Stage-specific selection of alternative transcriptional initiation sites from the 5C actin gene of Drosophila melanogaster

Jim O. Vigoreaux1 and Sara L. Tobin2
Department of Biochemistry and Molecular Biology, University of Oklahoma Health Sciences Center, Oklahoma City, Oklahoma 73190 USA

The transcription unit of the 5C actin gene exhibits a complex organization that is unique among the six actin genes of Drosophila melanogaster. Three different mRNA size classes showing distinct patterns of accumulation throughout development are detected on Northern blots. We have determined the structure of the various 5C actin transcripts by exon mapping using strand-specific RNA probes, primer extension analysis, and DNA sequences analysis of both cDNA and genomic clones. All the transcripts share a single protein-coding nucleotide sequence but are heterogeneous in the 5' and 3' untranslated regions. The 5' untranslated region of each transcript consists of either one of two small exons (exon 1 and exon 2) which are alternatively spliced to a single acceptor site 8 bp upstream from the translation initiation codon in exon 3. Results from primer extension analysis suggest that transcription can initiate from either exon 1 or exon 2, and also from a third site within exon 2. We detect an increase in the relative abundance of exon 1-containing transcripts at larval and pupal stages, as well as a change in the proportion of transcripts that initiate at either of the two exon 2 sites. Five polyadenylation sites have been found within three termination/processing regions that define the three size classes of polyadenylated transcripts. The results of our experiments indicate the existence in vivo of all possible combinations of 5' exon with 3' polyadenylation site. However, particular combinations of 5' initiation site and 3' polyadenylation site are preferred at certain developmental stages.

[Key Words: 5C actin gene; alternate promoters; developmental regulation; Drosophila; 3' end heterogeneity]

Received May 28, 1987; revised version accepted October 14, 1987.

During recent years a number of genes have been shown to consist of complex transcriptional units that produce a variety of combinations of differing primary transcripts coupled to differential post-transcriptional processing events (for review, see Leff et al. 1986). Analysis of these transcriptional units has defined new levels of gene regulation and has modified our concept of the genetic unit. Diverse strategies have been revealed by which differential expression of a single gene can be accomplished. Some examples include alternative use of transcription initiation sites (Carlson and Botstein 1982; Benyajati et al. 1983; Schibler et al. 1983) or 3' polyadenylation sites (Capetanaki et al. 1983), differential splicing of primary transcripts (Breitbart et al. 1985; Cooper and Ordahl 1985), or any combination of these mechanisms (Early et al. 1980; Rosenfeld et al. 1983; Nabeshima et al. 1984; Periasamy et al. 1984; Robert et al. 1984). In these cases, the particular choice of promoter, polyadenylation site, or splicing pathway is utilized in a tissue-specific and/or development-specific mode. This diversity can result in different transcripts coding for variants of the same protein that have become specialized for the particular cell type or developmental period during which they are expressed (Caplan et al. 1983). On the other hand, several examples have been described in which multiple transcripts from a single gene encode identical proteins but differ in the length and sequence of their 5' and/or 3' untranslated regions (Benetzen and Hall 1982; Freytag et al. 1984; McGrogan et al. 1985; Schneuwly et al. 1986; Urano et al. 1986). The function of these transcribed but untranslated sequences is not known, but it has been speculated that they might be involved in translational efficiency (Davau et al. 1985), mRNA instability (Belasco et al. 1986), or cellular targeting of transcripts (Lawrence and Singer 1986). The study of genes that produce multiple transcripts differing in non-coding sequences may elucidate whether these untranslated regions have a function in post-transcriptional regulation of gene expression.

A common feature of most eukaryotic genomes is an actin multigene family (Firtel 1981; Buckingham and Minty 1983). Since actin is a ubiquitous protein involved in a variety of cellular, morphological, and developmental events, multigene families can offer several advantages to the organism. Repeated copies of a gene

1Present address: Department of Biology 16-719, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139 USA.

