(-357,-34)}$ strains with the *Ddc* gene oriented with the 5′ end proximal to *Adh*: strain 35-4 (○); strain 35-11 (×); strain 35-13 (□).

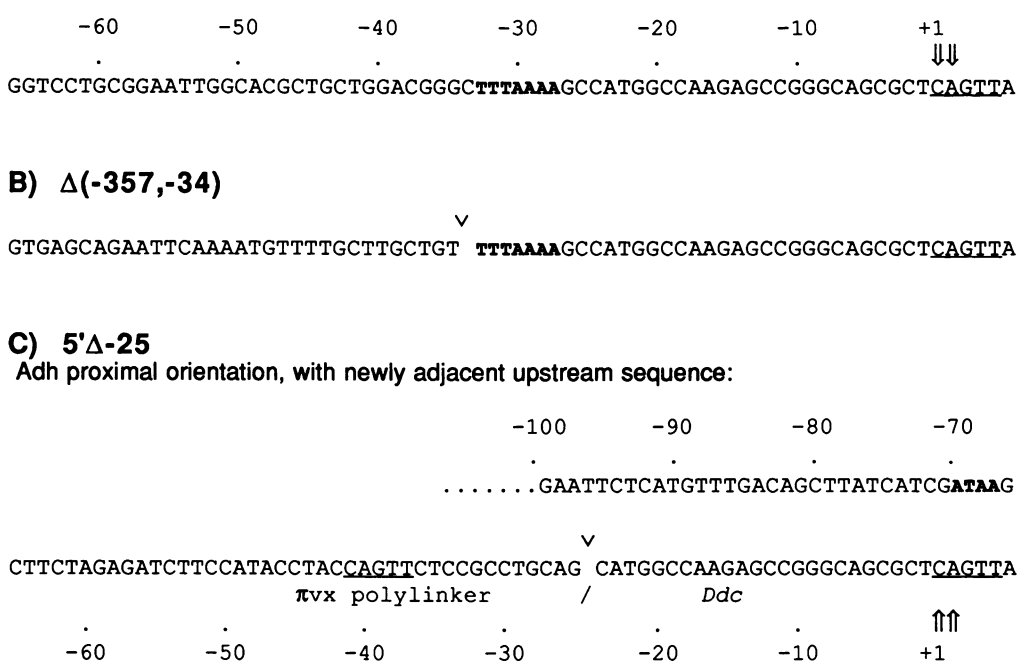

FIG. 3. The 5′-flanking sequences of wild-type and deleted *Ddc* genes. Deletions are numbered indicating the deleted bases *x* through *y*, Δ(*x*,*y*). Sequences showing homology to the insect start point consensus sequence (14) are underlined, and TATA-like sequences are shown in boldface type. The major RNA start points are indicated by arrows at bases 1 and 2. Deletions are marked by inverted carets. The deletion from −34 was drawn and numbered as if the deletion occurred upstream from the normal TATA box.

relative to the wild type. In contrast to normally regulated expression, there was no significant decrease in enzyme activity during adult development. Thus, removal of sequences between 24 and 209 bp upstream from the RNA start point affected both the level and several aspects of the developmental profile of expression.

Increased levels of expression were observed when the $Ddc^{5'\Delta-25}$ genes were oriented with the 5′ end proximal to the *Adh* gene (Fig. 2C). Both of these strains contained two reintegrated genes; the scale on the right of Fig. 2C should be used to compare gene expression per gene copy with other strains shown in Fig. 2C. *Ddc* enzyme activity in these strains is about 40% of normal at pupariation when averaged on a per gene basis. As was observed when *Ddc* was oriented with the 5′ end distal to *Adh*, induction was retarded by about 4 h during embryonic development, and there was no significant induction at pupariation. Levels of enzyme activity per gene copy were three- to fivefold higher during larval development than in the initial orientation and were abnormally high relative to that of the wild type. *Ddc* enzyme activity decreased after pupariation with nearly normal kinetics, but none of the lines was induced significantly at adult eclosion. Thus, the $Ddc^{5'\Delta-25}$ genes showed similar, though not identical, aberrations in developmental expression in both orientations relative to vector sequences, although orientation can have had a significant effect on the level of gene expression.

Restoring a TATA box to a deleted *Ddc* gene. Expression of the $Ddc^{\Delta(-357,-34)}$ gene was examined to determine whether the aberrant developmental regulation observed with the $Ddc^{5'\Delta-25}$ genes was due solely to the lack of a TATA box sequence. The only TATA boxlike sequence 5′ of *Ddc* was a sequence TTTAAAA at −33 relative to the RNA startpoint (Fig. 3A). This sequence was deleted from the $Ddc^{5'\Delta-25}$ genes (Fig. 3C) yet was retained in the $Ddc^{\Delta(-357,-34)}$ genes (Fig. 3B). Developmental profiles of three strains containing $Ddc^{\Delta(-357,-34)}$ genes are shown in Fig. 2D. The developmental profiles of these strains resembled those of the strains containing similarly oriented $Ddc^{5'\Delta-25}$ genes in that they showed inappropriately high levels of expression during larval development and no induction at pupariation; levels of expression at pupariation were 35 to 55% those of the wild type. However, in contrast to the similarly oriented $Ddc^{5'\Delta-25}$ genes, these genes showed a small degree of induction at adult eclosion. Thus, retention of the TATA-like sequence has relatively little effect on expression of these genes, and its absence is not the cause of the aberrant developmental expression of the $Ddc^{5'\Delta-25}$ genes.

