Analysis of the Promoter of the *ninaE* Opsin Gene in *Drosophila melanogaster*

Drzislav Mismer and Gerald M. Rubin

Department of Biochemistry, University of California, Berkeley, California 94720

Manuscript received April 17, 1987
Accepted May 16, 1987

ABSTRACT

We have analyzed the *cis*-acting regulatory sequences of the *ninaE* gene. This gene encodes the major *Drosophila melanogaster* opsin, the protein component of the primary chromophore of photoreceptor cells R1–R6 of the adult eye. DNA fragments containing the start point of transcription of the *ninaE* gene were fused to either the *Escherichia coli* chloramphenicol acetyltransferase or *lacZ* (β-galactosidase) gene and introduced into the Drosophila germline by *P*-element-mediated transformation. Expression of the *E. coli* genes was then used to assay the ability of various sequences from the *ninaE* gene to confer the *ninaE* pattern of expression. Fragments containing between 2.8 kb and 215 bp of the sequences upstream of the start of transcription plus the first 67 bp of the untranslated leader were able to direct nearly wild-type expression. We have identified three separable control regions in the *ninaE* promoter. The first, which has the properties of an enhancer element, is located between nucleotides −501 and −219. The removal of this sequence had little effect on promoter function; this sequence appears to be redundant. However, it appears to be able to substitute for the second control region which is located between nucleotides −215 and −162, and which also affects the level of output from this promoter. Removal of these two control regions resulted in a 30-fold decrease in expression; however tissue specificity was not affected. The third control region, located downstream from nucleotide −120, appears to be absolutely necessary for promoter function in the absence of the first two regulatory sequences. Examination of larvae containing fusion genes expressing β-galactosidase suggests that the *ninaE* gene is also expressed in a subset of cells in the larval photoreceptor organ.

ADULT Drosophila, like many other insects, have compound eyes consisting of a number of identical units called ommatidia (for review of insect ommatidia, see CHAPMAN 1982). In the Drosophila eye, each of the approximately 800 ommatidia contains 20 precisely arranged cells, eight of which are photoreceptor cells. The six outer photoreceptor cells, R1–R6, encircle the two inner cells, R7 and R8. The Drosophila compound eye develops from cells in the eye-antennal imaginal disc. At the start of the third larval instar, cells in the eye disc are undifferentiated. Differentiation of the disc epithelium starts at the posterior edge of the disc; associated with the differentiation process a morphological depression, the morphogenetic furrow, moves across the disc in a posterior-anterior direction over a 3-day period (READY, HANSON and BENZER 1976). Posterior to the furrow cells assemble into clusters that will give rise to adult ommatidia. During the assembly process, undifferentiated cells apparently become determined by interpreting positional cues in their local environment (READY, HANSON and BENZER 1976; TOMLINSON 1985; TOMLINSON and READY 1987). No determinative lineage relationships have been detected between cells in the developing eye (READY, HANSON and BENZER 1976; LAWRENCE and GREEN 1979). The end result is a highly ordered retina made up of several different cell types that differ not only in the positions that they occupy within the ommatidium, but also in their patterns of gene expression.

Each of the four different opsin genes that have been identified in the *Drosophila melanogaster* genome is expressed in only a subpopulation of photoreceptor cells (ZUKER, COWMAN and RUBIN 1985; O'TOUSA *et al.* 1985; COWMAN, ZUKER and RUBIN 1986; ZUKER *et al.* 1987; MONTELL *et al.* 1987; FRYXELL and MEYEROWITZ 1987). Photoreceptor cells R1–R6 appear to contain the same rhodopsin (OSTROY, WILSON and PAK 1974; HARRIS, STARK and WALKER 1976) which is encoded by the *ninaE* locus (ZUKER, COWMAN and RUBIN 1985; O'TOUSA *et al.* 1985). This gene is also referred to as Rh1, being the first rhodopsin gene isolated. The predicted product of the *ninaE* locus is 373 amino acids long and has seven hydrophobic domains (ZUKER, COWMAN and RUBIN 1985; O'TOUSA *et al.* 1985), a characteristic shared by all known rhodopsins (OVCHINNIKOV 1982; NATHANS and HOG-

The sequence data presented in this article have been submitted to the EMBL/GenBank Data Libraries under the accession number Y00601.

NESS 1983, 1984; NATHANS, THOMAS and HOGNESS 1986).

We have begun an analysis of the promoter sequences of *ninaE* locus using *in vitro* mutagenesis coupled with *P*-element-mediated germline transformation. We have made fusions between segments of the *ninaE* gene and two bacterial indicator genes (chloramphenicol acetyltransferase and *lacZ*, which encodes β-galactosidase). We show that as little as 120 bp upstream of the transcription initiation site plus 67 bp of 5′-untranslated leader are sufficient to direct expression to retinula cells R1–R6, albeit at reduced levels. Sequences located further upstream influence the level of expression from this promoter.

MATERIALS AND METHODS

P-Element-mediated transformation and modification of preexisting Drosophila transformation vectors: *P*-element-mediated transformation was carried out essentially as previously described (SPRADLING and RUBIN 1982; RUBIN and SPRADLING 1982). Drosophila embryos were dechorionated either mechanically or by submersion in 1:1 (v/v) commercial bleach:dH_2O for 2–3 min. For injection, DNA was dissolved in 5 mM KCl, 0.1 mM NaH_2PO_4 (pH 6.8) and mixed with the helper plasmid pπ25.7wc (KARESS and RUBIN 1984) to give final concentrations of approximately 2 mg/ml and 0.5 mg/ml, respectively.

Transformation vectors, containing either *rosy* (RUBIN and SPRADLING 1983) or G418 resistance markers (STELLER and PIRROTTA 1985) were used. Selection for G418 resistant flies was carried out using commercially available dehydrated fly food (Carolina Biologicals) made according to the supplier's instructions but containing G418 sulfate at 350–500 μg/ml. G418 sulfate was purchased from Gibco. Drosophila strains injected were either *ry506* when derivatives of the *rosy*-containing Carnegie vectors were used or Canton-S when pUChsneo derivatives were used. These stocks were reared on cornmeal-agar food.

To facilitate manipulations of the *ninaE* promoter, a number of modifications were introduced into preexisting transformation vectors. These vectors are diagrammed in Figure 1. In order to increase their unique cloning sites, Carnegie 30 and pUChsneo were modified by replacing part of their polylinkers by a portion of the M13 mp19 polylinker. pDM23 is a Carnegie 30 derivative in which *Xba*I, *Kpn*I and *Sal*I are unique cloning sites. pDM24 is a pUChsneo derivative containing unique *Sal*I, *Xba*I, *Bam*HI, *Sma*I, *Kpn*I, *Sac*I and *Eco*RI sites. By inserting an 8-bp oligonucleotide encoding the cleavage site of *Not*I (GCGGCCGC) into the *Hpa*I site of Carnegie 20 and into the *Sma*I site of pUChsneo, the vectors pDM30 and pDM25 were obtained.

The plasmid pC4*CAT* (gift of C. THUMMEL and D. HOGNESS) contains a chloramphenicol acetyltransferase transcription fusion module within the polylinker of Carnegie 4. In order to place a *Kpn*I site at the 5′-end of this module, pC4*CAT* was digested with *Sma*I and an 8-bp *Kpn*I linker sequence (CGGTACCG) was inserted. The *CAT* module was then excised from the resulting plasmid by digestion with *Kpn*I and *Sac*I and cloned into pDM24 to generate the plasmid pDM26. The plasmid pDM79 was constructed by transferring the *lacZ* (β-galactosidase) transcriptional fusion module from pβ-galAUG (gift of C. THUMMEL and D. HOGNESS) as an *Eco*RI fragment into pDM24. The translation start of this modified *lacZ* gene was taken from Drosophila alcohol dehydrogenase gene. The Adh segment used was the *Sau*3AI fragment extending from +34 to +160, which contains the translation initiator codon AUG (+1 represents the start of transcription from the proximal Adh