2Corresponding author: Dr. Sara L. Tobin, Department of Biochemistry and Molecular Biology, O.U.H.S.C., P.O. Box 26001, Oklahoma City, Oklahoma 73190 USA.
Vigoreaux and Tobin

can produce a greater amount of product than could a single-copy gene, a phenomenon referred to as dosage repetition (Finnegan et al. 1977; Long and Dawid 1980). Either alternatively or in addition to dosage repetition, each member of a multigene family could code for an isoform with a specialized function (Whalen et al. 1981; Weydert et al. 1983; Caplan et al. 1983) and/or possess a unique profile of transcript accumulation throughout development (Fyrberg et al. 1980; Garcia et al. 1984; Mohun et al. 1984; Cox et al. 1986). In Drosophila melanogaster, six actin genes have been shown to encode proteins whose amino acid sequences share greater than 85% homology (Tobin et al. 1980; Fyrberg et al. 1981; Sanchez et al. 1983). While the protein products are similar, each actin gene exhibits a unique profile of transcript accumulation throughout development (Fyrberg et al. 1983; Sanchez et al. 1983), suggesting that they have evolved independent regulatory programs.

The actin gene found at polytene chromosome locus 5C has been identified as a cytoplasmic actin gene (Fyrberg et al. 1980). In addition, Couderc et al. (1983) have shown that expression of the 5C actin gene is modulated by the molting hormone ecdysterone. The 5C actin gene differs from the other five Drosophila actin genes in one more respect: It is the only actin gene that is transcribed into developmentally regulated transcripts of different size classes (Fyrberg et al. 1983; Vigoreaux 1987). To understand the mechanism(s) responsible for differential expression of the 5C actin gene, we have determined the molecular organization of this gene and examined the structure of its multiple mRNA species. In this paper, we demonstrate that multiple mature 5C actin transcripts share a common protein coding region but differ in their utilization of 5' exons and 3' polyadenylation sites during development. Bond and Davidson (1986) have presented evidence regarding the structure of the various 5C actin transcripts. Our results confirm some of their conclusions and in addition, we present the first report of developmentally regulated alternative 5' initiation sites and splicing patterns in an actin gene in any organism.

Results

Isolation and characterization of 5C actin genomic DNA and transcripts

Figure 1A shows a partial restriction endonuclease map of RI-19, an 8.4-kb EcoRI genomic fragment containing the entire 5C actin gene isolated from a recombinant λ phage Drosophila genomic library (Maniatis et al. 1978). The entire amino acid coding region is located within a 1.9-kb HindIII-HindIII fragment, as determined by DNA sequence analysis (Fig. 1B). The protein-coding nucleotide sequences are highly conserved among Drosophila actin genes (Fyrberg et al. 1981; Sanchez et al. 1983). To study the accumulation of 5C actin homologous RNA, we used gene-specific probes that represent single-copy genomic sequences, do not contain protein coding nucleotides, and do not cross-react with heterologous transcripts. The 5' 3.7-kb EcoRI-HindIII fragment and the 3' 1.6-kb HindIII-HindIII fragment (Fig. 1A) fulfill these criteria (data not shown). Our preliminary studies on the patterns of accumulation of 5C actin transcripts were carried out with these probes.

These studies confirmed the finding that the 5C actin gene produces multiple transcripts (Fyrberg et al. 1983) and permitted us to focus our experiments on the developmental stages at which the greatest variation in accumulation of the different transcript size classes was observed. The 5C actin gene had been suggested by R-loop studies to consist of two exons: a short [70–170 bp] 5' 'leader' exon and a 1.55-kb exon containing the major portion of the gene (Fyrberg et al. 1980). However, this structure cannot, by itself, account for the generation of the various transcripts from this gene. We have carried out a detailed analysis of the structure of the 5C actin gene to determine which regions of the gene are represented in the several transcript size classes.

Characterization of different cDNAs and DNA sequence of corresponding genomic regions

We characterized several different actin cDNA clones that were identified as 5C-derived on the basis of their homology to either the 5' or 3' gene-specific probes [Fig.