***Ddc* RNA start point selection does not require a TATA box.** The orientation dependence of the $Ddc^{5'\Delta-25}$ expression suggests that a portion of the increased expression observed when the 5′ end of *Ddc* is proximal to *Adh* might be caused by readthrough from the *Adh* gene segment. To explore this possibility we determined the RNA start point used by the $Ddc^{5'\Delta-25}$ *Ddc* genes in this orientation. Poly(A)$^+$ RNA was prepared from early third instar larvae. Expression of the host Ddc^{ts2} genes was low at this stage, so most of the *Ddc* transcripts must have originated from the reintegrated genes. The RNA start point was determined by cDNA extension from a synthetic oligonucleotide primer complementary to *Ddc* RNA sequences about 100 bases downstream from the RNA start point. The major start points occurred within the sequence CAGTT, a sequence that matches the consensus sequence ATCA$^{G}_{T}$T$^{C}_{T}$ found at the start point of several insect mRNAs (15a). The same start points are used by the wild-type gene during the late third instar development (Fig. 4, lanes A and B), and by the $Ddc^{5'\Delta-25}$ genes during early (Fig. 4, lanes D to F) or late (Fig. 4, lane C) third instar

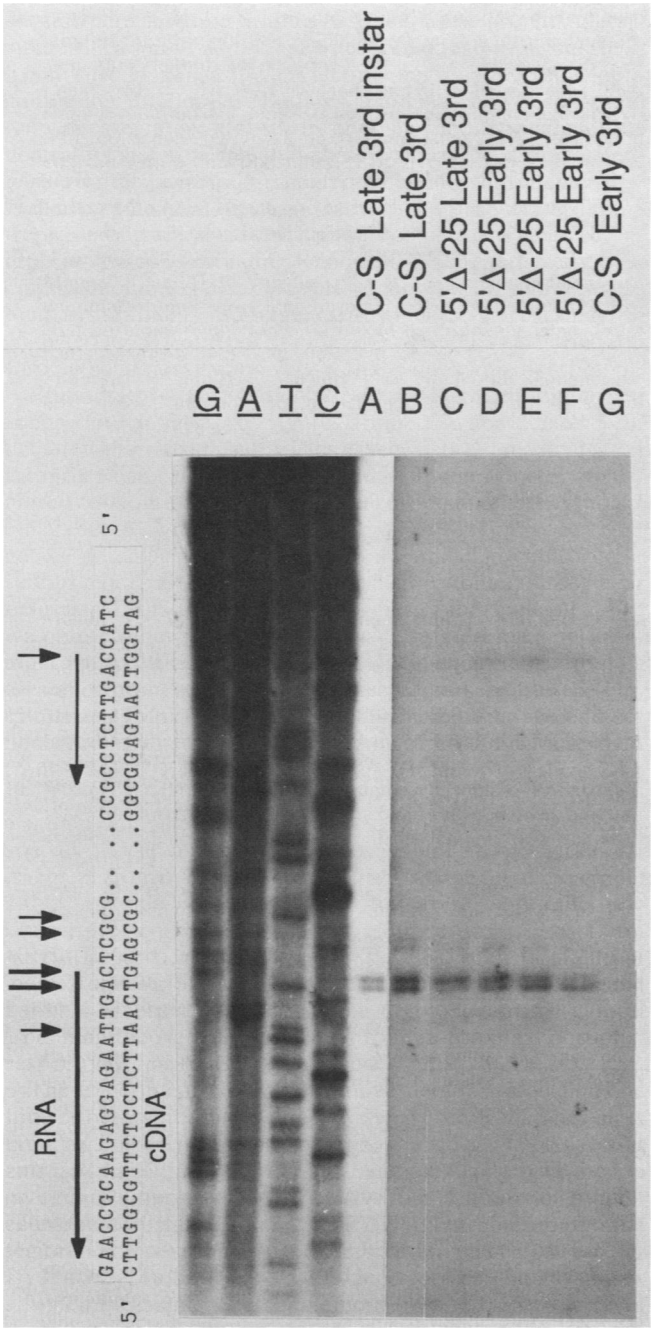

FIG. 4. RNA start points in wild-type and mutant strains. Poly(A)$^+$ RNA from the indicated strains and stages was used as template for primer extension from a synthetic oligonucleotide primer complementary to *Ddc* RNA. The products were detected with a *Ddc* gene probe after electrophoretic separation and transfer of the unlabeled extension products to a filter. The leftmost four lanes are size standards consisting of sequencing reactions that were produced by using a single-stranded *Ddc* DNA template in the presence of the indicated dideoxynucleotides. The sequence at the left shows the cDNA sequence and the inferred RNA sequence of the $Ddc^{5'\Delta-25}$ gene, with arrows indicating the observed start points. Lanes A and B, cDNAs synthesized from Canton-S (C-S) wild-type RNA; lane B, a darker exposure of lane A; lanes C to F, cDNAs synthesized from two 5′ Δ-25 strains by using early or late third instar larval RNA; lanes C and D, strain 27-36; lanes E and F, strain 27-48; lane G, a control showing that little *Ddc* RNA is present in the early third instar stage of a wild-type strain.

