MOLECULAR AND CELLULAR BIOLOGY, May 1986, p. 1464–1470
0270-7306/86/051464-07$02.00/0

Vol. 6, No. 5

The Memory Gene Dunce$^+$ Encodes a Remarkable Set of RNAs with Internal Heterogeneity

RONALD L. DAVIS[1]* AND NORMAN DAVIDSON[2]

Department of Biochemistry, Michigan State University, East Lansing, Michigan 48824,[1] and Department of Chemistry, California Institute of Technology, Pasadena, California 91125[2]

Received 4 October 1985/Accepted 22 January 1986

We have previously isolated the region of the *Drosophila melanogaster* X chromosome which contains the dunce$^+$ gene and mapped the dunce^2 mutation by recombination to a 10- to 12-kilobase (kb) interval (R. L. Davis and N. Davidson, Mol. Cell. Biol. 4:358–367, 1984). Here, we examine the expression of the dunce$^+$ chromosomal region and identify the dunce$^+$ gene within that region. A region of ca. 25 kb which contains the 10- to 12-kb interval to which dunce^2 was mapped codes for polyadenylated RNAs of 9.6, 7.4, 7.2, 7.0, 5.4, and 4.5 kb in adult flies. These transcripts are encoded by the same DNA strand and share sequences of some exons, indicating that the transcripts arise from the same gene. Some genome probes internal to the ca. 25-kb coding region show transcript-specific hybridization, demonstrating alternate usage of exonic sequence information in the formation of the mature transcripts. The basis for this internal heterogeneity in RNAs is most likely alternative splicing. Two dunce mutants examined show aberrant RNA expression from this coding region, confirming that this region is the dunce gene. The developmental expression of these transcripts has been examined. The 5.4-kb RNA is present at all developmental stages. The 9.6-, 7.4-, 7.2-, and 7.0-kb RNAs are not expressed at detectable levels in embryos, but are detected in late embryogenesis and in later developmental stages. The 4.5-kb species is found in early embryos and adults, but not in intermediate stages. We discuss the remarkable transcript heterogeneity and expression pattern with respect to the important function this gene performs in neurobiological and other physiological processes.

We have commenced the molecular analysis of the dunce (*dnc*) gene of *Drosophila melanogaster* to study and interrelate aspects of its genetics and molecular biology and its role in insect biochemistry and neurobiology. Mutations at this locus lead to a variety of behavioral defects, disrupt normal cyclic AMP (cAMP) metabolism, cause sterility in adult females, and alter the expression level of some other genes.

The *dnc* mutant flies are best classed as memory mutants since the principal behavioral phenotypes are abbreviated memories of conditioned behaviors (1). For example, these flies perform normally or almost normally immediately after training in an associative learning situation which couples olfactory cues to the negative reinforcement of electrical shock, but fail in these tests a few minutes after training (10). In addition, their memory of the nonassociative learning behavior sensitization is short lived (11). Other behaviors are abnormal in these flies, such as those of operant conditioning (2) and certain aspects of normal courtship behavior (15). These studies demonstrate that *dnc* performs a critical role in higher behavioral processes, including conditioned behavior.

The observation that *dnc* mutations perturb cAMP metabolism (4, 8) has been interpreted as an indication that cAMP metabolism and effector functions are intimately involved in learning and memory processes (1, 7). One form of cAMP phosphodiesterase, the cellular enzyme which degrades cAMP, is reduced in activity or missing in *dnc* flies, depending on the strength of the particular allele (8). As a result, cAMP levels are increased in *dnc* flies to a surprising 6 times over normal in the most extreme *dnc* mutants (8). The idea that cAMP metabolism is coupled directly to learning and memory processes and is not just an incidental connection in *dnc* mutants is supported by behavioral and biochemical studies of *Aplysia californica* (17) and from the biochemical analysis of other *Drosophila* learning-memory mutants (20). These studies also demonstrate that cAMP metabolism is intimately involved in learning and memory.

Several genes have been isolated which are expressed at altered abundance levels in *dnc* mutants relative to wild-type flies (Y. Yun and R. L. Davis, unpublished data). Some are expressed at increased levels, and some are expressed at decreased levels. The specific functions of these genes are unknown, but the altered gene expression is presumably related to the changes in cAMP metabolism and possibly related to fly behavior.

Using the methodology of chromosomal walking, we have isolated a large region of the X chromosome which contains *dnc*$^+$ and delimited the gene on this cloned DNA by using two different methods (6). First, we mapped the breakpoints of chromosomal aberrations known to break near *dnc*. Since the genetic relationship of these breakpoints and the *dnc* gene had been previously established, the data delimited *dnc* to a 50-kilobase (kb) region. In addition, we mapped the *dnc*^{2} lesion to a 10- to 12-kb interval by recombination, using restriction site polymorphisms as genetic markers.

In this paper, we examine the RNA expression of the *dnc*$^+$ chromosomal region. We find a single gene expressed at detectable levels; this gene spans the 10- to 12-kb region to which the *dnc*^{2} mutation was mapped. It codes for multiform transcripts, which are internally heterogeneous and are expressed during development in a complex manner. We present additional evidence indicating that this complex gene is *dnc*.

MATERIALS AND METHODS

RNA isolation. Total RNA was isolated by homogenizing animals in 4 M guanidine thiocyanate–1 M 2-mercaptoethanol–0.05 M sodium acetate (pH 4.5)–0.001 M

* Corresponding author.