Figure 1. Structure and partial restriction map of the 5C actin genomic clone and cDNA clones. (A) RI-19 is an 8.4-kb Drosophila genomic fragment containing the entire 5C actin gene. Boxed areas represent exons, filled boxes represent the protein-coding domain, and striped boxes represent the noncoding regions. Horizontal arrows indicate the extent and direction of sequencing from the indicated restriction site. Arrows with filled circles indicate fragments generated by BglII. The strategy shown was that used to generate the sequence shown in Fig. 1B. Only restriction sites used to generate probes or sequencing fragments are shown. The two subfragments used as gene-specific probes are indicated above the map. (B) DNA sequences of the 5'- and 3'-flanking regions of the 5C actin genomic and cDNA clones. +1 refers to the first amino acid-coding nucleotide. Nucleotides upstream from the open reading frame are given negative numbers, beginning with −1 for the first nucleotide preceding the ATG translation initiation codon. The region from −1062 to −1026 was not sequenced; numbering upstream from this 37-nucleotide gap was calculated following the sequence of Bond and Davidson (1986). Exons are in boldface italics with arrows indicating the 5' and 3' boundaries of the different cDNA clones. Locations of transcriptional start sites as determined by primer extension analysis are indicated by plus signs. The coding region sequence, which is common among all 5C actin transcripts, is not shown in its entirety. Putative regulatory elements (TATA box and polyadenylation signals) are indicated by brackets. Asterisks indicate mismatches between this sequence and the following previously published sequences: Parker and Topol (1984) from −1883 to −1781; Bond and Davidson (1986) from −2024 to −1608, −1148 to −371, and −52 to +36; Fyrberg et al. (1981) from −8 to +225. Relevant restriction endonuclease cleavage sites are shown, with the points of cleavage indicated by an open triangle.
induced third instar salivary gland cDNA library mined for all these clones. A comparison with the sequence of the corresponding genomic regions is shown in Figure 1B. All cDNAs exhibit a common 5’ splice site located eight nucleotides upstream from the translation initiation codon ATG. However, the sequences preceding the splice site differ. cDNAs p3A and p1D contain sequences that are homologous to a genomic region located 1.66 kb upstream from the splice site. On the other hand, the 5’ ends of pD11 and pD12 are homologous to a genomic region separated from the splice site by a 0.59-kb intron.

The cDNAs differ in the lengths of their 3’ noncoding regions [see Fig. 1B]. The DNA sequence of all four cDNAs is collinear with the genomic sequence, indicating that there are no splices in the 3’ region of the gene. The difference in length arises from differential termination and/or cleavage/polyadenylation sites. Ribonuclease-protected fragments corresponding to the length of the 3’ ends of these cDNA clones confirm that these cDNAs represent genuine 3’ ends [see below].

5’ end analysis of the 5C actin transcripts

The 5’ ends of the mature 5C actin transcripts were defined by quantitative RNase protection mapping, primer extension analysis, and DNA sequence analysis of cDNA clones and the corresponding regions of the genomic clone. Figure 2 shows results of ribonuclease protection experiments in which in vitro-generated single-stranded RNA probes corresponding to the 5’ region of the gene were hybridized to poly[A] RNA from different developmental stages. Three small fragments of approximately 150 nucleotides, 114 nucleotides, and 104 nucleotides are protected by probe 1. Identical results were obtained with a probe corresponding to the entire 5’ 3.7-kb EcoRI–HindIII region [results not shown]. Probe 2, on the other hand, only protects the 114- and 104-nucleotide fragments. These results suggest that the 150-nucleotide fragment corresponds to an exon located between the NdeI and BglII sites [exon 1 in Fig. 1a]. The 114- and 104-nucleotide protected fragments could correspond to two separate exons whose boundaries lie within the BglII and HindIII sites or, alternatively, one exon with two different 5’ or 3’ termini. This question was resolved by primer extension analysis using a 69-nucleotide, 5’-end-labeled SacII–TaqI fragment corresponding to nucleotides +32 to +100. As template, we used hybrid-selected mRNA homologous to a 5C actin gene specific probe to eliminate cross-hybridization with other actin gene transcripts. Three extension products of 142, 152, and 190 nucleotides were obtained [Fig. 3A]. These results suggest that there is a one-to-one correlation between extension products and the fragments protected by RNase digestion, since the difference in length among the three extension products is similar to that among the three protected fragments. This conclusion is further supported by the available sequence from the cDNAs [Fig. 1B].