TABLE 1. Relative fraction of *Ddc* activity in the hypoderm

Construct	Strain	Fraction (%) of activity in hypoderm (Mean [individual larvae])	Pupal phenotype[a]
	Canton-S	94 (87, 92, 93, 99, 98)	+
Adh ← $Ddc^{5'\Delta-209}$ ←	24-16	78 (94, 84, 77, 70, 66)	+
	24-54	91 (92, 90, 92, 86, 94)	+
Adh → $Ddc^{5'\Delta-209}$ →	21-26	88 (86, 87, 86, 91, 90)	+
Adh ← $Ddc^{5'\Delta-25}$ →	27-23	73 (69, 90, 78, 58, 71)	±
	27-36	89 (96, 73, 99)	±
	27-48	77 (50, 88, 92, 79)	±
Adh ← $Ddc^{5'\Delta-25}$ ←	47-7	49[b]	~/−
	47-8	68[b]	~/−
Adh ← $Ddc^{\Delta(-357,-34)}$ →	35-4	53 (47, 51, 62)	+
	35-11	69 (87, 90, 46, 54)	+
	35-13	70 (56, 76, 82, 83, 54)	+

[a] Pupal phenotypes: +, wild type; ±, intermediate; ~/−, barely distinguishable from the Ddc^{ts2} phenotype.
[b] Assays were performed on pools of five larvae each.

development. In Fig. 4 (lane G) is a control showing the comparatively low level of *Ddc* transcripts in early third instar larvae from a wild-type strain. Thus, the enhanced expression from these severely deleted genes cannot be explained by readthrough transcription and must therefore be due to an element within the vector affecting the *Ddc* promoter. The $Ddc^{5'\Delta-25}$ genes do employ at low efficiency a start point about 40 bp upstream from the normal start point. This start point is within the adjacent πvx polylinker at a sequence CAGTT, which is identical to the sequence at the normal *Ddc* RNA start point.

Tissue-specific expression of the altered *Ddc* genes. We further characterized the developmental expression of the deleted *Ddc* genes by examining the tissue-specific expression of the resulting *Ddc* enzyme activity. The majority of *Ddc* activity is normally found in the hypoderm, where it functions in a pathway leading to pigmentation and cross-linking of the cuticle. The pupal cases of strains grossly deficient in *Ddc* enzyme activity are abnormally pigmented (33). Reconstruction experiments using low-activity *Ddc* alleles isolated by classical genetic methods (42) demonstrate that normal pigmentation can be observed in puparia that express *Ddc* enzyme activity at or above 15% of the wild-type levels (unpublished data). The pigmentation of all strains containing $Ddc^{5'\Delta-209}$ and $Ddc^{\Delta(-357,-34)}$ genes was indistinguishable from that of the wild type. However, strains containing $Ddc^{5'\Delta-25}$ genes in either orientation were abnormally pigmented, showing intermediate phenotypes that were not completely normal but that were clearly distinguishable from the severe phenotype of the Ddc^{ts2} host strain (Table 1).

These abnormal phenotypes imply that *Ddc* expression in the hypoderm is aberrant, since some abnormally pigmented strains contained nearly normal levels of *Ddc* enzyme activity at pupariation. However, no abnormality in tissue-specific expression correlating with the abnormal pigmentation was detectable by our crude tissue fractionation assay (Table 1). By this assay nearly normal hypodermal expression was found in most strains containing $Ddc^{5'\Delta-209}$ genes. However, some strains, such as 24-16, had a significantly lower fraction of activity in the hypoderm than did the wild