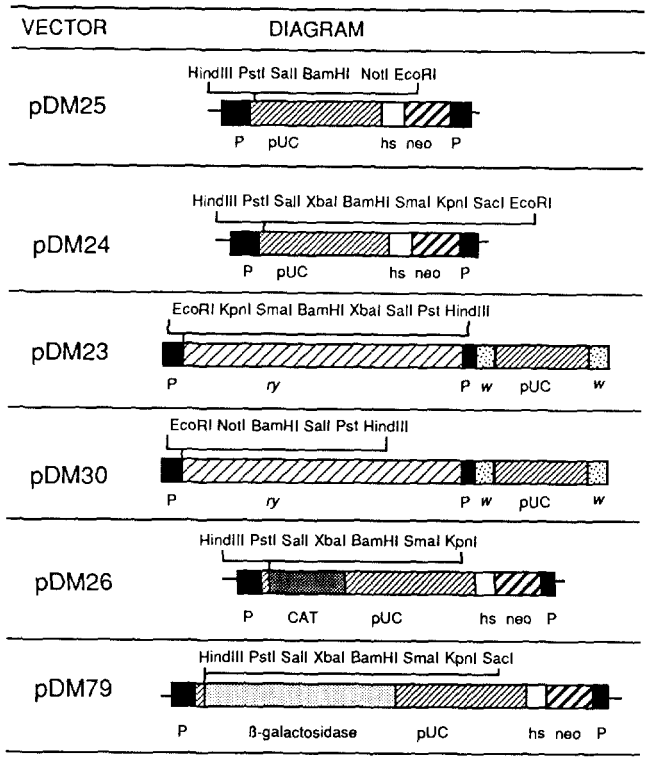


FIGURE 1.—Modification of vectors for *P*-element-mediated germline transformation. Both Carnegie and pUChsneo vectors were modified by introducing novel unique restriction enzyme sites into their polylinkers as described in MATERIALS AND METHODS. Only the polylinker sequences were modified; the remainder of the vector is the same as the previously published parent plasmids (RUBIN and SPRADLING 1983; STELLER and PIRROTTA 1985). In the Carnegie vector system, transformants are recognized by the rescue of *ry* eye color. The utility of these vectors is limited by the scarcity of unique restriction enzyme sites suitable for the cloning of DNA fragments to be transformed. Two vectors, pDM23 and pDM30, have additional unique restriction enzyme sites within their polylinkers which were introduced as described in MATERIALS AND METHODS. Plasmid pDM23, a Carnegie 30 derivative, contains unique *Xba*I, *Sal*I and *Kpn*I sites in its polylinker sequence. Plasmid pDM30, a Carnegie 20 derivative, contains an unique *Not*I site within its polylinker. This enzyme possesses an octanucleotide recognition sequence that rarely occurs in the *D. melanogaster* DNA. pDM30 is used in conjunction with plasmid pHSS7 (SEIFERT *et al.* 1986) which has a polylinker bordered by *Not*I sites. The desired fragment is first cloned into the polylinker of the pHSS7 vector, removed from pHSS7 by *Not*I digestion and then recloned into either pDM30 or pDM25, a pUChsneo derivative with a *Not*I site. Other unique polylinker sites in pDM25 are *Eco*RI, *Bam*HI and *Sal*I. The plasmid pDM24 is a pUChsneo derivative where the unique polylinker sites are: *Eco*RI, *Kpn*I, *Sac*I, *Sma*I, *Sal*I, *Bam*HI and *Xba*I. The *E. coli* indicator genes *CAT* and β-galactosidase were cloned into pDM24 to create plasmids suitable for gene fusion experiments. Plasmid pDM26 contains a *CAT* module as a *Kpn*I/*Sac*I fragment within its polylinker. A promoter can be inserted into the *Sal*I, *Xba*I, *Bam*HI, *Sma*I or *Kpn*I sites in order to effect the formation of a transcriptional fusion between the *CAT* gene and the promoter. Although *Sac*I is also a unique site, it cannot be used since it is located 3′ to the *CAT* gene. The plasmid pDM79 permits the insertion of a promoter into a *Sal*I, *Bam*HI or *Kpn*I site in order to form a transcriptional fusion between the promoter and *lacZ* gene, which has been cloned into pDM24 as an *Eco*RI fragment.

promoter). No Adh control regions have been reported in these sequences (GOLDBERG, POSAKONY and MANIATIS 1983; POSAKONY, FISCHER and MANIATIS 1985; HEBERLEIN, ENGLAND and TJIAN 1985). In pDM79, unique *Sal*I, *Bam*HI and *Kpn*I sites are located in the 5′-untranslated leader of this Adh/*lacZ* fusion gene.

DNA sequence analysis of *ninaE* promoter: Nucleotide sequence analysis was carried out by the dideoxy chain termination method (SANGER, NICKLEN and COULSON 1977). Randomly sheared fragments, 300–600 bp long, of the *ninaE* promoter were cloned into *Sma*I-cut M13 mp18 and sequenced. The host strain used for the growth of M13 was *Escherichia coli* TG1. Sequencing reactions utilized α-[^{35}S]dATP as the radioactively labeled nucleotide and were carried out essentially as described by BANKIER and BARRELL (1983). The sequence of the *ninaE* promoter was determined on both strands in the region from −833 bp to +67 bp. The DNA sequence from −260 to −1 was also confirmed by the chemical degradation method (MAXAM and GILBERT 1977). The single-stranded DNAs used as templates for dideoxy chain termination sequencing were obtained from either M13 of pEMBL vectors according to published procedures (DENTE, CESARENI and CORTESE 1983).

DNA manipulations: DNA manipulations, such as restriction endonuclease digests and nick-translation labeling of DNA, were generally performed as previously described (MANIATIS, FRISCH and SAMBROOK 1982). Restriction enzymes were purchased from New England Biolabs and Boehringer Mannheim Biochemicals. DNA polymerase large fragment (Klenow) was purchased from Bethesda Research Laboratories. Bal31 degradation of DNA was carried out in 0.6 M NaCl, 12 mM $MgCl_2$, 12 mM $CaCl_2$, 1 mM EDTA and 20 mM Tris (pH 8.0). For each 3.5 μg of linear DNA, about 1 unit of Bal31 was used at room temperature which gave a digestion rate of approximately 50 bp/min. In order to insert the *Xho*I linkers (CCTCGAGG) at random into the *ninaE* promoter, plasmids containing the promoter were linearized by the DNaseI/Mn^{2+} procedure (MELGAR and GOLDTHWAITE 1968).

Construction of *ninaE* promoter-*CAT* fusions: Various *ninaE* promoter fragments were fused with the *E. coli CAT* gene in order to facilitate the analysis of the regulatory sequences of this rhodopsin gene. Diagrams of all of the fusions described are shown in Figure 2.

pP[*hs-neo*; Rh1(~−2800/+67)-*CAT*]: A plasmid subclone of the genomic *Hin*dIII/*Bam*HI fragment of the *ninaE* gene (see Figure 3) that contains the 5′-untranslated leader and flanking sequences was made linear by digestion with *Bam*HI. Following limited Bal31 degradation and *Hin*dIII digestion, a population of partially deleted *ninaE* promoter fragments was isolated from a preparative agarose gel and subcloned into the vector pEMBL19+ which has been cut with *Hin*dIII and *Sma*I. pEMBL19+ is the same as pEMBL9+ (DENTE, CESARENI and CORTESE 1983) except that it contains the polylinker of M13 mp19 instead of M13 mp9 (gift of H. ROIHA). Single-stranded DNA was then obtained from a number of subclones and the end-points of Bal31 deletion were determined by DNA sequencing. One of the clones contained an end-point at position +67 which is between the start of transcription and the start of translation of the *ninaE* gene. This *ninaE* promoter fragment was then cloned as a *Hin*dIII (site filled in using DNA polymerase I large fragment)/*Kpn*I (this *Kpn*I site at the 3′ end of the promoter fragment, originates from the polylinker of pEMBL19+) fragment into *Sma*I- and *Kpn*I-digested pDM26.

pP[*hs-neo*; Rh1(−833/+67)-*CAT*]: This plasmid was constructed by cleaving pP[*hs-neo*; Rh1(~−2800/+67)-*CAT*] with *Bam*HI (the *Bam*HI site is in the polylinker of pDM26) and *Bgl*II (position −833 of the Rh1 promoter) and religating.

pP[*ry*; Rh1(−448/+67)-*CAT*]: The Rh1 promoter fragment was subcloned from pEMBL19+ as a *Hin*dIII/*Eco*RI fragment (the *Eco*RI site is in the pEMBL19+ polylinker) into pHSS7. The plasmid pHSS7 (SEIFERT *et al.* 1986) has a polylinker bordered by *Not*I sites. The resulting plasmid was then subjected to DNaseI cleavage in the presence of Mn^{2+}. This results in a population of randomly linearized molecules into which *Xho*I linkers were then inserted to create plasmid molecules with unique *Xho*I sites. Since the vector pHSS7 contains virtually no nonessential sequences (SEIFERT *et al.* 1986), only one out of 51 independent linker insertions was outside the *ninaE* promoter fragment. The positions of *Xho*I linker insertions were first mapped by restriction enzyme digestion and the position of those downstream from *Bgl*II site at −833 were determined by sequencing. One of the insertions mapped at position −448; by cleaving the plasmid with *Xho*I and *Hin*dIII, repairing the ends with DNA polymerase I large fragment in the presence of all four dNTPs and recircularization, a deletion to position −448 was obtained. The *CAT* gene was then subcloned into this plasmid as a *Kpn*I/*Sac*I fragment. The Rh1-*CAT* fusion was then excised from pHSS7 by *Not*I cleavage and cloned into the *Not*I site of pDM30.

pP[*ry*; Rh1(−162/+67)-*CAT*]: The same set of manipulations as above were carried out using a linker insertion at position −162 to obtain this plasmid. In the plasmid containing the original *Xho*I linker insertion, an internal deletion occurred, resulting in the *Xho*I linker being placed between nucleotides −219 and −162.

In order to construct plasmids pP[*ry*; Rh1(−252/+67)-*CAT*], pP[*ry*; Rh1(−215/+67)-*CAT*], pP[*ry*; Rh1(−181/+67)-*CAT*], pP[*ry*; Rh1(−166/+67)-*CAT*], pP[*ry*; Rh1(−105/+67)-*CAT*], pP[*ry*; Rh1(−86/+67)-*CAT*], pP[*ry*; Rh1(−46/+67)-*CAT*] and pP[*ry*; Rh1(−35/+67)-*CAT*], the plasmid with the *Xho*I linker insertion at position −448 was cleaved with *Xho*I and then subjected to limited Bal31 digestion. After the Bal31 digestion, plasmid molecules were cleaved with *Hin*dIII, repaired with DNA polymerase I large fragment in the presence of all four dNTPs, and recircularized by DNA ligation. The endpoints of the resultant deletions were determined by subcloning the appropriate portion of the Rh1 promoter into M13 vectors and sequencing. The *Kpn*I/*Sac*I *CAT* module was then cloned into each deletion plasmid and the Rh1-*CAT* fusion was transferred into pDM30 as a *Not*I fragment.

pP[*ry*; Rh1(~−2800/−701, Δ, −448/+67)-*CAT*]: This plasmid was contructed by combining two different *Xho*I linker insertions, one at −701 and the other at −448. A *Hin*dIII/*Xho*I fragment containing the 5′-part of the *ninaE* promoter fragment was isolated from the *Xho*I linker insertion at position −701. It was then cloned into the plasmid containing the *Xho*I linker insertion at −448, after the corresponding *Hin*dIII/*Xho*I fragment was removed. This results in an internal deletion between the sites of the two *Xho*I linker insertions (*i.e.*, from −701 to −448). After the *Kpn*I/*Sac*I *CAT* fragment was cloned into this plasmid, the Rh1-*CAT* fragment was transferred to pDM30 as a *Not*I fragment.

pP[*ry*; Rh1(−162/+67)-*CAT*-Rh1(−501 and −219)] and pP[*ry*; Rh1(−162/+67)-*CAT*-Rh1(−219/−501)]: The plasmid containing the *Xho*I linker insertion between nucleotides −162 and −219 was digested with *Eco*RI (+67) and *Aha*III (−501) and cloned into *Eco*RI/*Sma*I-cut M13 mp18. This places the part of the M13 polylinker adjacent to the *Aha*III site of the *ninaE* promoter. The −219 to −501 region was then isolated from the replicative form of the M13 phage as a *Xho*I (−219)/*Sal*I (M13 polylinker; −501) frag-