These results, together with those from ribonuclease protection experiments, indicate that exon 1 is approximately 150 bp in length. A variant of the Drosophila cap consensus sequence ATCA[G/T][T/C] (Hultmark et al. 1986) is found 8 nucleotides upstream from the 5’ end of p1D [ATCACTA, from –1820 to –1814, Fig. 1B]. If transcription initiates at this site, then exon 1 would be 147 nucleotides, which agrees very closely with the size determined by both protection assays and primer extensions. Consequently, we assigned the A at position –1820 as the transcription initiation site for exon 1.

A second extension analysis was performed using a 53-bp primer isolated from cDNA clone pD11. The 5’ end of this clone extends only 57 nucleotides from the splice site to exon 3 [Fig. 1B]. Its genomic counterpart corresponds to nucleotides –657 to –601, which are located within the region where the 114- and 104-nucleotide ribonuclease-protected fragments map. To determine the relationship between the protected fragments and the 5’ end of pD11, primer extension analysis was carried out with RNAs from various developmental stages [Fig. 3B]. With the exception of 90–99 hr RNA, extension products of approximately 54 and 63 nucleotides were obtained with RNA from all the other stages. These results suggest that the 114- and 104-nucleotide protected fragments correspond to a single exon with two alternative initiation sites. In addition, we were unable to detect an extension product that would correspond to the usage of exon 1 and exon 2 in the same RNA molecule, confirming our conclusion that exon 1 and exon 2 are alternatively spliced to exon 3.

Hybridization of 5’ probes to Northern blots of poly[A] RNA indicate that both exon 1 and exon 2 are utilized in all three size classes of transcripts [Fig. 2B–D]. To estimate the relative abundance of exon 1-containing transcripts versus exon 2-containing transcripts, we quantitated the intensity of the signals generated in three comparable ribonuclease protection experiments [Fig. 2E] using a scanning densitometer. As seen in Table 1, exon 1-containing transcripts are more abundant than exon 2-containing transcripts in all three stages examined. However, there is a twofold increase in the proportion of transcripts containing exon 1 in 90–99 hr larval RNA and 150–158 hr pupal RNA as compared to 0–12 hr embryo RNA. A second observation is that transcripts containing exon 2A are twice as abundant as transcripts containing exon 2B at 90–99 hr and 150–158 hr, while their ratio is almost 1 : 1 at 0–12 hr. This modulation of 5C actin transcripts could result from a preferential stability of particular transcripts as a function of development or from state-specific control mechanisms involved in the selection of 5’ initiation site.

Coding region and 3’ end analysis of the 5C actin transcripts

Ribonuclease protection experiments were also carried
Multiple *Drosophila* 5C actin transcripts

![Image of Figure 2](https://example.com/figure2.png)

**Figure 2.** (A) RNase protection analysis of 5C actin transcripts. Mapping the 5' end of the gene. The 5' 3.7-kb *EcoRI*-*HindIII* genomic fragment was cloned into the SP6 polymerase vector pSP64 and used to generate overlapping probes by linearizing at the *NdeI* site (probe 1) and *BglII* site (probe 2). (m) *ϕX 174*-*HaeIII* digest used as molecular weight standards; (P) undigested probe; 0–12 hr embryo poly(A) RNA, 90–99 hr larvae poly(A) RNA, 150–158 hr pupae poly(A) RNA, and yeast tRNA (c) used as control. Numbers on the right margin refer to the size of the protected fragments. (B–D) Northern blots hybridized to probe 1, probe 2, and a 0.8-kb *Clai*-*SacII* probe (exon 1 probe in diagram), respectively. All blots were exposed for 20 hr (E) Overexposure of autoradiogram A (144-hr exposure).