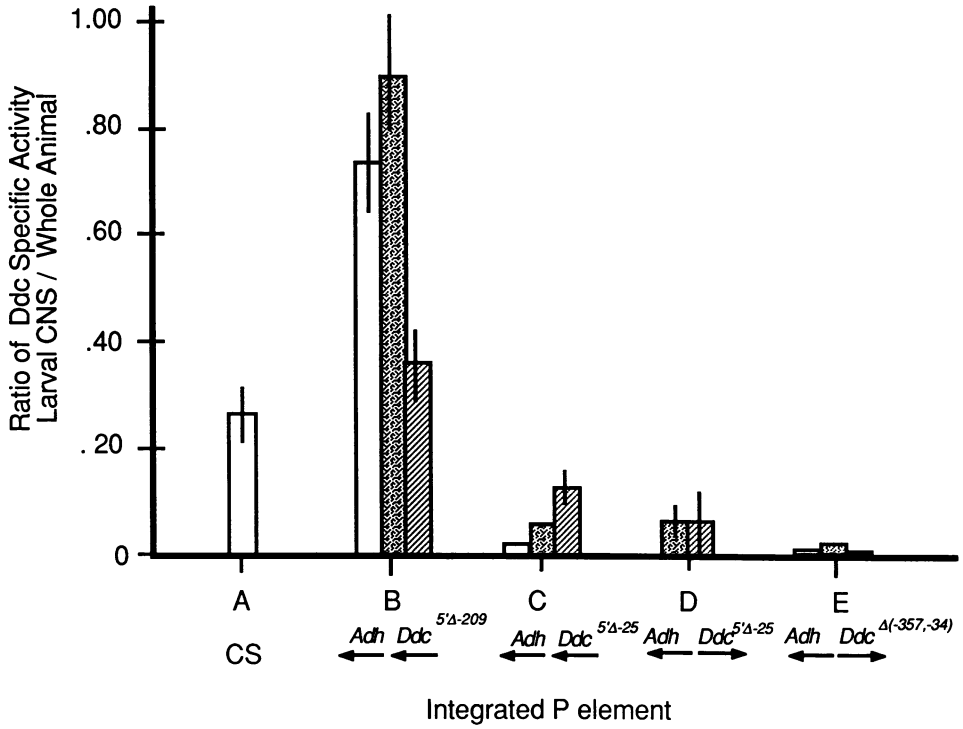

FIG. 5. Central nervous system (CNS) *Ddc* activity in wild-type and transformed lines. Relative levels of *Ddc* specific activity in each strain are shown, taking the ratio of specific activity in the larval central nervous system versus the whole white prepupae to normalize for different levels of expression. Each bar shows the value obtained from multiple measurements of a given strain. Errors shown are the standard errors of the mean. Sketches of the *Ddc* vectors are shown at the bottom of the figure, and the arrows indicate the 5'→3' orientation of *Ddc* and *Adh*. (A) Canton-S (CS) wild-type strain. (B) $Ddc^{5'\Delta-209}$ strains: strain 24-22 (open); strain 24-44 (stippled); strain 24-54 (striped). (C) $Ddc^{5'\Delta-25}$ strains with the *Ddc* gene oriented with the 5' end distal to *Adh*: strain 47-7 (open); strain 47-8 (stippled); strain 47-17 (striped). (D) $Ddc^{5'\Delta-25}$ strains with the *Ddc* gene oriented with the 5' end proximal to *Adh*: strain 27-36 (stippled); strain 27-48 (striped). (E) $Ddc^{\Delta(-357,-34)}$ strains with *Ddc* gene oriented with the 5' end proximal to *Adh*: strain 35-4 (open) strain 35-11 (stippled) strain 35-13 (solid).

type, even though they were normally pigmented. Paradoxically, abnormally pigmented strains could have had a nearly normal fraction of activity in the integument. This paradox is most apparent in the 27 series ($Ddc^{5'\Delta-25}$) strains. Both strains assayed had nearly wild-type levels of *Ddc* activity at pupariation (Fig. 2C), and many individual flies had a normal fraction of activity in the integument, yet all were abnormally pigmented. The premature induction of *Ddc* activity observed in these strains cannot be the explanation for the abnormal phenotype, since the $Ddc^{\Delta(-357,-34)}$ strains also showed this abnormal developmental profile and an even lower fraction of integument-restricted *Ddc* enzyme activity, yet were normally pigmented. We conclude that the loss of sequences between −209 and −24 leads to a subtle defect in hypodermal expression that is beyond the detection of our crude tissue fractionation methodology. Note, however, that in none of the strains was the fraction of *Ddc* enzyme in the integument at the 30% level, which would be expected for random tissue distribution.

A minor fraction of *Ddc* enzyme activity is found in the central nervous system, where it appears to be necessary for normal learning behavior (37). We assayed for expression of *Ddc* in the central nervous system by measuring *Ddc* enzyme activity in hand-dissected preparations of brain lobes and ventral ganglia from late third instar larvae. Relative levels of *Ddc* specific activity in wild-type controls and in strains containing altered *Ddc* genes are shown in Fig. 5. The data are expressed as the ratio of specific activity in the central nervous system relative to maximal levels in whole puparia, to normalize for different levels of overall expression. Normalizing to levels in whole puparia is justified because *Ddc* expression in the central nervous system of wild-type strains does not change between early third instar and pupariation (14). Strains containing $Ddc^{5'\Delta-209}$ genes are expressed at relative levels near or above those of the wild type in the central nervous system. Since the whole puparia expression of these strains was near normal (Fig. 2A), this relative central nervous system overexpression was also an absolute overexpression. Whether this overexpression was caused by the loss of upstream sequences that normally serve to down-modulate expression or to position effects from newly adjacent chromosomal sequences cannot be determined from these data. However, all strains containing *Ddc* genes lacking sequences between −209 and −34 had significantly reduced expression in the central nervous system. This effect was seen for genes integrated in both orientations within the P-element vector and for genes expressed at widely varying levels at pupariation. Thus, sequences between −209 and −34 are necessary for normal expression of *Ddc* in the central nervous system.