CONSTRUCT	LINE	CHROMOSOME	RELATIVE LEVEL OF CAT ACTIVITY IN THE HEAD INDIVIDUAL LINES		MEAN
~-2800 ninaE CAT pP[hs-neo; Rh1(~-2800/+67)-CAT]	1	X (18A)	100 ± 10	(n = 5)	100 ± 10
-833 pP[hs-neo; Rh1(-833/+67)-CAT]	4	3 (72B)	179 ± 4	(n = 3)	151 ± 63
	20	3 (95C)	73 ± 22	(n = 2)	
	26	3 (61A)	133 ± 0.5	(n = 2)	
	41	3 (77B)	220 ± 4	(n = 2)	
-448 pP[*ry*; Rh1(-448/+67)-CAT]	6	3	67 ± 5	(n = 3)	76 ± 10
	21	3	91 ± 6	(n = 3)	
	25	3	71 ± 3	(n = 3)	
	36	2	75 ± 9	(n = 3)	
-252 pP[*ry*; Rh1 (-252/+67)-CAT]	1	2	32 ± 3	(n = 3)	57 ± 25
	2	3	82 ± 15	(n = 3)	
	3	2	59 ± 13	(n = 3)	
-215 pP[*ry*; Rh1(-215/+67)-CAT]	1	X	52 ± 16	(n = 2)	48 ± 7
	2	2	47 ± 8	(n = 3)	
	3	3	40 ± 2	(n = 3)	
	6	X	41 ± 4	(n = 3)	
	7	2	52 ± 10	(n = 3)	
	8	3	59 ± 7	(n = 3)	
-181 pP[*ry*; Rh1(-181/+67)-CAT]	2	X	10.5 ± 0.2	(n = 2)	11.8 ± 2.8
	3	2	15 ± 4.2	(n = 2)	
	5	X	10 ± 2	(n = 2)	
-166 pP[*ry*; Rh1 (-166/+67)-CAT]	1	X	0.16 ± 0.01	(n = 3)	0.5 ± 0.7
	3	2	1.26 ± 0.64	(n = 3)	
	4	3	0.05 ± 0.01	(n = 2)	
-162 pP[*ry*; Rh1(-162/+67)-CAT]	9	X	0.75 ± 0.23	(n = 2)	1.3 ± 1.6
	11	2	3.15 ± 0.89	(n = 3)	
	17	2	0.04 ± 0.01	(n = 3)	
-120 pP[*ry*; Rh1(-120/+67)-CAT]	1	3	6.0 ± 0.17	(n = 2)	5.5 ± 4.0
	2	X	3.3 ± 0.76	(n = 2)	
	3	3	5.1 ± 0.7	(n = 2)	
	5	2	12 ± 0.13	(n = 2)	
	6	2	1.3 ± 0.15	(n = 2)	
-105 pP[*ry*; Rh1(-105/+67)-CAT]	1	3	<0.005		
	2	3	<0.005		
	3	2	<0.005		
	4	3	<0.005		
-86 pP[*ry*; Rh1(-86/+67)-CAT]	1	X	<0.005		
	2	2	<0.005		
	3	2	<0.005		
-46 pP[*ry*; Rh1(-46/+67)-CAT]	1	X	<0.005		
	2	X	<0.005		
	4	2	<0.005		
	5	2	<0.005		
-35 pP[*ry*; Rh1(-35/+67)-CAT]	1	3	<0.005		
~-2800 -701 -448 +67 pP[*ry*; Rh1(~-2800/-701,Δ;-448/+67)-CAT]	8	2	88 ± 17	(n = 3)	88 ± 17

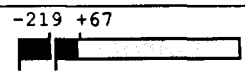
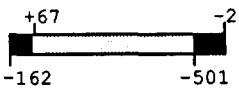
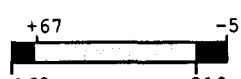
CONSTRUCT	LINE	CHROMOSOME	RELATIVE LEVEL OF CAT ACTIVITY IN THE HEAD INDIVIDUAL LINES		MEAN
 pP[*ry*; Rh1(-501/-219,Δ,-162/+67)-CAT]	1	3	49 ± 3	(n = 3)	39 ± 17
	2	X	57 ± 6	(n = 3)	
	3	X	18 ± 3	(n = 3)	
	6	3	33 ± 0.1	(n = 2)	
 pP[*ry*; Rh1(-162/+67)-CAT-Rh1(-501/-219)]	1	2	25 ± 2	(n = 3)	15 ± 10
	2	3	27 ± 1	(n = 3)	
	4	3	6 ± 2	(n = 3)	
	5	2	13 ± 0.1	(n = 2)	
	6	X	4 ± 1	(n = 2)	
 pP[*ry*; Rh1(-162/+67-CAT-Rh1(-219/-501)]	1	X	58 ± 3	(n = 3)	40 ± 15
	2	2	32 ± 7	(n = 3)	
	4	3	30 ± 8	(n = 3)	

FIGURE 2.—Diagrams of the *ninaE-CAT* fusions and quantitative analysis of the CAT level in corresponding transformant lines. The *CAT* assay is described in MATERIALS AND METHODS. The chromosome of *P*-element insertion is shown; the cytological position of transposon insertion, when determined, is shown in parentheses. The level of *CAT* activity found in each line after a number of separate assays (usually two to three) is expressed as a percentage value of the reference line P[*hs-neo*; Rh1(~−2800/+67)-*CAT*]1. The mean level of activity for a construct is also presented. A level of activity below the limit of detection of the assay is represented as <0.005. The diagrams outline the constructs used in this study. *ninaE* promoter sequences are shown as a solid black rectangle and *CAT* sequences as a stippled rectangle. The 5′ extent of the Rh1 promoter fragment is shown as the number of bp upstream of the transcription initiation site that are still retained in the construct; the 3′ end of the promoter fragment was always at +67 bp in the 5′ untranslated leader. The constructs which were cloned into the pDM30 transformation vector (see MATERIALS AND METHODS) as *Not*I fragments are all in the same orientation; the Rh1 promoter sequence is always close to 5′ end of the *P* element with the *CAT* gene being between the Rh1 promoter and *ry* sequences.

ment and cloned into the pIBI-76 vector digested with *Sal*I and *Xho*I (obtained from IBI Inc., pIBI-76 contains unique *Sal*I and *Xho*I sites). The plasmid pHSS7-Rh1(−162/+67)/*CAT* was modified by repairing the *Sac*I site at the 3′ end of the CAT gene using T4 DNA polymerase in the presence of all four dNTPs and then inserting an eight bp *Xho*I linker. Plasmids pP[*ry*; Rh1(−162/+67)-*CAT*-Rh1(−501/−219)] and pP[*ry*; Rh1(−162/+67)-*CAT*-(−219/−501)] were obtained by inserting into this *Xho*I site the −501/−219 *Sal*I/*Xho*I fragment from the pIBI-76 sublone and then transferring the Rh1/CAT fusion as a *Not*I fragment into pDM30. The plasmid pP[*ry*; Rh1(−501/−219; −162/+67)-*CAT*] was obtained in a similar manner by cloning the −501/−219 *Sal*I/*Xho*I into the *Xho*I site at −162.

pP[*ry*; Rh1(~−2800/−816; Δ; −120/+67): The Rh1 promoter subclone (~−2800/+67) in pHSS7 was digested by *Sal*I and the digestion products were religated. Into the resultant plasmid the *CAT* gene was cloned as a *Kpn*I/*Sal*I fragment. The Rh1/*CAT* fusion was then transferred into pDM30 as a *Not*I fragment resulting in the plasmid pP[*ry*; Rh1(~−2800/−816; Δ; −120/+67)-*CAT*].

Construction of *ninaE* promoter − *lacZ* gene fusions: To more precisely determine the pattern of expression from the *ninaE* promoter, two promoter fragments were fused to β-galactosidase indicator gene. Plasmid pP[*hs-neo*; Rh1-(−833/+67)-β-gal] was constructed by cloning the *Bgl*II (−833)/*Kpn*I (+67) *ninaE* promoter fragment into the vector pDM79 which had been digested with *Bam*HI and *Kpn*I. The plasmid pP[*hs-neo*; Rh1(−120/+67)-β-gal] was obtained by cleaving pP[*hs-neo*; Rh1(−833/+67)-β-gal] with *Sal*I and religating the plasmid. Diagrams of these two constructs are shown in panel E of Figure 6.

Choramphenicol acetyltransferase assays: Flies to be assayed were manually decapitated using a razor blade and their heads and bodies were placed into separate 1.5 ml microcentrifuge tubes. After freezing in a −80° freezer for 15 min fly heads or bodies were homogenized in 100 μl of 0.25 M Tris (pH 7.7). The homogenate was frozen by placing in a −80° freezer for several minutes, thawed and sonicated. Microfuge tubes were then placed in a 65° water bath for 5 min followed by centrifugation in a microcentrifuge for 10 min at 4°. An aliquot of the supernatant, 20 μl of 4 mM acetyl-coenzyme A, 1 μl of ^{14}C-labeled chloramphenicol (0.2 μCi) and 0.25 M Tris (pH 7.7) were combined to give a final volume of 140 μl. After incubation at 37° reactions were stopped by extraction with 500 μl of ethyl acetate. Further treatment of the samples is identical to a previously published procedure (GORMAN, MOFFAT and HOWARD 1981).

After preliminary assays of the transformant lines belonging to each construct, assays were quantitated by allowing the acetylation reaction to proceed for a limited period of time during which the reaction was in the linear range (10 min to 2 hr depending on the level of activity observed in the preliminary assays for that particular construct). To more accurately quantitate the low activity found in certain lines, an extract of the control line P[*hs-neo*; Rh1(~−2800/+67)-*CAT*]1 was diluted 1:1000, 1:100 and 1:20 with an extract of *ry506* host flies and the observed level of acetylation in extracts of the low-level lines was directly compared to that observed in the diluted control extracts. The amount of acetylation was determined by liquid scintillation counting of that portion of the thin layer chromatography plates, used to separate the reaction products, that contained the ^{14}C-labeled reaction products (1-*O*-acetyl- and 3-*O*-acetyl-chloramphenicol). The scintillation cocktail used was Betamax (Westchem Corp.).

***In situ* hybridization to polytene chromosomes:** To determine the number and position of *P*-element vector insertions into the genomes of the transformed flies, *in situ* hybridization to polytene chromosomes was carried out for lines where G418 selection was utilized for the isolation of transformant lines. Polytene chromosome squashes were

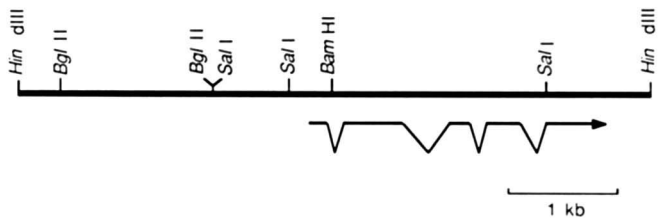


FIGURE 3.—A restriction enzyme map of the 5.4 kb *Hin*dIII fragment containing the entire coding region of the *ninaE* gene and approximately 2.8 kb of 5′-flanking sequence is shown. Below the map, the intron/exon arrangement of the *ninaE* transcript is outlined.

prepared as previously described (ZUKER, COWMAN and RUBIN 1985). Squashes were hybridized with biotinylated DNA probes corresponding to plasmid pUC18. The procedure was carried out as described by LANGER-SOFER, LEVINE and WARD (1982) except that the DNA was labeled by nick translation using Bio-16-dUTP (ENZO Biochem) and hybridization sites were detected by using the Detek-I-HRP detection kit produced by ENZO Biochem.