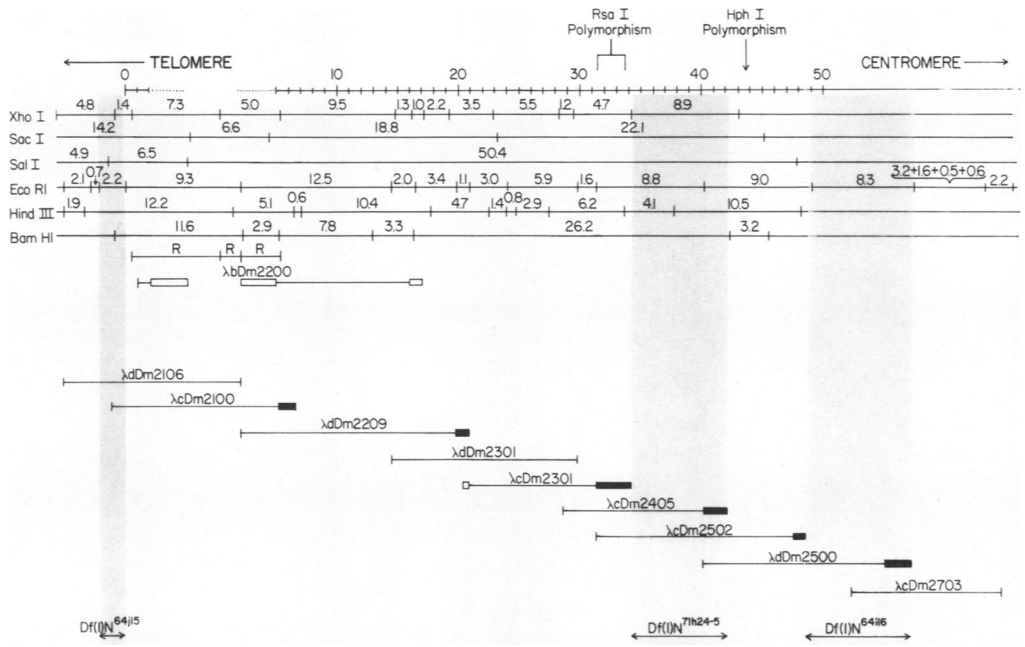

FIG. 1. *dnc*+ chromosomal region. Superimposed on the restriction map of the *dnc*+ chromosomal region are the locations of breakpoints associated with three different chromosomal aberrations. These are indicated by shaded areas, the width of each indicating the region to which the breakpoints have been defined. The extents of various λ clones are shown as horizontal line segments below the restriction map. Darkened blocks represent probes used during each successive step of the chromosomal walk (6). Open boxes represent uncertainties in the length of cloned DNA segments. All clones were isolated from Canton S genome libraries except λbDm2200, which was selected from an Oregon R genome library. This clone does not carry repetitive sequences carried by Canton S-derived clones in the same vicinity. The restriction fragments (R) have been shown to carry the repeated sequences. The coordinate system (in kb) used for analyzing restriction site polymorphisms is shown above the map. Not counted in the coordinate system is the 7.3-kb repetitive insertion element found in Canton S clones but not Oregon R clones. The insertion site of this element is within coordinates 2 and 5.

EDTA. The homogenate was cleared by centrifugation at 8,000 × *g* for 20 min, and RNA in the supernatant was banded in cesium chloride as described by Fyrberg et al. (13).

Polyadenylated [poly(A)+] RNA was selected by one pass over an oligo(dT)-cellulose column.

RNA blots and hybridizations. Poly(A)+ RNA was fractionated on formaldehyde-agarose gels, blotted onto nitrocellulose, and hybridized to nick-translated probes or single-stranded probes by the procedures of Rozek and Davidson (23) and Davis and Davidson (6). Single-stranded probes were prepared as described previously (16). Size markers for RNA blots were *Escherichia coli* rRNA and *Drosophila* myosin heavy-chain RNA (23).

Other procedures. Synthetic deficiencies of the *dnc* region were constructed by mating males of the genotype $Df(1)N^{64i16}$; *SM1*, $CyDp(1;2)w^{+51b7}$ to *C(1)DX*, *yw*f*l*$w^{+}Y$ females (19). Males of the genotype $Df(1)N^{64i16}/w^{+}Y$ were selected and used as a source of RNA.

RESULTS

***dnc*+ chromosomal region.** The restriction map and other important features of the *dnc*+ chromosomal region are shown in Fig. 1. The right breakpoint of $Df(1)N^{64j15}$, positioned on the restriction map in the figure, resides genetically to the left of *dnc*+ and marks the left boundary of what we term the *dnc*+ chromosomal region. The right side of this region is marked by the breakpoints of $Df(1)N^{71h24\text{-}5}$ and $Df(1)N^{64i16}$. These deficiencies remove sequences to the left of the breakpoints shown and remove *dnc*+ function. Therefore, these breakpoints must be within the gene or to its right. One repetitive sequence is located near the left margin of the *dnc*+ chromosomal region in Canton S flies; the rest of the DNA in the region is a unique sequence as determined by genome blotting experiments.

The polymorphisms for the restriction sites of *Rsa*I and *Hph*I detected previously between Canton S and Amherst wild-type strains have been used as genetic markers and are known to flank the dnc^{2} mutation on the left and right, respectively (6). This mutation is a marker for a protein-coding portion of the gene, since it affects the kinetics of cAMP hydrolysis by phosphodiesterase (7, 18). Thus, at least part of the coding region for a *dnc* gene product must reside in the 10- to 12-kb interval defined by these polymorphisms.

A 25-kb region containing *dnc* comprises a single gene which codes for a complex set of transcripts. The locations of the *Rsa*I and *Hph*I restriction site polymorphisms which bracket the dnc^{2} lesion are shown in Fig. 1. These reside at coordinates 32 to 34 and 44, respectively. We have prepared subclones spanning this region from coordinates 15 to 57 to use as probes for RNA blots (Fig. 2) to define coding regions. The results presented below indicate that a set of RNA transcripts of molecular lengths 9.6, 5.4, and 4.5 kb and several of about 7.2 kb are all derived from a gene extending across the region between coordinates 21 and 46 in Fig. 2.