DISCUSSION

All regulatory elements essential for normal tissue-specific and temporally regulated expression of the *D. melanogaster* *Ddc* gene are located very close to or within the gene. Genes retaining only 208 bp of 5'-flanking sequences and 1 kb of 3'-flanking sequences are expressed normally. More se-

verely deleted *Ddc* genes, retaining 24 to 33 bp of 5'-flanking sequences, are expressed abnormally, indicating that important regulatory sequences are contained in the upstream region between −34 and −209.

Defects in tissue-specific expression associated with deletion of sequences between −209 and −24. The loss of expression in the larval central nervous system is the most clear-cut aberration in tissue-specific expression observed with genes lacking sequences between −24 and −209. *Ddc* expression in the larval central nervous system is reduced to below 20% of expected levels in all strains containing these genes (Fig. 5). We can rule out one trivial explanation for this effect, namely, that the central nervous system transcript uses an upstream RNA start point relative to the predominant *Ddc* transcript, since primer extension assays show that the larval central nervous system transcript has the same RNA start point (Morgan et al., submitted). It is not clear whether expression in the adult central nervous system is similarly reduced. Given the low level of expression of the gene in adults of these strains, a selective loss of expression in the adult central nervous system would be difficult to detect.

In contrast to the striking loss of *Ddc* expression in the larval central nervous system observed with the severely deleted *Ddc* genes, expression in the larval hypoderm is more subtly affected. The strains containing $Ddc^{5'\Delta\text{-}25}$ genes with the 5' end proximal to *Adh* express *Ddc* activity at 80% of normal levels at pupariation, with nearly normal hypodermal tissue specificity, yet develop abnormal pigmentation following pupation. Paradoxically, this pigmentation abnormality resembles that seen in mutant strains expressing greatly reduced levels of *Ddc* enzyme activity. Pupal cases formed by *Ddc* mutant strains expressing *Ddc* enzyme activity at several percent of normal levels are a yellow-green color, with blackening at both anterior and posterior ends (33). This phenotype is clearly distinguishable from the normal red-brown pupal coloration of wild-type strains, although mutant strains expressing 20% of normal *Ddc* levels are indistinguishable from the wild type (unpublished data). The phenotype of the $Ddc^{5'\Delta\text{-}25}$ strains is moderate, as would be expected of a mutant strain expressing about 10% the normal levels.

The mutant phenotype of a $Ddc^{5'\Delta\text{-}25}$ strain is more severe if the strain is hemizygous and the phenotype is recessive in the presence of a wild-type *Ddc* gene (unpublished data). Both of these findings would be expected if the lack of normal pupal pigmentation is caused by a loss of expression, perhaps in some specialized cell type in the hypoderm. Whatever the nature of the anomalous hypodermal expression, it was not detectable by our crude biochemical assay of hypodermal tissue specificity (Table 1). The hypoderm-specific expression of these strains will be better studied with immunologic probes to examine cell-specific expression.

A role for the *Ddc* TATA box? The *Ddc* TATA box sequence TTTAAAA does not appear to be necessary for specifying the RNA start point. The $Ddc^{5'\Delta\text{-}25}$ genes lack this sequence or any A+T-rich sequence in a position near the normal TATA box site, yet these genes use the normal start point for the majority of transcripts. Only a small fraction of transcripts use an upstream start point in a vector polylinker sequence which is identical to the sequence at the normal start point. These results imply that the sequence at the RNA start point is more important for start point determination than the TATA box sequence. Results of a recent study of several yeast genes have reached a similar conclusion (7). Similar results were obtained with the rabbit β-globin gene following cell line transfection, although an alteration in the β-globin TATA box leads to a severe reduction in expression (9).

A mutant gene which retains the normal TATA box sequence, the $Ddc^{\Delta(-357,-34)}$ gene, shows a grossly similar developmental profile to the $Ddc^{5'\Delta\text{-}25}$ genes in the same orientation within the P-element vector. This deletion was constructed so as to retain a normal TATA box sequence, yet sequences immediately upstream were altered. Curiously, the $Ddc^{\Delta(-357,-34)}$ strains showed a wild-type pupal pigmentation phenotype. Thus, the *Ddc* TATA box may play some role in specifying proper hypoderm-specific expression. An alternative explanation, which cannot be ruled out by our data, is that a sequence element near −350 replaces a deleted element to enable normal pigmentation when it is closely apposed to *Ddc*. However, analysis of the DNA sequences in the −350 region failed to show any significant homologies with sequences between −33 and −209 (data not shown).