Isolation of *D. melanogaster* genomic DNA and DNA blotting: In lines transformed with constructs in the pDM30 vector the chromosome of insertion was determined by segregation against appropriate balancers in a *ry* background. Subsequently, the number of insertions in the particular chromosome was determined by carrying out DNA

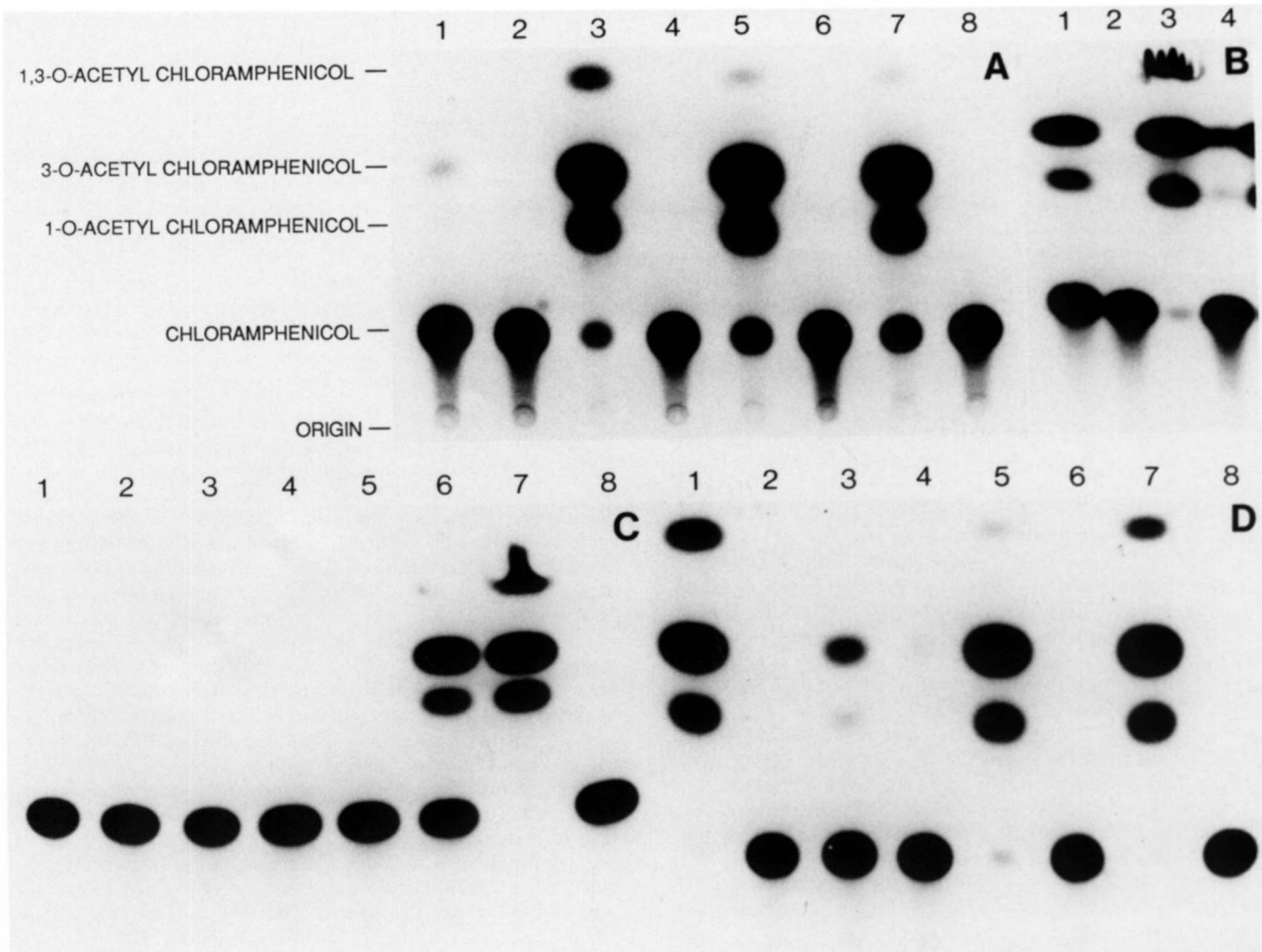


FIGURE 4.—Autoradiographs of chloramphenicol acetyltransferase assays of flies belonging to transformant lines carrying Rh1-*CAT* fusions. *CAT* assays were carried out as described in MATERIALS AND METHODS and visualized by autoradiography. Panel A shows an autoradiograph of an assay in which extracts from either five heads or five bodies of each transformant line were incubated with substrate for two hours. Lanes 1 and 2 represent head and body assays of line P[*ry*; Rh1(−162/+67)-*CAT*]17, respectively. There is little accumulation of major monoacetylated product in this assay indicating the presence of only a low level of *CAT* activity in the heads of these flies. No activity can be detected in the body extract. In contrast, head extracts from transformant lines P[*hs-neo*; Rh1(−833/+67)-*CAT*]4, P[*ry*; Rh1(−448/+67)-*CAT*]6 and P[*hs-neo*; Rh1(~−2800/+67)-*CAT*]1 (lanes 3, 5 and 7, respectively) show an abundance of both monoacetylated products as well as the appearance of the diacetylated chloramphenicol indicating a high level of *CAT* activity. No *CAT* activity is detectable under these conditions in the extract from the bodies of flies from these same transformant lines (lanes 4, 6 and 8, respectively). In order to test whether the low level of *CAT* activity detected in transformant line P[*ry*; Rh1(−162/+67)-*CAT*]17 is completely head-restricted, we have carried out same type of assay using ten-times more material (50 *vs.* 5 flies). The results are shown in panel B. Although an appreciable level of chloramphenicol acetylation is now observed in the head extract of line P[*ry*; Rh1(−162/+67)-*CAT*]17 (lane 1), no activity can be detected in the body extract (lane 2). On the other hand, in transformant line P[*ry*; Rh1(−162/+67)-*CAT*]11, in addition to a high level of chloramphenicol acetylation in the head extract (lane 3), there is a low level of acetylation in the body extract (lane 4). This transformant

blots to genomic DNA isolated from each transformant line. The restriction enzyme digest utilized to determine the number of ends was either *Sac*I or *Xho*I depending on the construct and the hybridization probe was ^{32}P-labeled *CAT* DNA. The number of bands observed corresponds to the number of 5′-*P*-element ends. DNA blotting was carried out as previously described (MANIATIS, FRISCH and SAMBROOK 1982). Rapid small scale isolation of DNA from flies of the different transformant lines was performed by the procedure of STELLER and PIRROTTA (1986).

***In situ* hybridization to tissue sections:** Adult *D. melanogaster* heads were sectioned (8 μm frozen cryostat sections) and treated as described by HAFEN *et al.* (1983), except that the acid and pronase pretreatments were omitted.

β-Galactosidase histochemical staining: Histochemical staining for β-galactosidase was carried out as described by D. R. KANKEL and M. FITZGERALD (personal communication). Cryostat sections (8 μm) were fixed in 2% glutaraldehyde in disodium phosphate buffer (pH 7.5). After the sections were washed free from the fixative, a staining gel consisting of 100 μg/ml 5-bromo-4-chloro-3-indolyl-β-D-galactopyranoside (X-gal), 5 mM potassium ferricyanide, 5 mM potassium ferrocyanide, 2% gelatin in citric acid/phosphate buffer (pH 7.8) was applied. Although *D. melanogaster* adults possess endogenous β-galactosidase activity, the pH optimum for this enzyme is approximately 5 and it is largely inactive at the pH used during the histochemical staining procedure. For the wholemount staining of the anterior part of larvae and adult heads, the same procedure was used except that the staining solution did not contain gelatin.

RESULTS

Fragments containing either 2.8 kb or 0.9 kb of the *ninaE* promoter appear to be able to direct *CAT* and β-galactosidase expression to retinula cells R1–R6: The plasmid pP[*hs-neo*; Rh1(~−2800/+67)-*CAT*] contains a transcriptional fusion between nucleotides ~−2800/+67 of the *D. melanogaster ninaE* promoter (transcription initiation site is at nucleotide +1) and the *E. coli* chloramphenicol acetyltransferase (*CAT*) gene. The 5′ end of this fragment coincides with that of the 5.4 kb genomic *Hin*dIII fragment shown in Figure 3. This 5.4-kb DNA segment is sufficient to rescue the *ninaE* mutant phenotype when introduced into the Drosophila germline by *P*-element transformation (C. S. ZUKER, D. MISMER and G. M. RUBIN, unpublished data). Following microinjection of pP[*hs-neo*; Rh1(~−2800/+67)-*CAT*] into *D. melanogaster* Canton-S embryos, a transformant line was obtained. *CAT* activity in this transformant line was found to be restricted to the head (Figure 4A). The *ninaE* gene is expressed solely in the visual system (O'TOUSA *et al.* 1985; ZUKER, COWMAN and RUBIN 1985), which in the adult is found only in the head. Thus, the observed head-specific *CAT* activity is an indication that the promoter fragment reestablishes, at least in part, the pattern of expression of the *ninaE* gene. The absence of body *CAT* activity is not due to body-specific degradation of *CAT* protein or its mRNA as can be seen by examining the *CAT* activity found in a transformant line containing a nearly identical *CAT* module fused to the *D. melanogaster* hsp70 promoter (gift of M. C. MULLINS). In this line high levels of *CAT* activity are present in both the head and body (data not shown).

The 2.8-kb promoter fragment was shortened to contain 833 bp of upstream sequences. When transformant lines were obtained for this construct, pP[*hs-neo*; Rh1(−833/+67)-*CAT*], their *CAT* activity was also restricted to the head and was similar in level to that found in the transformant line pP[*hs-neo*; Rh1-(~−2800/+67)-*CAT*]1 (Figure 2A). The nucleotide sequence of this promoter fragment, extending from −833 bp to +67 bp, is shown in Figure 5.