The definition of a gene becomes ambiguous when overlapping transcripts are encountered. For this situation, we define a gene as that genome region which encodes the overlapping transcripts and add the criteria that the transcripts be encoded by the same strand of DNA and that they share sequence information from one or more exons. This definition makes no statement about the number of promoters used by a given gene. Therefore, it is to be distinguished from that for the term transcription unit, which has been

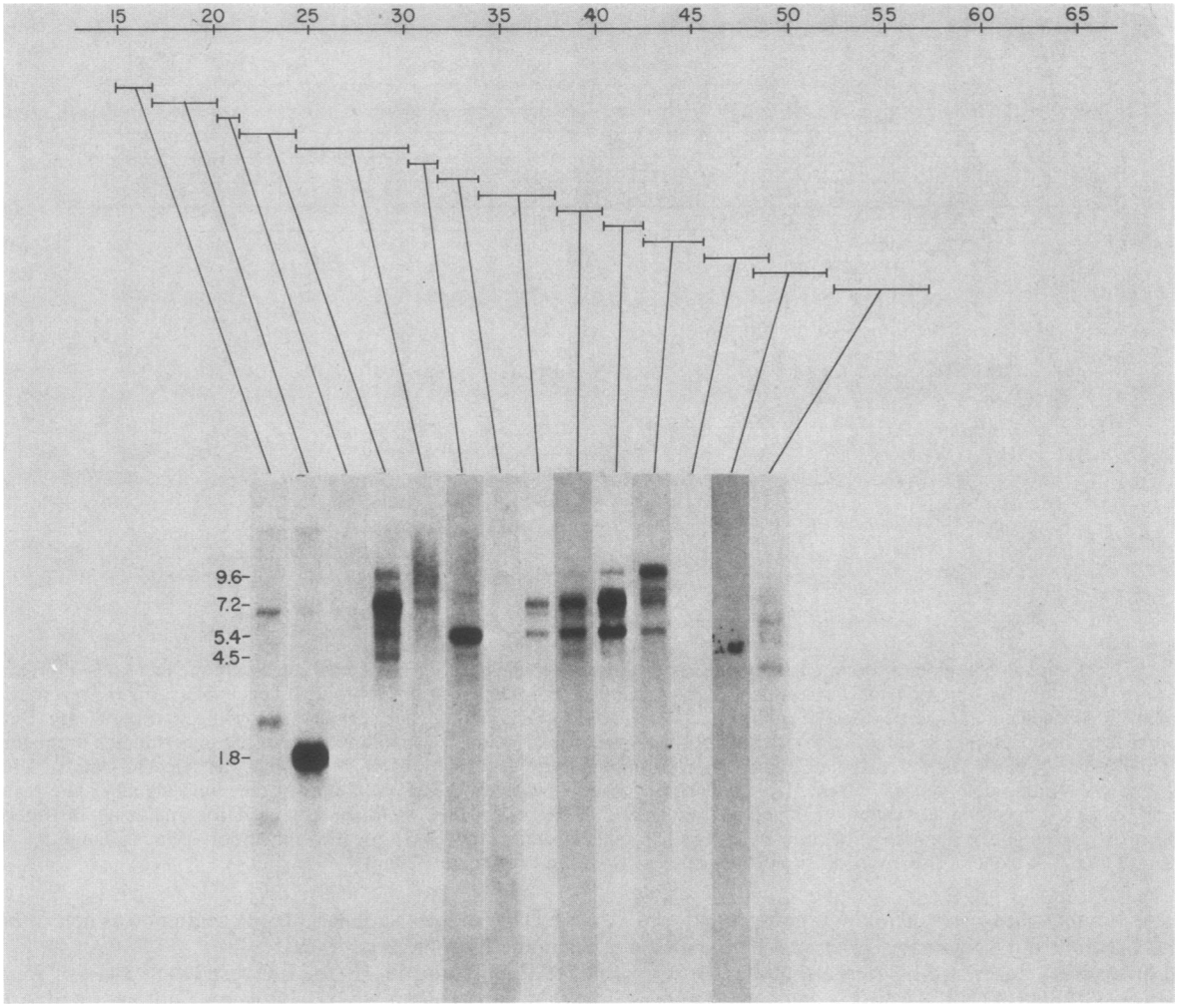

FIG. 2. RNA blots probed with *dnc*$^+$ region subclones. Subclones of the region are indicated by horizontal line segments. These were nick translated and used to probe 10 μg of adult poly(A)$^+$ RNA fractionated by size. The size of genome DNA in the subclones ranges from 1.1 to 5.9 kb. The intensity of the signal reflects both the abundance of the RNAs detected and the extent of sequence homology of any particular RNA species to the probe. The autoradiogram exposure times (from left to right) were 5, 5, 5, 5, 5, 2.5, 5, 1, 0.16, 0.16, 1, 2, 2, and 5 days. Some faint signals did not reproduce well.

defined by Darnell (5) as the genome region between a promoter and the transcription stop site.

All of the subclones from coordinates 21 to 46 which display homology to adult poly(A)$^+$ RNA hybridize to RNAs of 9.6, 5.4, and 4.5 kb and several of about 7.2 kb, or a subset of these. For example, a subclone spanning the region of coordinates 21 to 24 (herein referred to as subclone 21–24) hybridizes to the 9.6- and 7.2-kb broad band and the 5.4- and 4.5-kb regions. This same hybridization pattern is observed with a subclone at coordinates 41 to 43 (subclone 41–43). These observations suggest that a single gene codes for these RNAs and extends over this region of the genome. This suggestion has been verified by demonstrating that the RNAs to which probe 21–24 will hybridize are identical to those detected by probe 41–43. Poly(A)$^+$ RNAs homologous to probe 21–24 were purified from total poly(A)$^+$ RNA by hybridization-selection. These RNAs were fractionated by gel electrophoresis, blotted, and then probed with subclone 41–43. This probe detects RNAs selected by subclone 21–24, and these RNAs are the same as those detected by subclone 41–43 in the unfractionated RNA population (Fig. 3). It is not likely that the hybridization observed in Fig. 3B is due to total poly(A)$^+$ RNA contaminating the RNA obtained by hybridization-selection. The procedures afford a purification of about 10,000-fold (9, 27), and we cannot visualized *dnc*$^+$ RNAs on RNA blots with less than 1 μg of total poly(A)$^+$ RNA.