Expression of the severely deleted genes: position effects and implications for localization of *Ddc* regulatory elements. The $Ddc^{5'\Delta\text{-}25}$ *Ddc* genes are expressed at roughly threefold higher levels when oriented with the 5' end proximal to the adjacent *Adh* gene than when oppositely oriented. In the proximal orientation the RNA start points are separated by about 1.4 kb. This activation cannot be due to transcriptional readthrough from *Adh*, since the majority of mutant transcripts initiate at the proper site. The larval overexpression of *Ddc* must be due to transcriptional stimulation of the remaining *Ddc* promoter elements by some *Adh* sequence.

The enhanced larval expression of the $Ddc^{5'\Delta\text{-}25}$ mutants resembles the pattern of expression of the wild-type *Adh* gene, a truncated version of which is adjacent to *Ddc* in these constructs. *Adh* is abundantly expressed in both larvae and adults but at stages and in tissues that are distinct from those observed for *Ddc*. Although the cloned copy of *Adh* contained within these constructs is underexpressed fivefold during larval development (J. Posakony, personal communication), it may still contain an element(s) capable of stimulating *Ddc* expression throughout larval life. However, the expression of these *Ddc* mutants is not consistent with simple regulation of the *Ddc*-coding region by *Adh*-controlling elements. First, *Ddc* expression from these genes begins at about 20 h into embryonic development, which is later than *Adh* expression, which begins at 9 h (32), and later than normal *Ddc* expression, which commences at about 16 h into embryonic development. Second, the tissue specificity resembles normal *Ddc* expression. We assayed these strains for *Ddc* expression in larval fat body, a tissue rich in *Adh* enzyme activity, and found negligible *Ddc* enzyme activity (data not shown). Third, the temporal regulation at pupariation does not fit simply with an effect of *Adh*, in that *Ddc* enzyme activity does not decrease until after pupariation, which is significantly after the decay of the majority of *Adh* transcripts (32). These data suggest that an *Adh* element interacts with remaining *Ddc* regulatory elements, resulting in a novel pattern of expression with features of both wild-type genes.

Reversing the orientation of the $Ddc^{5'\Delta\text{-}25}$ gene such that the 5' end is distal from *Adh* lowers larval *Ddc* expression roughly threefold, but even in this configuration significant expression remains. In each of the three strains examined, larval *Ddc* enzyme levels were near normal, although there was no significant induction at pupariation. Levels dropped following pupariation and then were induced at adult eclosion, although to a smaller extent than in wild-type strains. It

is unclear whether the decrease in postpupariation *Ddc* enzyme activity represents gene-specific regulation or the histolysis of the larval cells in which expression is occurring. The most developmentally significant aspect of this profile is the induction at adult eclosion. An induction at adult eclosion was also seen with $Ddc^{\Delta(-357,-34)}$ genes oriented in the 5′ direction proximal to *Adh*, although this induction was not observed in any of the $Ddc^{5'\Delta-25}$ genes in the same orientation. We surmise that this lack of induction at eclosion in this latter strain is caused by an incompatibility of sequences immediately flanking the $Ddc^{5'\Delta-25}$ gene with regulatory elements responsible for this induction.

From the analysis presented above, it appears likely that an element from within the *Adh* gene can positively influence expression from the severely deleted $Ddc^{5'\Delta-25}$ genes when the 5′ ends of the two genes are oriented proximally. However, it also appears unlikely that all of the residual expression can be accounted for by this interaction or that this element could account for all of the expression of the $Ddc^{5'\Delta-25}$ genes when they are oriented oppositely. We therefore expect that the severely deleted *Ddc* genes, some of which retain as little as 24 bp of upstream-flanking sequences, retain some components of their normal regulatory and expression apparatuses. Although regulatory elements within eucaryotic genes transcribed by RNA polymerase II are not common, several examples have been described. The best defined examples are elements within the mouse immunoglobulin light and heavy chain genes, consisting of both an enhancer (1, 4, 10, 26–28) and other intragenic sequences (12). Other examples include intragenic sequences within the chicken *tk* gene that regulate expression in differentiating muscle cells (22) and sequences within the mouse β-globin gene that determine the responses to enhancers and to dimethyl sufoxide stimulation (6, 38, 41). The mechanisms of action of these intragenic regulatory sequences have yet to be defined.

Our results indicate that the tissue and temporal expression of *Ddc* has complex *cis*-acting sequence requirements that certainly include elements located in the immediate 5′-flanking sequences and that probably include elements very close to or within the gene. More subtle sequence alterations will be necessary to finely localize the sequences responsible for the various phases of this regulation.

ACKNOWLEDGMENTS

We thank Gardiner Morse and Carol Ann McCormick for excellent technical assistance during many phases of this work. We thank Jim Posakony for providing the *Adh* gene primer, the pSXΔ1 P-element vector, and *Adh* gene balancer strains; Alex Nussbaum for synthesis of the *Ddc* gene primer; and the other members of our laboratory for a critical reading of this manuscript.

This work was supported by Public Health Service grant GM27318 from the National Institutes of Health, grant 1-908 from the March of Dimes National Foundation, and fellowship PF2073 from the American Cancer Society.