In order to more precisely determine the tissue specificity of expression directed by the −833/+67 promoter fragment we constructed a fusion between this fragment and the *E. coli lacZ* gene, pP[*hs-neo*; Rh1(−833/+67)-β-gal]. The pattern of β-galactosid-

line contains the highest level of *CAT* activity observed in the three transformant lines obtained with this construct, suggesting that this transposon inserted into a genomic region close to another enhancer-like element. This element may partially substitute for the absence of *ninaE*'s quantitative control elements, resulting in a higher level of *CAT* activity in the head, but may also override *ninaE* control elements maintaining head-restricted expression, resulting in a low level of body expression. The third transformant line belonging to this construct P[*ry*; Rh1(−162/+67)-*CAT*]9 also exhibited head-restricted *CAT* activity (data not shown). Panel C shows a developmental *CAT* assay performed on the transformant line P[*ry*; Rh1(−215/+67)-*CAT*]7 by analysing five individuals belonging to different developmental stages. No activity can be detected in following stages: embryo (lane 1), first, second and third larval instar (lanes 2, 3 and 4, respectively), early pupa (white, no adult structures discernible; lane 5) and body (lane 8). *CAT* activity can be detected in the stages that have begun or completed the morphogenesis of the rhabdomere, the light-sensing organelle of the retinula cells: late pupa (red eyes, darkening of the cuticle but no adult bristles; lane 6) and head (lane 7). Panel D shows an experiment in which sequences of the *ninaE* promoter extending from −501 to −219 have been placed 3′ to a *CAT* fused to nucleotide −162 to +67 of the *ninaE* promoter. All reactions contained five heads or bodies and were incubated for 2 hr. Lanes 1 and 2 show control *CAT* assay of heads and bodies of flies belonging to the transformant line P[*ry*; Rh1(−215/+67)-*CAT*]3; a high level of product accumulation is seen in the head lane while no activity was detected in the body extract. Lane 3 shows an assay of a head extract of line P[*ry*; Rh1(−166/+67)-*CAT*]1; a low level of chloramphenicol acetylation is observed. Although the body extract of the same line (lane 4) contains some activity, this was not reproducibly seen and probably represents minor contamination of this extract. These two transformant lines are controls and contain a level of *CAT* activity characteristic of lines where quantitative control elements are present (P[*ry*; Rh1(−215/+67)-*CAT*]3) or absent (P[*ry*; Rh1(−166/+67)-*CAT*]1). Lanes 5 and 7 show the assays of head extracts belonging to lines P[*ry*; Rh1(−162/+67)-*CAT*-Rh1(−501/−219)]3 and P[*ry*; Rh1(−162/+67)-*CAT*-Rh1(−219/−501)]4; there is a similar level of *CAT* activity to that observed for the control line P[*ry*; Rh1(−215/+67)-*CAT*]3. The assays shown in this figure are not in the linear range; assays of these lines were carried out with less material to obtain the numbers shown in Figure 2. Body extracts of these two lines (lanes 6 and 8) show no detectable *CAT* activity. Therefore, the sequences from −501 to −219 can function as an enhancer element when placed downstream from the truncated *ninaE* promoter.

```
-833 AGATCTGCCG ACATGTGTCG ACTGTCACAC CAGCTGACAG TTATTTCGGT TTAGTTCCAT ATTTGGGTTG  -764
     Bgl II          Sal I

-763 GGCGCAAAGC GTGGCATCCC ATTTGGCCAT TTGGATTAGC TAAAACTTGA CCCAATTCAA GCTGATGTGT  -694

-693 TGATCCATCG CACAAAGGCG AACAAAAGGC TTTTATGCCA AAGCTGGACT GAAAAAATAA AATCCAAGCT  -624

-623 GCCCACACCA AAAAGAAAAG AAACAAAACA TGTTCGGAAA ATTCAATGGA GGTGCGAAAA GATCAATCAG  -554

-553 CACAAAAAAT CCATAAGCGG CTTAGCGTCT GGGGTGTTCA TATTCTTAAT TTAAATAGAT TATTTTACTA  -484
                                                          Aha III

-483 CTTCCAAAAG CTGGCATTCT TTTTAAGATA ATCCAAGATT AGCAGAGCCC TCAAGATGTG CGTTCATCG   -414

-413 CGCACACACT GCGTATACTT GATTGCGCAT ACGCACTGGT GTCGCAGGCG AACGATACTG ATATTTCTTC  -344

-343 TGCTACGATA TGTTTTCATA ATATTCTTCA AATTTGTCTA GGCATTGACA CATTAAATCG CTGACTTTCG  -274

-273 CCATTGCTTT CTGCACAATC TGGAGACTCA AGAATAATAC TCGGCCAGAA CACCCAGTGG AAACCCTTGA  -204

-203 AATGCCTTTA ACTATTGACA TTGGCCCATT GCGATGTGCG CTTGCTTTCC GTTATGGATA TTATGATTAT  -134

-133 AAGCCAAAAG CTGTCGACAC TTTCCTCTGC ACATTGCAGA CATATTTTTA GGCTGGCCTA ATTGAATTTC   -64
                 Sal I
                                                                    +1
 -63 AAGAGCAGGG GTTACACGAG CATTGCGAAA ACTATAAAAG CCACGCGGCA GAATGCAGAC ATTGCAGGTT    +7

  +8 TCCAACGACC AATCGCCGCG ACTAGTCCGC CCCAGTGAAA TATTCAGAAT CCAGGACCCT              +67
```

FIGURE 5.—DNA sequence of the *ninaE* promoter region. The *ninaE* gene was sequenced from the *Bgl*II site located 833 bp upstream from the site of transcription initiation to the position +67 in the 5′-untranslated leader. Selected restriction enzyme sites are indicated. The sequence TATAAAA that conforms to the consensus sequence for the TATA box (GOLDBERG 1979) and is located between nucleotides −31 and −25, is also underlined. No large (*i.e.*, greater than 10 bp) direct or inverted repeats exist within the sequence.

ase expression in the adult head and larval photoreceptor organ (stemmata) of transformants can be seen in Figure 6. The β-galactosidase staining observed in cryostat sections of adults from two independent transformant lines was restricted to the retina of the compound eye, the site of the rhodopsin gene expression, and the lamina, the first synaptic region of the optic lobe. No staining of the ocelli was observed either in cryostat sections or wholemounts of the heads of transformant flies. Transcripts of the *lacZ* gene were limited to the retina as determined by *in situ* hybridization to head cryostat sections with a tritiated fragment from the *lacZ* gene (data not shown), indicating that diffusion of β-galactosidase molecules down the axons of the R1–R6 cells rather than transcription of the gene in the cells of the lamina accounts for the laminar staining. β-Galactosidase activity was also detected in a subset of cells belonging to the larval photoreceptor organ. The larvae of hombometabolous insects contain visual organs termed stemmata (for review, see CHAPMAN 1982; CARLSON and CHI 1979). They are found at the anterior end of the larva; the stemmata of Musca are associated with the cytopharyngeal skeleton (BOLWIG 1946). The Drosophila stemmata are found in the same approximate location (FUJITA *et al.* 1982; HOTTA and KENG 1984) and contain 12 photoreceptor cells (H. STELLER, personal communication). In the larvae of transformant line P[*hs-neo*; Rh1(−833/+67)-β-gal]1 four or five of the 12 cells present within each organ appear to stain (Figure 6).

By using fragments containing as little as 833 bp of the *ninaE* promoter and about one third of the 5′-untranslated leader we have succeeded in conferring the pattern of expression characteristic of *ninaE* gene onto two different transcriptional fusion modules (*CAT* and β-galactosidase) in the absence of any coding regions. These results indicated that no essential control elements exist within the *ninaE* coding sequence, introns or 3′-flanking sequences.

Further deletion analysis of the *ninaE* promoter: A set of further deletions was constructed using two techniques: DNaseI cleavage and Bal31 deletion (see MATERIALS AND METHODS). In this way we were able to generate promoter fragments extending from −448, −252, −215, −181, −166, −162, −120, −105, −86, −46 and −35 to +67. Each of these was fused to *CAT* and assayed following its reintroduction into the genome by *P*-element-mediated transformation. Special care was taken to minimize the source of variability between the constructs: multiple independent transformant lines were obtained for the majority