Subclones 34–38 and 43–46 appear to distinguish the broad 7.2-kb region of hybridization into 3 species. RNAs of 7.2 and 7.0 kb are detected by subclone 34–38, and 7.4- and 7.2-k RNAs are detected by subclone 43–46. This suggests that the gene extends from coordinates ca. 21 to 46 and encodes species of 9.6, 7.4, 7.2, 7.0, 5.4, and 4.5 bk.

We have used single-stranded probes to establish the direction in which the multiple transcripts are synthesized. The genomic region at coordinates 41 to 43 (an *Eco*RI-*Bam*HI fragment) was directionally cloned into the vectors M13mp8 and M13mp9, and single-stranded probes were prepared (16). Both strands were used to probe RNA blots. The blot shown in Fig. 4A shows that one strand hybridizes to RNA regions of 9.6, 7.4 to 7.0, 5.4, and 4.5 kb, whereas the other strand (Fig. 4B) does not hybridize to adult RNA. This demonstrates that the RNAs identified in Fig. 4 are all encoded by the same strand, and that the direction of transcription (5′ to 3′) is from left to right in Fig. 1 and 2.

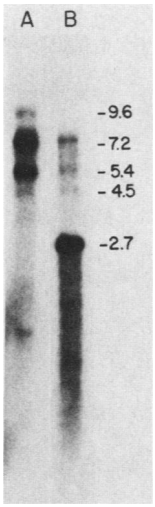

FIG. 3. Hybridization-selection of *dnc*$^+$ RNAs. Lane A is a blot of 10 μg of adult poly(A)$^+$ RNA probed with clone 41–43. Lane B is RNA selected from 50 μg of poly(A)$^+$ RNA by hybridization to probe 21–24, fractionated, blotted, and probed with cloned 41–43. The intense band at 2.7 kb is plasmid DNA (pUC8) from the linearized clone 21–24, which was removed from the filter during elution of the bound RNA and which hybridized to the plasmid DNA in the probe 41–43.

A single-stranded probe of a 0.4-kb cDNA clone recently recovered which defines portions of two exons (unpublished data) was hybridized to adult RNA (Fig. 4C), and all of the previously identified RNAs hybridize. This indicates that the transcripts of 9.6, 7.4 to 7.0, 5.4, and 4.5 kb share sequences represented by one or both of the two exons defined by this cDNA clone. Thus, the transcripts of 9.6, 7.4, 7.2, 7.0, 5.4, and 4.5 kb are synthesized by the genome region at coordinates 21 to 46. These transcripts are coded for by the same DNA strand, and they share exonic sequence information. Therefore, this region represents a single gene, according to the criteria outlined above. Since this gene spans the region to which *dnc*^{2} was mapped, we conclude that it is *dnc*$^+$. Further support for this is presented below.

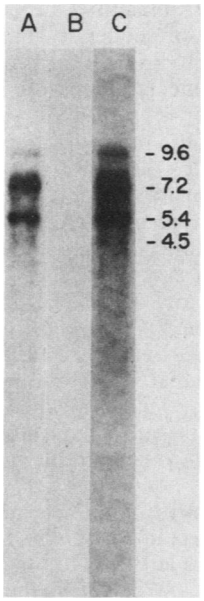

FIG. 4. Hybridization of adult RNA to single-stranded probes from the *dnc*$^+$ region.

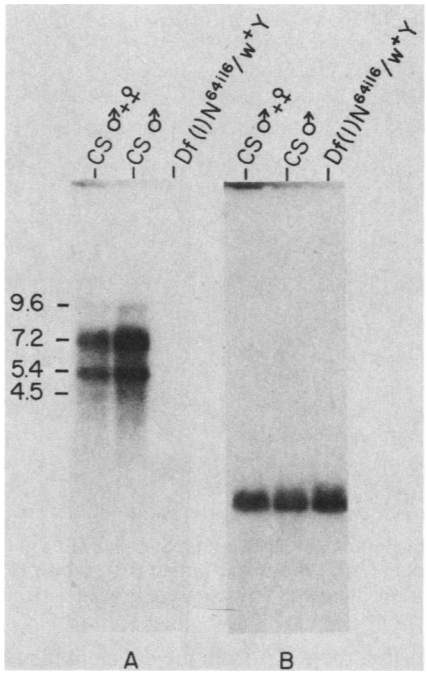

FIG. 5. RNAs homologous to *dnc*$^+$ in *dnc* deficient flies. Samples (10 μg) of poly(A)$^+$ RNA from *Df(1)N*^{64i16}*/w*$^+$*Y* males and Canton S (CS) flies were visualized by RNA blotting. The blot shown in A was reprobed in B with an actin 87E probe (14).