ADDENDUM IN PROOF

Recent experiments (S. B. Scholnick, S. J. Bray, B. A. Morgan, C. A. McCormick, and J. Hirsh, Science, in press) have localized an element that is selectively required for expression of *Ddc* in the CNS to lie between −83 and −59, within the region defined in this manuscript. However, an immunochemical examination of *Ddc* expression demonstrates that at least one additional element upstream of −208 is required for normal cell-specific expression within the CNS (C. Beall and J. Hirsh, in preparation).

LITERATURE CITED

1. **Banerji, J., L. Olson, and W. Schaffner.** 1983. A lymphocyte-specific cellular enhancer is located downstream of the joining region in immunoglobulin heavy chain genes. Cell **33**:729–740.
2. **Beall, C., and J. Hirsh.** 1984. High levels of intron containing RNAs are associated with expression of the *Drosophila* dopa decarboxylase gene. Mol. Cell. Biol. **4**:1669–1674.
3. **Benyajati, C., N. Spoerel, H. Haymerle, and M. Ashburner.** 1983. The messenger RNA for alcohol dehydrogenase in *Drosophila melanogaster* differs in its 5′ end in different developmental stages. Cell **33**:125–133.
4. **Bergman, Y., D. Rice, R. Grosschedl, and D. Baltimore.** 1984. Two regulatory elements for immunoglobulin kappa light chain gene expression. Proc. Natl. Acad. Sci. USA **81**:7041–7045.
5. **Brunet, P. C. J.** 1980. The metabolism of the aromatic amino acids concerned in the cross-linking of insect cuticle. Insect Biochem. **10**:467–500.
6. **Charnay, P., R. Treisman, P. Mellon, M. Chao, R. Axel, and T. Maniatis.** 1984. Differences in human α- and β-globin gene expression in mouse erythroleukemia cells: the role of intragenic sequences. Cell **38**:251–263.
7. **Chen, W., and K. Struhl.** 1985. Yeast mRNA initiation sites are determined primarily by specific sequences, not by the distance from the TATA element. EMBO J. **4**:3273–3280.
8. **Church, G. M., and W. Gilbert.** 1984. Genomic sequencing. Proc. Natl. Acad. Sci. USA **81**:1991–1995.
9. **Dierks, P., A. van Ooyen, M. D. Cochran, C. Dobkin, J. Reiser, and C. Weissman.** 1983. Three regions upstream from the cap site are required for efficient and accurate transcription of the rabbit β-globin gene in mouse 3T6 cells. Cell **32**:695–706.
10. **Gillies, S. D., V. Folsom, and S. Tonegawa.** 1984. Cell type specific enhancer element associated with a mouse MHC gene, E_β. Nature (London) **310**:594–597.
11. **Goldberg, D., J. Posakony, and T. Maniatis.** 1983. Correct developmental expression of a cloned alcohol dehydrogenase gene transduced into the *Drosophila* germ line. Cell **34**:50–73.
12. **Grosschedl, R., and D. Baltimore.** 1985. Cell-type specificity of immunoglobulin gene expression if regulated by at least three DNA sequence elements. Cell **41**:885–887.
13. **Henikoff, S.** 1983. Cloning exons for mapping transcription: characterization of the *Drosophila melanogaster* alcohol dehydrogenase gene. Nucleic Acids Res. **11**:4735–4752.
14. **Hirsh, J.** 1986. The *Drosophila melanogaster* dopa decarboxylase gene: progress in understanding the in vivo regulation of a higher eukaryotic gene, p. 103–116. *In* L. Bogorad (ed.), Molecular developmental biology. Alan R. Liss, Inc., New York.
15. **Hirsh, J., and N. Davidson.** 1981. Isolation and characterization of the dopa decarboxylase gene of *Drosophila melanogaster*. Mol. Cell. Biol. **1**:475–485.

15a. **Hultmark, D., R. Klemenz, and W. J. Gehring.** 1986. Translational and transcriptional control elements in the untranslated leader of the heat-shock gene hsp22. Cell **44**:429–438.

16. **Karess, R. E., and G. M. Rubin.** 1984. Analysis of P transposable element functions in Drosophila. Cell **38**:135–146.
17. **Kraminsky, G. P., W. C. Clark, M. A. Estelle, R. D. Gietz, B. A. Sage, J. D. O'Connor, and R. B. Hodgetts.** 1980. Induction of translatable mRNA for dopa decarboxylase in *Drosophila*: an early response to ecdysterone. Proc. Natl. Acad. Sci. USA. **77**:4175–4179.
18. **Livingstone, M. S., and B. L. Tempel.** 1983. Genetic dissection of monamine neurotransmitter synthesis in *Drosophila*. Nature (London) **303**:67–70.
19. **Maniatis, T., E. F. Fritsch, and J. Sambrook.** 1982. Molecular cloning, a laboratory manual. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
20. **Marsh, J. L., P. D. L. Gibbs, and P. M. Timmons.** 1985. Developmental control of transduced dopa decarboxylase genes in *D. melanogaster*. Mol. Gen. Genet. **198**:393–403.
21. **Marsh, J. L., and T. R. F. Wright.** 1980. Developmental relationship between dopa decarboxylase, dopa acetyl transferase, and ecdysone in *Drosophila*. Dev. Biol. **80**:379–387.