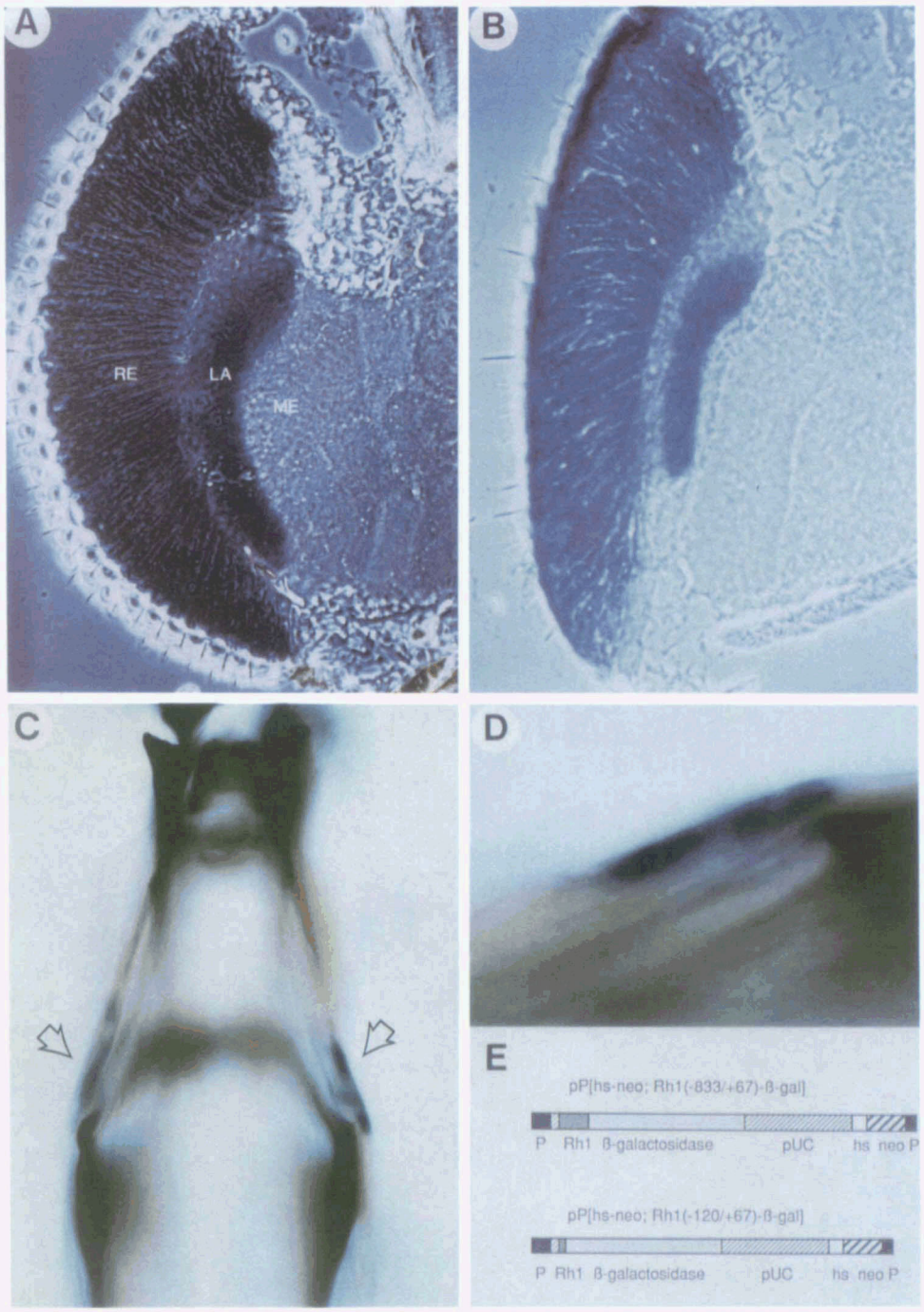


FIGURE 6.—β-Galactosidase histochemical staining of Rh1-β-galactosidase fusions. Panel E shows the diagram of the plasmids containing Rh1-β-galactosidase fusions used to obtain the transformant lines analyzed. Plasmids pP[*hs-neo*; Rh1(−833/+67)-β-gal] and pP[*hs-neo*; Rh1(−120/+67)-β-gal] contain 833 and 120 bp upstream of the transcription initiation site of the *ninaE* gene, respectively. Transformant lines containing these fusion genes were crossed into a *white*$^-$ genetic background to facilitate the histochemical staining. Two transformant lines were obtained for plasmid pP[*hs-neo*;; Rh1(−833/+67)-β-gal]; the sites of transposon insertion are at *72DE* and *96F*. Panel A shows a phase photomicrograph of a head section of a fly from the transformant line P[*hs-neo*; Rh1(−833/+67)-β-gal]1 after β-gal histochemical staining. Staining can be seen in the retina and lamina (labeled RE and LA, respectively; ME specifies medulla, another synaptic region of the optic lobe). The same pattern of staining has been observed in another transformant line P[*hs-neo*; Rh1(−833/+67)-β-gal]5 which carries the same construct (data not shown). Panels C and D show wholemount β-gal histochemical staining of the larval photoreceptor organ in the larvae belonging to line P[*hs-neo*; Rh1(−833/+67)-β-gal]1. Arrows in C indicate the position of larval photoreceptor organ. Only four to five cells out of approximately 12 cells present in larval photoreceptor show detectable levels of β-gal activity. Panel B shows the pattern of staining observed in an adult head section of a fly from transformant line P[*hs-neo*; Rh1(−120/+67)-β-gal]1 in which the transposon is inserted at position *43E* on 2R. The pattern of staining observed corresponds closely to the observed in the line P[*hs-neo*; Rh1(--833/+67)-β-gal]1 indicating that the tissue-specificity of expression has not been altered.

of the constructs analyzed, flies assayed were of the same age (7 days) and every line contained only a single *P*-element insertion into the genome. The results of this analysis are shown in Figure 2. During the construction of each deletion, a novel junction was created between the *P*-element sequences and *ninaE* promoter sequences. Although it cannot be formally excluded, we believe that the creation of this novel junction did not influence tissue-specificity since the characteristic pattern of expression was never altered. At least two different control regions have been detected. Removal of the first control region, located between nucleotides −215 and −162, resulted in at least a 30-fold reduction in *CAT* activity. However, the low activity observed was still head-restricted (Figure 4, A and B). Moreover, in a transformant line, P[*hs-neo*; Rh1(−120/+67)-β-gal]1, in which nucleotides −120 bp to +67 bp are fused to *lacZ*, β-galactosidase expression is limited to the retinula cells R1–R6 (Figure 6B). These results indicate that sequences upstream of nucleotide −120 are not essential for the proper tissue-specificity of the *ninaE* promoter. The second control region found 3′ to the nucleotide −120 is essential for promoter function in the absence of the upstream element(s) controlling the level of expression; removal of sequences between −120 and −105 reduces the *CAT* activity to an undetectable level.

The smallest promoter fragment that still behaves essentially as the largest promoter fragment used in this study extends from −215 to +67 (Figure 2). A transformant line containing this promoter fragment, P[*ry*; Rh1(−215/+67)-*CAT*]7, was assayed to determine the developmental timing of *CAT* expression (Figure 4C). Expression of the *ninaE* gene, as detected by RNA blotting (ZUKER, COWMAN and RUBIN 1985; O'TOUSA *et al.* 1985), begins during the late pupal period (48–60 hr after puparium formation) when the final stages of photoreceptor cell differentiation occur. The temporal pattern of *CAT* expression in this transformant line appears to follow that of the endogenous gene. The absence of detectable *CAT* activity in larval stages is probably due not to the lack of expression of the fusion gene in the larval photoreceptor organ but rather to the small number of cells expressing the gene. That the level of activity in adults is higher than in pupae is expected since *ninaE* mRNA levels are known to increase during adult life.

The removal of promoter sequences to position −215 did have a small quantitative effect. We wished to test whether this small effect resulted from the removal of an enhancer-like element. Enhancers are DNA sequences that can elevate the level of expression of a given promoter in position- and orientation-independent manner (for review, see SERFLING, JASIN and SCHAFFNER 1985). Often, promoters contain multiple, partially redundant sequences capable of exerting an enhancer effect. Deletion of any one of these sequences reduces the quantitative level of expression from the promoter only slightly in comparison with the reduction observed when all enhancer sequences are absent. By analogy, the 5′-flanking sequence of the *ninaE* gene may contain a number of enhancer sequences which may, at least in part, substitute for one another. We constructed the plasmids pP[*ry*; Rh1(−162/+67)-*CAT*-Rh1(−501/−219)] and pP[*ry*; Rh1(−162/+67)-*CAT*-Rh1(−219/−501)] in which the sequences from −501 to −219 are placed downstream from a *ninaE/CAT* gene fusion containing nucleotides −162 to +67 of the *ninaE* promoter. *CAT* expression in the lines obtained after transformation of these plasmids was restricted to the head and was comparable in level to that found in the transformant lines containing more than 215 bp of upstream sequences (Figure 4D) or about 20- to 30-fold greater than in transformants with only 162 bp of upstream sequences (Figure 2). This strongly suggests that the sequences located between nucleotides −219 and −501 behave as an enhancer; they can function 3′ to the promoter fragment, in both orientations, and over a distance of approximately 1.6 kb. Moreover, it appears that these sequences are partially redundant since in their absence there is only a two-fold reduction in expression.

In one of our constructs, an internal deletion within the 2.8-kb promoter fragment that removes the segment of DNA between the two *Sal*I sites (−816 to −120 bp), a low level of expression was seen in bodies of the five independent transformant lines while no activity was detected in heads (data not shown). The repression of head activity is particularly striking since the lines corresponding to the plasmid pP[*ry*; Rh1(−120/+67)-*CAT*] have a significant level of head-specific *CAT* activity and no detectable activity in the body. We believe that this pattern of expression is due to the overriding effect of another enhancer-like element located between ~−2800 and −833 bp. The relevance, if any, of this control element to the regulation of the endogenous *ninaE* gene is unclear.

DISCUSSION

We have performed a preliminary analysis of the *cis*-acting sequences involved in regulation of the *ninaE* locus. We have succeeded in transferring the tissue-specific pattern of expression of this gene onto two distinct bacterial indicator genes by using solely *ninaE* sequences upstream of its transcription start point and a part of its 5′-untranslated leader. Our results indicate that no essential regulatory elements exist in the *ninaE* coding sequence, introns or 3′-flanking sequences. Deletion analysis revealed the presence of at least three separable *cis*-acting promoter elements (see Figure 7). Removal of both control elements I and II (Figure 7) resulted in at least a 30-fold reduction in the expression from this pro-

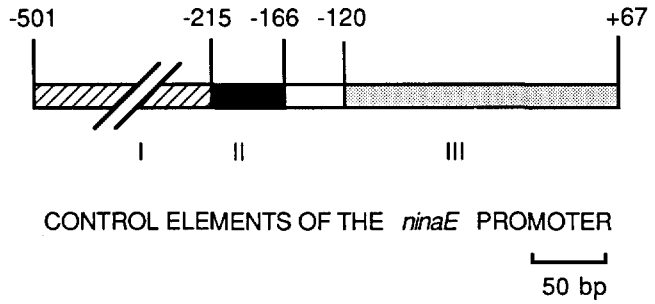


FIGURE 7.—Schematic representation of the *cis*-acting regulatory elements of the *ninaE* promoter. The position and properties of the control elements II and III have been determined from the deletion analysis of the promoter sequences carried out by *P*-element-mediated transformation. Control element I, which has the properties of an enhancer, has been detected by placing the DNA fragment containing sequences between nucleotides −219 and −501 downstream from a truncated *ninaE* promoter. This enhancer element is redundant; its removal has only a twofold effect on promoter efficiency. Control element II appears also to be a quantitative regulatory element. Control element III, located between nucleotides −166 and +67, contains all the sequences necessary for proper tissue-specificity of this promoter and can function independently from elements I and II, albeit at reduced levels.

moter. The remaining low level of expression was still head-restricted, indicating that elements I and II are not the sole determinants of tissue specificity; a fragment that retains only 120 bp upstream of the transcription initiation site still appears to direct expression to the retinula cells R1–R6. Therefore, all the regulatory elements necessary for the proper tissue-specific expression from this promoter are located between nucleotides −120 and +67 (element III in Figure 7) and can function independently from elements I and II which appear to control transcriptional efficiency. We have not rigorously determined whether the transcription initiation point characteristic of the endogonous gene (ZUKER, COWMAN and RUBIN 1985; O'TOUSA *et al.* 1985) is used in our transformant lines. However, our constructs initiate transcription in an *in vitro* transcription system derived form a Kc cell nuclear extract (prepared according to HEBERLEIN, ENGLAND and TJIAN 1985) at a site identical to that found for the endogenous *ninaE* gene (data not shown). It is, therefore, highly likely that same transcription initiation site is used in the transformant lines.