We have examined the possibility that some of the RNAs homologous to *dnc*$^+$ region probes may be the products of loci related to *dnc*$^+$ at the DNA sequence level, even though genome blotting experiments at normal stringency suggest that the *dnc*$^+$ gene is unique by this criterion. Flies that are homozygous deficient for a 86- to 100-kb region including the *dnc*$^+$ gene were prepared; RNA was extracted, and the RNAs homologous to a 2.1-kb cDNA clone from the *dnc*$^+$ gene were visualized by RNA blotting experiments. The RNAs detected with the cDNA probe in Canton S flies are not present at detectable levels in flies deficient for the *dnc*$^+$ locus (Fig. 5). As a control, the blot shown in Fig. 5 was reprobed with a *Drosophila* actin gene probe. The control

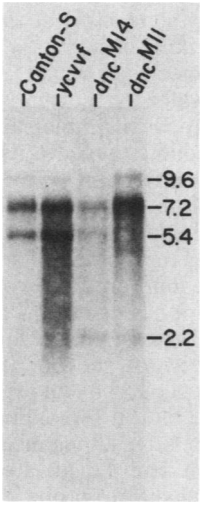

FIG. 6. RNAs in *dnc* mutants. Samples (6 to 8 μg) of poly(A)$^+$ RNA from adult organisms of each strain indicated were used for RNA blots and probed with clone 41–43. The 4.5-kb RNA is only visible on very long exposures.

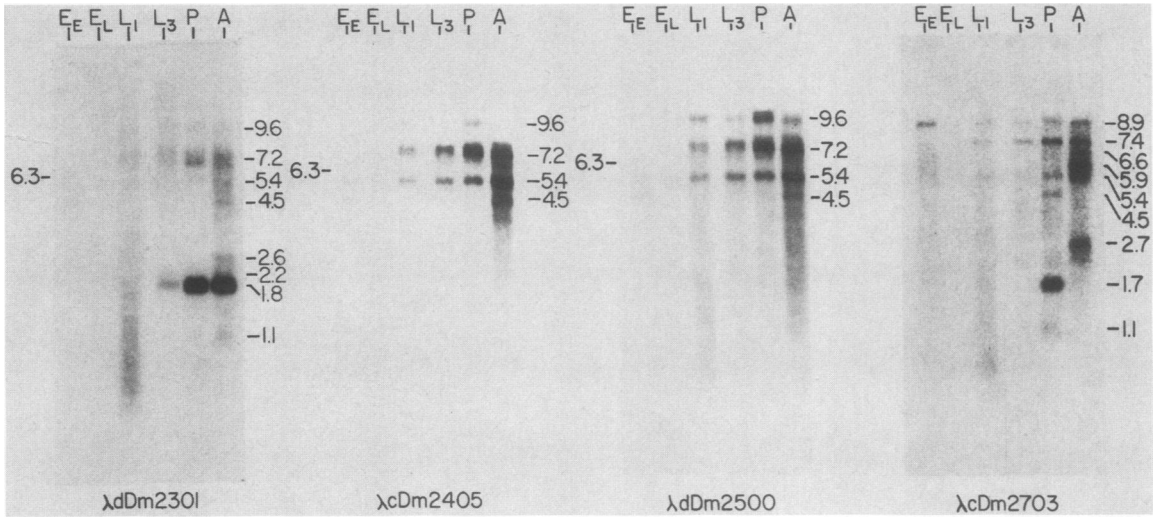

FIG. 7. Developmental RNA blots. Samples (10 μg) of poly(A)$^+$ RNA from early embryos (E_E, 0 to 4 h), late embryos (E_L, 16 to 20 h), first-instar larvae (L_1), third-instar larvae (L_3), pupae (P), and adults (A) were fractionated by denaturing agarose gel electrophoresis and blotted onto nitrocellulose. The blots were probed with nick-translated λ clones (Fig. 1). Some bands are very faint, i.e., the 6.3-kb transcript.

demonstrates that the RNA from the deficient flies was intact and present at levels approximating that in the Canton S lanes. This observation strongly argues against the possibility that the RNAs detected with *dnc*$^+$ gene probes are encoded by related genes elsewhere in the genome.

RNA expression in *dnc* mutants. We have examined the RNAs complementary to region 40 to 42 in two *dnc* mutants (Fig. 6). The *dnc*$^+$ strains Canton S and *ycvvf* both display RNA hybridization in the regions of 9.6 (faint), 7.4 to 7.0, and 5.4 (and 4.5 on very long exposures) kb. The mutants *dnc*^{M14} and *dnc*^{M11} exhibit alterations in this expression pattern. The *dnc*^{M14} mutant shows a decrease in the level of the RNAs detected. The same blot has been probed with a *Drosophila* clone representing a gene expressed at equal levels in *ycvvf*, *dnc*^{M14}, and *dnc*^{M11} strains (Yun and Davis, unpublished data). Densitometry of the signal displayed in this control blot demonstrates that the *dnc*^{M14} lane actually contains 28% more poly(A)$^+$ RNA in the sample loaded than does the *ycvvf* lane. Thus, the decreased signals in the *dnc*^{M14} lane probed with subclone 40–42 represent an actual decrease in the steady-state level of *dnc* RNA molecules. However, the most striking observation is that both mutants produce a new RNA species of 2.2 kb. An understanding of how this RNA is produced is currently not available, but these observations add considerable support to our conclusion that the gene encoding these RNAs is *dnc*.

Evidence for internal heterogeneity in *dnc*$^+$ transcripts. The intensities of the signals in Fig. 2 are dependent upon several parameters, the most important being the abundance of any given RNA species and the amount of sequence homology a given probe has to the different RNAs. From ongoing work to define the location and sizes of exons comprising the gene, we know that probe 21–24 contains a single detectable exon of 175 base pairs (data not shown). All of the heterogeneously sized RNAs encoded by the region 21–46 use this exon, as shown by the blot in Fig. 2 probed with the 21–24 clone. Therefore, the relative abundance levels of the RNAs of 9.6, 7.4 to 7.0, 5.4, and 4.5 kb are indicated by their relative hybridization signals to probe 21–24. Any deviation in these relative signal intensities with other probes must reflect differences in the amount of RNA-coding sequences in these particular probes.