out with probes corresponding to the entire coding region of the 5C actin gene and extending 2.0 kb downstream from the translational stop (data not shown). The conclusions drawn from these data suggest that exon 3, which contains the entire protein coding region, is contiguous with the 3' end of the gene. The differential lengths of the transcripts are generated by 3' cleavage/polyadenylation sites that attenuate the primary transcript at different points. Three major cleavage/polyadenylation regions with three major and two minor polyadenylation sites were identified. Each of the cDNA clones analyzed above corresponded to a site identified by these ribonuclease protection assays. The results of these experiments closely parallel those of Bond and Davidson (1986), with the addition of our detection of the two minor cleavage/polyadenylation sites summarized in Figure 4. These results indicate that the 3' ends of the 1.8-kb transcripts as well as the 3' ends of the 2.0-kb transcripts are heterogeneous.

**Discussion**

The major results of this study are summarized in Figure 4. We have determined the structure of the 5C actin gene and have shown that it consists of one large exon (exon 3) containing the protein-coding and 3' noncoding...
region of the gene which is spliced to one of two alternative 5' noncoding exons. Exon 1 and exon 2 are separated from the coding exon by introns of approximately 1.66 kb and 0.59 kb, respectively, and each exon accounts for all but 8 bp of the 5' untranslated sequences present in 5C actin transcripts. The coding exon containing the entire open reading frame extends from position −8 (relative to the translation initiation codon) to one of three different polyadenylation regions, generating three distinct molecular weight classes of polyadenylated transcripts that differ in the length of their 3' noncoding regions.

The transcription initiation sites were mapped by primer extension to define the 5' boundaries of exon 1 and exon 2. A third initiation site has been mapped within exon 2, 10 nucleotides downstream from the second initiation site. Microheterogeneity of 5' termini is a common feature of many genes [see, for example, Baker and Ziff 1981]; however, this heterogeneity is usually within a shorter region (two to seven bases). Microheterogeneity could also be due to an artifact in the primer extension reaction resulting from an encounter of the reverse transcriptase with the 5' cap structure [Chan et al. 1980]. However, such doublets are never seen for the longest extension product (exon 1); thus, they are not due to cap structure alone. Second, a protected fragment corresponding to transcription initiation from this site is routinely observed in the RNase protection experiments. Exon mapping experiments in which RNase T2 (which is not sequence specific) was used instead of RNase A and RNase T1 also resulted in the 114/104-nucleotide protected fragment doublet [results not shown]. The repeated appearance of this site in independent experiments leads us to propose that it represents a bona fide 5' initiation site. Moreover, the stage-specific modulation of exon 2A- versus exon 2B-containing transcripts suggests that the sites are under separate genetic control.
Multiple Drosophila 5C actin transcripts

EXON 1  EXON 2  EXON 3

transcript size

1.8Kb

2.0Kb

2.3Kb

Figure 4. Schematic summary of the multiple 5C actin transcripts. Open reading frame is indicated by filled boxes and noncoding regions by striped boxes. Transcription initiation sites and cleavage/polyadenylation sites are also shown.

An analysis of the nucleotide sequence in the 5' vicinity of exon 2 failed to reveal a consensus TATA box homology. This sequence has been shown to be involved in determining the proper site of transcription initiation (Benoist and Chambon 1981; Breathnach and Chambon 1981; Mathis and Chambon 1981). Several genes have been described that do not contain a typical TATA box and show multiple in vivo transcription initiation sites within a short region (Allen et al. 1983; Reynolds et al. 1984). Thus, the lack of a characteristic promoter region could be responsible for the two initiation sites associated with exon 2. Although we have not mapped the precise 5' end of exon 2, our results allow us to conclude that this exon has two initiation sites that are approximately 10 nucleotides apart. Using the primer extension results, we have placed these two 5' ends at positions -705 and -715 (Fig. 2). Bond and Davidson (1986) placed a single initiation site for exon 2 at position -711. Unlike exon 2, exon 1 is associated with a TATA box and only shows one transcription initiation site. Based on partial homology to a Drosophila cap consensus sequence (Hultmark et al. 1986), we assigned nucleotide -1820 as the transcription initiation site for exon 1. Bond and Davidson (1986) also designated this nucleotide as the start point, but Parker and Topol (1984) designated nucleotide -1822 as the start point.