22. **Merrill, B. F., S. D. Hauschka, and S. L. McKnight.** 1984. TK enzyme expression in differentiating muscle cells is regulated through an internal segment of the cellular *tk* gene. Mol. Cell. Biol. **4**:1777–1784.
23. **O'Hare, K., and G. M. Rubin.** 1983. Structures of P transposable elements and their sites of insertions and excision in the *Drosophila melanogaster* genome. Cell **34**:25–35.
24. **Okuno, S., and H. Fujisawa.** 1983. Accurate assay of dopa decarboxylase by preventing nonenzymatic decarboxylation of dopa. Anal. Biochem. **129**:412–415.
25. **O'Tousa, J. E., W. Baehr, R. L. Martin, J. Hirsh, W. L. Pak, and M. L. Applebury.** 1985. The *Drosophila ninaE* gene encodes an opsin. Cell **4**:839–850.
26. **Picard, D., and W. Schaffner.** 1984. A lymphocyte specific enhancer in the mouse immunoglobulin kappa gene. Nature (London) **307**:80–82.
27. **Potter, H., L. Weir, and P. Leder.** 1984. Enhancer-dependent expression of human χ immunoglobulin genes introduced into mouse pre-B lymphocytes by electroporation. Proc. Natl. Acad. Sci. USA **81**:7161–7165.
28. **Queen, C., and D. Baltimore.** 1983. Immunoglobulin gene transcription is activated by downstream sequence elements. Cell **33**:741–748.
29. **Richards, G., A. Cassab, M. Bourouis, B. Jarry, and C. Dissous.** 1983. The normal developmental regulation of a cloned sgs3 'glue' gene chromosomally integrated in *Drosophila melanogaster* by P element transformation. EMBO J. **2**:2137–2142.
30. **Rubin, G. M., and A. C. Spradling.** 1982. Genetic transformation of *Drosophila* with transposable element vectors. Science **218**:348–353.
31. **Sanger, F., S. Nicklen, and A. R. Coulsen.** 1977. DNA sequencing with chain terminating inhibitors. Proc. Natl. Acad. Sci. USA **74**:5463–5467.
32. **Savakis, C., M. Ashburner, and J. H. Willis.** 1986. The expression of the gene coding for alcohol dehydrogenase during the development of *Drosophila* melanogaster. Dev. Biol. **114**:194–207.
33. **Scholnick, S., B. A. Morgan, and J. Hirsh.** 1983. The cloned dopa decarboxylase gene is developmentally regulated when reintegrated into the *Drosophila* germline. Cell **34**:37–45.
34. **Seed, B.** 1983. Purification of genomic sequences from bacteriophage libraries by recombination and selection in vivo. Nucleic Acids Res. **11**:2427–2445.
35. **Spradling, A. C., and G. M. Rubin.** 1982. Transposition of cloned P elements into *Drosophila* germline chromosomes. Science **218**:341–347.
36. **Spradling, A. C., and G. M. Rubin.** 1983. The effect of chromosomal position on the expression of the *Drosophila* xanthine dehydrogenase gene. Cell **34**:47–57.
37. **Tempel, B. L., M. S. Livingston, and W. G. Quinn.** 1984. Mutations in the dopa decarboxylase gene affect learning in *Drosophila*. Proc. Natl. Acad. Sci. USA **81**:3577–3581.
38. **Treisman, R., M. R. Green, and T. Maniatis.** 1983. Cis- and trans-activation of globin gene transcription in transient assays. Proc. Natl. Acad. Sci. USA **80**:7428–7432.
38a. **Vallés, A. M., and K. White.** 1986. Development of serotonin-containing neurons in *Drosophila* mutants unable to synthesize serotonin. J. Neurosci. **6**:1482–1491.
39. **Vieira, J., and J. Messing.** 1982. The pUC plasmids, an M13mp7-derived system for insertion mutagenesis and sequencing with synthetic universal primers. Gene **19**:259–268.
40. **Vigue, C., and W. Sofer.** 1976. *Adh^{n5}*: a temperature-sensitive mutant at the *Adh* locus in *Drosophila*. Biochem. Genet. **11**:387–396.
41. **Wright, S., A. Rosenthal, R. Flavell, and R. Grosveld.** 1984. DNA sequences required for regulated expression of β-globin genes in murine erythroleukemia cells. Cell **38**:265–273.
42. **Wright, T. R. F., B. C. Black, C. P. Bishop, J. L. Marsh, E. S. Pentz, R. Steward, and E. Y. Wright.** 1982. The genetics of dopa decarboxylase in *Drosophila melanogaster*. V. *Ddc* and *l(2)amd* alleles: isolation, characterization and intragenic complementation. Mol. Gen. Genet. **188**:18–26.