Inspection of the signals in Fig. 2 for coordinates 21 to 46, in this light, reveals an interesting and complex picture. Note that a probe at coordinates 30 to 32 strongly detects the 5.4-kb species and gives faint hybridization to the 4.5- and 7.2-kb regions. This differential hybridization intensity, when compared with that obtained with probe 21–24, indicates that more genomic sequences within probe 30–32 are used to construct the 5.4-kb message than the 4.5-, 7.4- to 7.0-, or 9.6-kb regions. Thus, internal heterogeneity exists among the transcripts. Additional internal heterogeneity is reflected with probe 34–38. This probe hybridizes detectably to RNAs of 7.2, 7.0, and 5.4 kb. The intensity of the signal to the 5.4-kb species is close to that with the 21–24 probe, yet there is a much weaker signal from the 34–38 probe to the 7.2- and 9.6-kb RNAs. It seems likely that observed internal heterogeneity is produced by alternative splicing of precursor RNA molecules (see Discussion).

One region apparently devoid of homology to RNAs is probe 32–34; this is presumably a large intron. Probe 24–31 reveals a low amount of sequence homology to the 7.2-kb RNA and a smear of hybridization from about 8 to 25 kb. Possibly, the smear reflects processing intermediates which are relatively stable.

The transcript-specific hybridization displayed by the subclone at coordinates 30 to 32 argues against one less interesting explanation for the multiple RNAs, that being that the larger RNAs are precursors for the smallest transcript. Since the 5.4-kb RNA contains sequences not found in the larger RNAs, it cannot be a processing intermediate derived from the 7.4- to 7.0-kb RNA or the 9.6-kb RNA or both. Otherwise, probe 30–32 would hybridize to the larger RNAs. It is improbable that the multiple RNAs represent products of the in vivo RNA(s), produced by degradation during RNA isolation procedures, since the RNAs shown in Fig. 2 have been observed in five different RNA preparations. It is also difficult to explain the internal heterogeneity from this viewpoint.

Preliminary Assignment of the *dnc* gene boundaries. The clone furthest to the left in Fig. 2 which hybridizes to *dnc*$^+$ RNAs is 21–24. To the left of coordinate 21 exists a 1.8-kb RNA coded by region 17–20. This is the RNA detected with probe 17–20. We presume that the existence of this RNA

marks the left limit to the *dnc*$^+$ coding region. The clone furthest to the right which clearly hybridizes to *dnc*$^+$ transcripts is 43–46. To the right of coordinate 46 exists a 6-kb region devoid of any detectable coding sequences, and beyond this are coding sequences for very rare transcripts within coordinates 53 to 57. The sizes of some of the faint RNAs visualized with probe 53–57 in Fig. 2 are different from those detected by subclones within coordinates 21 to 46. We believe, therefore, that the right end of the *dnc*$^+$ coding region for RNAs of 9.6, 7.4 to 7.0, 5.4, and 4.5 kb is probably within coordinates 43 to 46.

These assignments are the simplest provided by the data in Fig. 2. If the coding region for the multiple *dnc*$^+$ RNAs extends beyond coordinates 21 to 46, then the existence of other detectable RNAs, like that within coordinates 17 to 20, requires a more novel explanation, such as being located within a *dnc*$^+$ intron.

Developmental expression of *dnc*$^+$ region RNAs. We are interested in how *dnc*$^+$ transcripts are expressed during development, and we have therefore probed RNA populations from several different stages during the *Drosophila* life cycle. Figure 7 displays developmental RNA blots probed with several different lambda clones from the *dnc*$^+$ chromosomal region. The developmental pattern of RNA expression from this region is complex and not fully understood at present. Nevertheless, certain themes are established from these blots and are emphasized here.

The clones λdDm2301, λcDm2405, and λdDm2500 all hybridize to *dnc*$^+$ expressed poly(A)$^+$ RNA regions of 9.6, 7.4 to 7.0, 5.4, and 4.5 kb in adult flies. The probe λdDm2301 exhibits limited sequence homology to these RNAs indicated by the faint hybridization. Inspection of the two middle panels in Fig. 7 reveals that the 5.4-kb species is expressed in all stages examined, although its abundance increases with development. The 4.5-kb RNA is clearly present in early embryos and adults, but is not detected in RNA populations from intermediate stages. The hybridization intensity and complexity around the 9.6- and 7.4- to 7.0-kb regions increases during development, being first detected with RNA from late embryos. There are probably at least two RNAs detected of ca. 9.6 kb; a doublet in this region is suggested by the hybridization to pupal RNA with λdDm2500. Early embryos also express an RNA of 6.3 kb that is not expressed in later stages (except possibly in adults). These additional RNAs are possibly transcripts from the *dnc*$^+$ gene, since λcDm2405 resides within coordinates 21 to 46, although we do not know whether they share exon sequence information with and are transcribed in the same direction as the RNAs expressed in adults. Thus, *dnc*$^+$ expresses at least six poly(A)$^+$ RNAs, ranging in size from 4.5 to 9.6 kb, whose steady-state level during development is regulated in a complex fashion. The biological significance of this is discussed in more detail below.

The developmental expression of the 1.8-kb transcriptional unit to the left of *dnc*$^+$ is displayed by hybridization to λdDm2301. This transcript appears in the third larval instar stage and increases in abundance to adulthood. This expression pattern, which differs from that of any of the *dnc*$^+$-encoded transcripts, is consistent with our assignment of the left end of *dnc*$^+$ to the right of this RNA-coding region.

The RNAs detected by λcDm2703 present a more complex picture. Coding sequences for many transcripts expressed in pupae and adults are located within this clone. The sizes and developmental expression profile of many of the RNAs detected by λcDm2703 are different from the *dnc*$^+$ transcripts visualized with λcDm2405 or λdDm2500 as probes. For example, the 8.9-kb transcript detected with λcDm2703 is expressed in embryos, unlike the 9.6-kb RNA encoded by region 21 to 46. A 2.7-kb RNA in adults and a 1.7-kb RNA expressed in pupae are detected with λcDm2703. The existence of coding sequences for these additional RNAs not detected by probes from within coordinates 21 to 46 is consistent with our assignment of the right boundary of *dnc*$^+$ to coordinate 46.