For several developmentally regulated Drosophila promoters, analysis by *P*-element-mediated transformation has revealed the existence of multiple control elements. Separate control elements required for the expression in testes and Malpighian tubules have been described for the *white* gene (LEVIS, HAZELRIGG and RUBIN 1985; PIRROTTA, STELLER and BOZZETTI 1985). Different control elements determine the expression of two yolk protein genes yp1 and yp2 in the ovaries and fat body (GARABEDIAN, HUNG and WENSINK 1985). Separate control elements have also been implicated in the activation of hsp27 and hsp23 by the heat shock and ecdysteroids (RIDDIHOUGH and PELHAM 1986; MESTRIL *et al.* 1986). COHEN and MESELSON (1985) demonstrated separate elements needed for heat-shock inducible and ovarian expression of hsp26 gene. BOUROUIS and RICHARDS (1985) have separated the quantitative and qualitative control elements in the Sgs3 glue protein gene. All of the regulatory elements identified in these studies are located 5′ to the gene. However, FISCHER and MANIATIS (1986) described a regulatory element of *Drosophila mulleri* Adh2 gene that is located at least 600 bp downstream of the Adh2 polyadenylation site; this element is essential for the expression of Adh2 gene in the larval Malpighian tubules after introduction of the *D. mulleri* Adh2 gene into the germline of *D. melanogaster*. In *D. melanogaster* dopa decarboxylase gene separate elements appear to regulate expression in CNS and hypoderm (SCHOLNICK *et al.* 1986). Thus, multiple control regions appear to be a common characteristic of developmentally regulated genes in *D. melanogaster*.

Many genes have been reintroduced into the *D. melanogaster* germline and, in general, these genes appear to be correctly regulated despite their presence at chromosomal positions that differ from that of the endogenous gene (SPRADLING and RUBIN 1983; SCHOLNICK, MORGAN and HIRSH 1983; GOLDBERG, POSAKONY and MANIATIS 1983; JOWETT 1985; MARSH, GIBBS and TIMMONS 1985; KALFAYAN, WAKIMOTO and SPRADLING 1984; KRUMM, ROTH and KORGE 1985; HAZELRIGG, LEVIS and RUBIN 1984; PIROTTA, STELLER and BOZZETTI 1985; LEVIS, HAZELRIGG and RUBIN 1985). However, the minimal length of the promoter needed to mimic the qualitative and quantitative regulation of the endogenous gene is highly variable and has been carefully determined in only a few cases. The *ninaE* promoter analyzed in this study requires one of the shortest promoters, less than 300 bp, to obtain apparently normal expression. On the other extreme, 6.1 kb of 5′-flanking sequence is necessary to obtain the rescue of *fushi tarazu* phenotype (HIROMI, KUROIWA and GEHRING 1985).

Some quantitative variation among different transformant lines carrying the same construct has been observed in all studies using *P*-element-mediated transformation. The variation observed in the present study appears to be the greatest for constructs missing control elements I and II; presumably the truncated *ninaE* promoter stripped of these elements becomes more responsive to other control sequences present in the region of transposon integration. Also, a higher level of expression was obtained when G418 was used to select transformants. This is probably due to the fact that high level expression of the G418 resistance gene is required to confer G418 resistance (STELLER

and PIRROTTA 1985). Therefore, one may be biased toward recovering G418-resistant lines that carry insertions into chromosomal regions favoring a high level of gene expression. Random chromosomal locations are more probable when using *ry* as a marker, since less than 2% of wild type ry^+ expression is needed to rescue the rosy eye color phenotype.

Control element I, located between nucleotides −501 and −219, appears to have the properties of an enhancer element. This sequence is capable of elevating the level of expression from the *ninaE* promoter, in the absence of control element II, in a position- and orientation-independent manner. Experiments are now in progress to determine whether the control element II, located between nucleotides −215 and −166, also behaves as an enhancer. Several other position- and orientation-independent control elements have been reported in the Drosophila genome: the yolk protein 1 enhancer (GARABEDIAN, SHEPHERD and WENSINK 1986), the alcohol dehydrogenase enhancer that acts on the proximal Adh promoter (POSAKONY, FISCHER and MANIATIS 1985), the *ftz* enhancer (HIROMI, KUROIWA and GEHRING 1985) and the Sgs4 enhancer (SHERMOEN *et al.* 1987). The Adh, yp1 and Sgs4 enhancers appear to be tissue-specific elements that are capable of conferring a pattern of expression characteristic of their parent genes onto a heterologous promoter element. Similar tissue-specific enhancers have been identified in the genes of the mammalian immune system (NEUBERGER 1983; BANERJI, OLSON and SCHAFFNER 1983; GILLIES *et al.* 1983; PICARD and SCHAFFNER 1984; QUEEN and BALTIMORE 1983). Although the removal of the enhancer sequence identified in this study does not alter the tissue-specific expression of the *ninaE* promoter, it cannot be ruled out that the enhancer itself carries determinants of tissue-specific expression. This enhancer is apparently redundant; in its absence expression is reduced at most twofold. Enhancer redundancy has not been previously demonstrated in *D. melanogaster*, although it is common in the enhancer regions of several animal viruses (for review, see SERFLING, JASIN and SCHAFFNER 1985). Also, three partially redundant enhancer elements have been identified in the mouse α-fetoprotein gene (HAMMER *et al.* 1987). Although all three enhancers are functionally equivalent in transient transfection assays, *in vivo* these elements varied in their ability to modulate expression in different tissues.

Fusions between the *ninaE* promoter fragments (extending from either −833 bp or −120 bp to +67 bp) and the *lacZ* gene have been useful in demonstrating the tissue-specific expression of this promoter fragment. It is noteworthy that β-galactosidase molecules diffuse into the axons of the photoreceptor cells R1–R6. If such intracellular diffusion generally occurs in nerve cells, then expression of β-galactosidase becomes a powerful technique for staining a subset of neurons.

Expression of the *ninaE*-β-galactosidase fusion gene in the larval photoreceptor was an unexpected observation since larvae lacking a functional *ninaE* gene behave as their wild-type counterparts in assays of phototactic behavior (HOTTA and KENG 1984). Our analysis suggested that the *ninaE* gene is expressed in only a subset of the cells belonging to the larval photoreceptor and thus it is likely that other adult opsins are expressed in different subsets of larval photoreceptor cells. This hypothesis is now being tested by constructing fusions between other adult opsins and appropriate indicator genes.

The functional differences between cells within the ommatidium must result in large part from transcriptional regulatory decisions that are made during ommatidial assembly and subsequent morphogenesis. Regulatory identities are established in the absence of cell division in a relatively short timespan in a previously homogenous cell population. Our long-term goal is to understand how the positional cues used in the assembly process are interpreted and lead to differences in transcription. By identifying the *cis*- and *trans*-acting regulators of genes which are restricted to specific photoreceptor subpopulations we hope to ultimately reconstruct the hierarchy of decisions that result in a particular cell type. Understanding the regulatory differences between the different cell types and how they are brought about should increase our knowledge of the regulatory mechanisms occurring during the final stages of organogenesis.

The authors would like to thank CHARLES ZUKER for generously providing the *ninaE* gene, CARL THUMMEL and DAVID HOGNESS for providing the chloramphenicol acetyltransferase and β-galactosidase genes, FRED HEFFRON for donating plasmid pHSS7, HELI ROIHA for plasmids of pEMBL 18+ and pEMBL19+, MARY MULLINS for donating a transformant line carrying hsp70-*CAT* fusion, W. MATTHEW MICHAEL for help in the construction of some of the enhancer constructs and TODD LAVERTY for determining the chromosomal position of transformant lines. We would like to thank HERMANN STELLER, DANIEL KALDERON, ANDREW TOMLINSON, ULRIKE HEBERLEIN, JAMES W. FRISTROM and ROBERT TJIAN for reading the manuscript and their insightful comments. D.M. would particularly like to thank DONALD C. RIO for his moral support and guidance during this study. During this study, D.M. was supported in part by an Ernest Brown Babcock Fellowship and a Regents Scholarship. This work was supported by National Institutes of Health grant GM33135 to G.M.R.

LITERATURE CITED

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.

BANKIER, A. T. and B. G. BARRELL, 1983 Shotgun DNA sequencing. pp. 1–33. In: *Techniques in Nucleic Acid Biochemistry*, Vol. B5, edited by R. A. FLAVELL, Elsevier, Amsterdam.

BOLWIG, N., 1946 Senses and sense organs of the anterior end of the house fly larvae. Vidensk. Medd. Dan. Naturhis. Foren. **109:** 81–196.

BOUROUIS, M. and G. RICHARDS, 1985 Remote regulatory sequences of the Drosophila glue gene Sgs3 as revealed by P-element transformation. Cell **40:** 349–357.

CARLSON, S. D. and C. CHI, 1979 The functional morphology of the insect photoreceptor. Annu. Rev. Entomol. **24:** 379–416.

CHAPMAN, R. F., 1982 *The Insects, Structure and Function.* Harvard University Press, Cambridge, Mass.

COHEN, R. S. and M. MESELSON, 1985 Separate regulatory elements for the heat-inducible and ovarian expression of the Drosophila hsp26 gene. Cell **43:** 737–746.

COWMAN, A. C., C. S. ZUKER and G. M. RUBIN, 1986 An opsin gene expressed in only one photoreceptor cell type of the Drosophila eye. Cell **44:** 705–710.

DENTE, L., G. CESARENI and R. CORTESE, 1983 pEMBL: a new family of single stranded plasmids. Nucleic Acids Res. **11:** 1645–1655.

FISCHER, J. A. and T. MANIATIS, 1986 Regulatory elements involved in Drosophila Adh gene expression are conserved in divergent species and separate elements mediate expression in different tissues. EMBO J. **5:** 1275–1289.

FRYXELL, K. J. and E. M. MEYEROWITZ, 1987 An opsin gene that is expressed only in the R7 photoreceptor cell of Drosophila. EMBO J. **6:** 443–451.

FUJITA, S. C., S. L. ZIPURSKY, S. BENZER, A. FERRUS and S. L. SHOTWELL, 1982 Monoclonal antibodies against the Drosophila nervous system. Proc. Natl. Acad. Sci. USA **79:** 7929–7933.

GARABEDIAN, M. J., M.-C. HUNG and P. C. WENSINK, 1985 Independent control elements that determine yolk protein gene expression in alternative Drosophila tissues. Proc. Natl. Acad. Sci. USA **82:** 1396–1400.