DISCUSSION

A 100-kb segment of the *Drosophila* X chromosome was previously isolated by chromosomal walking, and some information about the location of the *dnc*$^+$ locus on the DNA was obtained by using recombinational mapping with restriction site polymorphisms as genetic markers (8). The data place the *dnc*^{2} mutation between two polymorphisms marking a 10- to 12-kb window. We have examined the expression of the *dnc*$^+$ chromosomal region to further define the *dnc*$^+$ gene on the cloned DNA and to understand the environment in which *dnc* is located. We find a single gene which spans a 25-kb region including the 10- to 12-kb window. This gene encodes an amazing spectrum of RNA molecules and is defined as such, because these RNAs are transcribed from the same DNA strand and because they share exon sequence information.

We argue that the gene which produces multiple RNAs in the *dnc*$^+$ chromosomal region corresponds to the *dnc*$^+$ gene for two reasons. First, this gene spans the region to which the *dnc*^{2} mutation was mapped by recombination. Since the *dnc*^{2} mutation alters the kinetics of cAMP hydrolysis by the phosphodiesterase (7, 8, 18), the mutation served as a good marker for the protein-coding portion of the gene. Using single-stranded probes for regions 41 to 43 (Fig. 4) and 21 to 24 (data not shown), we detect no RNAs homologous to the antisense strand. Second, two different *dnc* mutants alter the pattern of RNAs encoded from the gene. The most striking observation is that a new RNA species of 2.2 kb is produced in both mutants. The basis for this is not known, but presumably the *dnc* mutants are producing this RNA by altering the normal site(s) for the initiation of transcription, changing the poly(A)$^+$ site(s) utilized, or modifying the splicing patterns of the primary transcript(s). Most importantly, these observations add support to the idea that this complex gene is *dnc*$^+$.

The picture we gain of *dnc*$^+$ from these studies is one of great complexity, especially with respect to the gene's RNA-coding role. The gene produces no less than six different RNAs in adult flies. These transcripts are unusually large and rare (unpublished data) in the RNA population extracted from flies. Some of the size heterogeneity is internal; that is, some coding regions within the gene are used differentially among the final RNAs. Most likely, this internal heterogeneity is produced by alternative splicing of exons, although we cannot rule out the possibility that genome rearrangements in certain cell types produce gene structure heterogeneity, which is then reflected in transcript heterogeneity. Other mechanisms of alternative RNA processing can produce heterogeneity at the 5′ or 3′ ends of transcripts. It is possible that *dnc* also employs these other mechanisms to produce some of the observed heterogeneity. Indeed, the RNA blotting experiments (Fig. 2, probe 43–46) suggest that the larger transcripts use more coding sequences from region 43 to 46 to form their 3′ ends compared with the smaller RNAs. A final judgement of this awaits the precise definition of the promoter region(s) and the 3′ end(s) of the gene. Our ongoing work, which involves using S1

nuclease analysis and the sequence analysis of cDNA and genome clones, should reveal the structural features of the gene and resolve some of these unknowns.

The RNA transcript complexity and heterogeneity came as a surprise, since dnc^+ is thought to be the structural gene for the enzyme cAMP phosphodiesterase (7, 18). The significance of this RNA heterogeneity is unknown, but insights are gained from knowledge of RNA heterogeneity exhibited by other genes. RNA heterogeneity and its significance have been recently discussed by several authors (3, 12, 21, 22). Many genes which produce multiple, internally heterogenous transcripts produce multiple protein products via alternative use of protein-coding exons. It is conceivable that dnc^+ encodes multiple protein products in an analogous fashion. In many of the well-documented cases of protein product heterogeneity from a single gene, the products are isoforms of a single type of cellular function. This scenario would predict that dnc^+ produces microheterogeneity in the presumed product, cAMP phosphodiesterase. However, no microheterogeneity in this enzyme has been uncovered in the characterizations performed to date (7). It also seems probable that some of the RNA heterogeneity is the result of tissue specific or cell type-specific processing of the primary transcript(s). This is likely because many other genes which produce heterogeneous transcripts do so as a function of the tissue or cellular environment. This possibly reflects control mechanisms for governing the quantity of the gene product synthesized in different cells.

The developmental regulation of the dnc^+ transcripts adds another level of complexity to its expression. For example, the 4.5-kb transcript is present during early development, disappears during the larval instars and during pupation, and reappears in adulthood. The expression of cAMP phosphodiesterase has been examined in first- and third-instar larvae and in adult flies (26). The specific activity in both larval instars is comparable to that detected in adults. If the enzyme activity is a close function of the mRNA level, this rules out the 4.5-kb transcript as a sole phosphodiesterase RNA.

Within the dnc^+ chromosomal region, there exists one other genetic function which has been identified by genetic analysis. This is the *sam* locus, which when mutated causes male sterility (24). Mapping experiments have placed *sam* to the left of dnc^+ (25). The 1.8-kb transcript which appears during the latter part of development is therefore a candidate for this gene.

These studies portray a gene of unusual complexity. Although this complexity was initially a surprise, perhaps it is fitting, since the gene functions in the complex biochemical processes underlying conditioned behavior.

ACKNOWLEDGMENTS

We thank B. Bond and M. Snyder for RNA samples and T. Freidman and W. Mattox for their critical review of the manuscript.

This research was supported by Public Health Service grants to N.D. and to R.L.D. from the National Institutes of Health and by funds to R.L.D. from the McKnight Foundation.