GARABEDIAN, M. J., B. M. SHEPHERD and P. C. WENSINK, 1986 A tissue-specific transcription enhancer from the Drosophila yolk protein 1 gene. Cell **45:** 859–867.

GILLIES, S. D., S. L. MORRISON, V. T. OI and S. TONEGAWA, 1983 A tissue-specific transcription enhancer element is located in the major intron of a rearranged immunoglobulin heavy chain gene. Cell **33:** 717–728.

GOLDBERG, D. A., J. W. POSAKONY and T. MANIATIS, 1983 Correct developmental expression of a cloned alcohol dehydrogenase gene transduced into the Drosophila germ line. Cell **34:** 59–73.

GOLDBERG, M., 1979 Ph.D. thesis, Stanford University, Stanford, Calif.

GORMAN, C. M., L. F. MOFFAT and B. H. HOWARD, 1982 Recombinant genomes which express chloramphenicol acetyltransferase in mammalian cells. Mol. Cell. Biol. **2:** 1044–1051.

HAFEN, E., M. LEVINE, R. L. GARBER and W. J. GEHRING, 1983 An improved in situ hybridization method for detection of cellular RNAs in Drosophila tissue sections and its application for localizing transcripts of the homeotic gene *Antennapedia* gene complex. EMBO J. **2:** 617–623.

HAMMER, R. E., R. KRUMLAUF, S. A. CAMPER, R. L. BRINSTER and S. M. TILGHMAN, 1987 Diversity of alpha-fetoprotein gene expression in mice is generated by a combination of separate enhancer elements. Science **235:** 53–58.

HARRIS, W. A., W. S. STARK and J. A. WALKER, 1976 Genetic dissection of the photoreceptor system in the compound eye of *Drosophila melanogaster.* J. Physiol. **256:** 415–439.

HAZELRIGG, T., R. LEVIS and G. M. RUBIN, 1984 Transformation of *white* locus DNA in Drosophila: dosage compensation, *zeste* interaction, and position effects. Cell **36:** 469–481.

HEBERLEIN, U., B. ENGLAND and R. TJIAN, 1985 Characterization of Drosophila transcription factors that activate the tandem promoters of the alcohol dehydrogenase gene. Cell **41:** 965–977.

HIROMI, Y., A. KUROIWA and W. J. GEHRING, 1985 Control elements of the Drosophila segmentation gene *fushi tarazu.* Cell **43:** 603–613.

HOTTA, Y. and Z. C. KENG, 1984 Genetic dissection of larval photoreceptors in Drosophila. pp. 49–60. In: *Animal Behavior, Neurophysiological and Ethological Approaches.* Edited by K. AOKI, S. ISHII and H. MORITA. Springer-Verlag, Berlin.

JOWETT, T., 1985 The regulatory domain of a larval serum protein gene in *Drosophila melanogaster.* EMBO J. **4:** 3789–3795.

KALFAYAN, L., B. WAKIMOTO and A. C. SPRADLING, 1984 Analysis of transcriptional regulation of the s38 chorion gene of Drosophila by P element-mediated transformation. J. Embryol. Exp. Morphol. **83** (Suppl.)**:** 137–146.

KARESS, R. E. and G. M. RUBIN, 1984 Analysis of P transposable element functions in Drosophila. Cell **38:** 135–146.

KRUMM, A., G. E. ROTH and G. KORGE, 1985 Transformation of salivary gland secretion protein gene Sgs-4 in Drosophila: stage- and tissue-specific regulation, dosage compensation, and position effect. Proc. Natl. Acad. Sci. USA **82:** 5055–5059.

LANGER-SOFER, P. R., M. LEVINE and D. C. WARD, 1982 Immunological method for mapping genes on Drosophila polytene chromosomes. Proc. Natl. Acad. Sci. USA **79:** 4381–4385.

LAWRENCE, P. A. and S. M. GREEN, 1979 Cell lineage in the developing retina of Drosophila. Dev. Biol. **71:** 142–152.

LEVIS, R., T. HAZELRIGG and G. M. RUBIN, 1985 Separable *cis*-acting control elements for expression of the *white* gene of Drosophila. EMBO J. **4:** 3489–3499.

MANIATIS, T., E. F. FRISCH and J. SAMBROOK, 1982 *Molecular Cloning, A Laboratory Manual.* Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.

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.

MAXAM, A. M. and W. GILBERT, 1977 A new method for sequencing DNA. Proc. Natl. Acad. Sci. USA **74:** 560–564

MELGAR, E. and D. A. GOLDTHWAITE, 1968 Deoxyribonucleic acid nucleases. II. The effects of methals on the mechanism of action of deoxyribonuclease I. J. Biol. Chem. **243:** 4409–4416.

MESTRIL, R., P. SCHILLER, J. AMIN, H. KLAPPER, J. ANANTHAN and R. VOELLMY, 1986 Heatshock and ecdysone activation of the *Drosophila melanogaster* hsp23 gene; a sequence element implied in developmental regulation. EMBO J. **5:** 1667–1673.

MONTELL, C., K. JONES, C. S. ZUKER and G. M. RUBIN, 1987 A second opsin gene expressed in the ultraviolet sensitive R7 photoreceptor cells of *Drosophila melanogaster.* J. Neurosci. **7:** 1558–1566.

NATHANS, J. and D. S. HOGNESS, 1983 Isolation, sequence analysis, and intron-exon arrangement of the gene encoding bovine rhodopsin. Cell **34:** 807–814.

NATHANS, J. and D. S. HOGNESS, 1984 Isolation and nucleotide sequence of the gene encoding human rhodopsin. Proc. Natl. Acad. Sci. USA **81:** 4851–4855.

NATHANS, J., D. THOMAS and D. S. HOGNESS, 1986 Molecular genetics of human color vision: the genes encoding blue, green and red pigments. Science **232:** 193–202.

NEUBERGER, M. S., 1983 Expression and regulation of immunoglobulin heavy chain gene transfected into lymphoid cells. EMBO J. **2:** 1373–1378.

OSTROY, S. E., M. WILSON and W. L. PAK, 1974 Drosophila rhodopsin: photochemistry, extraction and differences in the *norpA*^{P12} phototransduction mutant. Biochem. Biophys. Res. Commun. **59:** 960–966.

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 **40:** 839–850.

OVCHINNIKOV, Y. A., 1982 Rhodopsin and bacteriorhodopsin: structure-function relationships. FEBS Lett. **148:** 179–189.

PICARD, D. and W. SCHAFFNER, 1984 A lymphocyte-specific enhancer in the mouse immunoglobulin k gene. Nature **307:** 80–82.

PIRROTTA, V., H. STELLER and M. P. BOZZETTI, 1985 Multiple upstream regulatory elements control the expression of the *Drosophila white* gene. EMBO J. **4:** 3501–3508.

POSAKONY, J. W., J. A. FISCHER and T. MANIATIS, 1985 Identification of DNA sequences required for the regulation of Drosophila alcohol dehydrogenase gene expression. Cold Spring Harbor Symp. Quant. Biol. **50:** 515–520.

QUEEN, C. and D. BALTIMORE, 1983 Immunoglobulin gene transcription is activated by downstream sequence elements. Cell **33:** 741–748.

READY, D. F., T. E. HANSON and S. BENZER, 1976 Development of the Drosophila retina, a neurocrystalline lattice. Dev. Biol. **53:** 217–240.

RIDDIHOUGH, G. and H. R. B. PELHAM, 1986 Activation of the Drosophila hsp27 promoter by heat shock and by ecdysone involves independent and remote regulatory elements. EMBO J. **5:** 1653–1658.

RUBIN, G. M. and A. C. SPRADLING, 1982 Genetic transformation of Drosophila with transposable element vectors. Science **218:** 348–353.

RUBIN, G. M. and A. C. SPRADLING, 1983 Vectors for *P* element-mediated gene transfer in Drosophila. Nucleic Acids Res. **18:** 6341–6351.

Sanger, F., S. Nicklen and A. R. COULSON, 1977 DNA sequencing with chain terminating inhibitors. Proc. Natl. Acad. Sci. USA **74:** 5463–5467.

SCHOLNICK, S. B., B. A. MORGAN and J. HIRSH, 1983 The cloned dopa decarboxylase gene is developmentally regulated when reintegrated into the Drosophila genome. Cell **34:** 37–45.

SCHOLNICK, S. B., S. J. BRAY, B. A. MORGAN, C. A. MCCORMICK and J. HIRSH, 1986 CNS and hypoderm regulatory elements of the *Drosophila melanogaster* dopa decarboxylase gene. Science **234:** 998–1002.

SEIFERT, H. S., E. Y. CHEN, M. SO and F. HEFFRON, 1986 Shuttle mutagenesis: a method of transposon mutagenesis for *Saccharomyces cerevisiae*. Proc. Natl. Acad. Sci. USA **83:** 735–739.

SERFLING, E., M. JASIN and W. SCHAFFNER, 1985 Enhancers and eukaryotic gene transcription. Trends Genet. **1:** 224–230.

SHERMOEN, A. W., J. JONGENS, S. W. BARNETT, K. FLYNN and S. K. BECKENDORF, 1987 Developmental regulation by an enhancer from the Sgs-4 gene of Drosophila. EMBO J. **6:** 201–214.

SPRADLING, A. C. and G. M. RUBIN, 1982 Transposition of cloned P elements into Drosophila germ line chromosomes. Science **218:** 341–347.

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.

STELLER, H. and V. PIRROTTA, 1985 A transposable P vector that confers selectable G418 resistance to Drosophila larvae. EMBO J **4:** 167–171.

STELLER, H. and V. PIRROTTA, 1986 P transposons controlled by the heat shock promoter. Mol. Cell. Biol. **6:** 1640–1649.

TOMLINSON, A., 1985 The cellular dynamics of pattern formation in the eye of Drosophila. J. Embryol. Exp. Morphol. **89:** 313–331.

TOMLINSON, A. and D. F. READY, 1987 Neuronal differentiation in the Drosophila ommatidium. Dev. Biol. **120:** 366–376.

ZUKER, C. S., A. F. COWMAN and G. M. RUBIN, 1985 Isolation and structure of a rhodopsin gene from *D. melanogaster*. Cell **40:** 851–858.

ZUKER, C. S., C. MONTELL, K. R. JONES, T. LAVERTY and G. M. RUBIN, 1987 A rhodopsin gene expressed in photoreceptor cell R7 of the Drosophila eye: homologies with other signal transducing molecules. J. Neurosci. **7:** 1550–1557.

Communicating editor: A. SPRADLING