LITERATURE CITED

1. **Aceves-Pina, E. O., R. Booker, J. S. Duerr, M. S. Livingstone, W. G. Quinn, R. F. Smith, P. P. Sziber, B. L. Tempel, and T. P. Tully.** 1984. Learning and memory in *Drosophila*, studied with mutants. Cold Spring Harbor Symp. Quant. Biol. **48**:831–840.
2. **Booker, R., and W. G. Quinn.** 1981. Conditioning of leg position in normal and mutant *Drosophila*. Proc. Natl. Acad. Sci. USA **78**:3940–3944.
3. **Breitbart, R. E., H. T. Ngnyen, R. M. Medford, A. T. Destree, V. Mahdavi, and B. Nadal-Ginard.** 1985. Intricate combinatorial patterns of exon splicing generate multiple regulated troponin T isoforms from a single gene. Cell **41**:67–82.
4. **Byers, D., R. L. Davis, and J. A. Kiger.** 1981. Defect in cyclic AMP phosphodiesterase due to the *dunce* mutation of learning in *Drosophila melanogaster*. Nature (London) **289**:79–81.
5. **Darnell, J. E.** 1979. Transcription units for mRNA production in eukaryotic cells and their DNA viruses. Prog. Nucleic Acid Res. **22**:327–353.
6. **Davis, R. L., and N. Davidson.** 1984. Isolation of the *Drosophila melanogaster dunce* chromosomal region and recombinational mapping of *dunce* sequences with restriction site polymorphisms as genetic markers. Mol. Cell. Biol. **4**:358–367.
7. **Davis, R. L., and L. Kauvar.** 1984. *Drosophila* cyclic nucleotide phosphodiesterases. Adv. Cyclic Nucleotide Res. **16**:393–402.
8. **Davis, R. L., and J. A. Kiger.** 1981. *dunce* mutants of *Drosophila melanogaster*: mutants defective in the cyclic AMP phosphodiesterase enzyme system. J. Cell. Biol. **90**:101–107.
9. **Dawid, I. B.** 1977. DNA-DNA hybridization on membrane filters: a convenient method using formamide. Biochim. Biophys. Acta **477**:191–194.
10. **Dudai, Y.** 1983. Mutations affect storage and use of memory differentially in *Drosophila*. Proc. Natl. Acad. Sci. USA **80**:5445–5448.
11. **Duerr, J. S., and W. G. Quinn.** 1982. Three *Drosophila* mutations that block associative learning also affect habituation and sensitization. Proc. Natl. Acad. Sci. USA **79**:3646–3650.
12. **Falkenthal, S., V. Parker, and N. Davidson.** 1985. Developmental variations in the splicing pattern of transcripts from the *Drosophila* gene encoding myosin alkali light chain result in different carboxyl-terminal amino acid sequences. Proc. Natl. Acad. Sci. USA **82**:449–453.
13. **Fyrberg, E. A., K. L. Kindle, N. Davidson, and A. Sodja.** 1980. The actin genes of *Drosophila*: a dispersed multigene family. Cell **19**:365–378.
14. **Fyrberg, E. A., J. W. Mahaffey, B. J. Bond, and N. Davidson.** 1983. Transcripts of the six *Drosophila* actin genes accumulate in a stage and tissue specific manner. Cell **33**:15–123.
15. **Gailey, D. A., F. R. Jackson, and R. W. Siegel.** 1982. Male courtship in *Drosophila*: the conditioned response to immature males and its genetic control. Genetics **102**:771–782.
16. **Hu, N., and J. Messing.** 1982. The making of single stranded probes. Gene **17**:271–277.
17. **Kandel, E. R., and J. H. Schwartz.** 1982. Molecular biology of learning: modulation of transmitter release. Science **218**:433–443.
18. **Kauvar, L.** 1982. Defective cAMP phosphodiesterase in the *Drosophila* memory mutant *dunce*. J. Neurosci. **2**:1347–1358.
19. **Kiger, J. A., Jr., and E. Golanty.** 1977. A cytogenetic analysis of cyclic nucleotide phosphodiesterase activities in *Drosophila*. Genetics **85**:609–622.
20. **Livingstone, M. S., P. P. Sziber, and W. G. Quinn.** 1984. Loss of calcium/calmodulin responsiveness in adenylate cyclase of *rutabaga*, a *Drosophila* learning mutant. Cell **37**:205–215.
21. **Nawa, H., H. Kotani, and S. Nakanishi.** 1984. Tissue-specific generation of two preprotachykinin mRNAs from one gene by alternative RNA splicing. Nature (London) **312**:729–734.
22. **Rosenfeld, M. G., S. G. Amara, and R. M. Evans.** 1984. Alternative RNA processing: determining neuronal phenotype. Science **225**:1315–1320.
23. **Rozek, C. E., and N. Davidson.** 1983. *Drosophila* has one myosin heavy-chain gene with three developmentally regulated transcripts. Cell **32**:23–34.
24. **Salz, H. K., R. L. Davis, and J. A. Kiger.** 1982. Genetic analysis of chromomere 3D4: the *dunce* and *sperm-amotile* genes in *Drosophila melanogaster*. Genetics **100**:587–596.
25. **Salz, H. K., and J. A. Kiger.** 1984. Genetic analysis of chromomere 3D4 in *Drosophila melanogaster*. II. Regulatory sites for the *dunce* gene. Genetics **108**:377–392.
26. **Shotwell, S. L.** 1983. Cyclic adenosine 3′-5′-monophosphate phosphodiesterase and its role in learning in *Drosophila*. J. Neurosci. **3**:739–747.
27. **Stark, G. R., and J. G. Williams.** 1979. Quantitative analysis of specific labelled RNA's using DNA covalently linked to diazobenzyloxymethyl paper. Nucleic Acids Res. **6**:195–203.