The role of 5' leader sequences in bacterial transcription attenuation (Lee and Yanofsky 1977; Carter et al. 1986) and mRNA stability (Green and Inouye 1984; Belasco et al. 1986) is well documented. More recent studies have shown that untranslated leader regions of eukaryotic mRNA play a very important role in transcriptional and translational control (Darveau et al. 1985; McGarry and Lindquist 1985; Pelletier and Sonenberg 1985; Hultmark et al. 1986; Mueller and Hinnbusch 1986). In many cases, different length 5' noncoding regions reflect transcription initiation from separate promoters, indicating that regulation lies at a transcriptional level (Benyajati et al. 1983; Schibler et al. 1983). In the case of the 5C actin gene, the two 5' noncoding exons have very similar length [147 bp vs. 114 bp] and sequence composition [61% vs. 56.5% A + T], but they are utilized with different frequencies throughout development. The functional significance for these noncoding exons is not immediately evident, but there are several possibilities.

Each 5' exon could be under the control of a separate promoter that responds to different developmental, mitogenic, hormonal, or tissue-specific signals, thus conferring differential expression capabilities to this gene. The 5C gene encodes a cytoplasmic actin (Fyrberg et al. 1980) and it also has been shown to respond to hormonal stimulus (Couderc et al. 1983). Therefore, the existence of cis-regulatory elements that mediate 5C actin response to an array of transcription-controlling factors is not unlikely. The presence of multiple initiation sites could simply reflect the evolution of these transcriptional regulatory elements into separate entities. The divergence of the DNA sequence in the regions immediately preceding exon 1 and exon 2, together with the presence of a TATA box for exon 1 but not for exon 2, argues in favor of independent promoters for these two exons. Another possibility is that structural determinants exist within the two noncoding exons that are involved in separate aspects of post-transcriptional or translational control. These could include rapid changes in mRNA stability in response to a controlling signal such as a steroid hormone (Brock and Shapiro 1983), or differential ribosome binding efficiencies and therefore differential translational efficiencies (Pelletier and Sonenberg 1985). Finally, Lawrence and Singer (1986) have suggested that noncoding regions and, in particular, noncoding regions of actin transcripts could be involved in intracellular targeting (trafficking) of transcripts. Recent in situ hybridization experiments using probes representing either exon 1 or exon 2 sequences revealed...
completely different patterns of accumulation of exon 1- versus exon 2-containing transcripts in embryonic tissue sections [J. Vigoreaux, T. Burn, and S. Tobin, unpubl.].

The steady-state level of each transcript size class varies through the developmental stages examined, suggesting that the mechanism[s] involved in polyadenylation site selection and/or 3' end processing are under developmental control. Although 3' untranslatable sequences are a common feature of eukaryotic mRNAs, their function[s], if any, is unclear. Recent studies have revealed segments of homology that have been highly conserved among homologous genes from several species [Martin et al. 1981; Gunning et al. 1984; Yaffe et al. 1985], or among structurally unrelated genes that are functionally related [Caput et al. 1986]. The existence of these conserved sequences suggests that 3' noncoding regions, or at least subsegments of 3' noncoding regions, may be involved in some aspect of gene regulation, possibly at either the transcriptional or post-transcriptional level. This is supported by recent studies identifying a conserved sequence in the 3' noncoding region of certain mRNAs that appears to mediate the steady-state level of these molecules (Shaw and Kamen 1986). In addition, several studies have indicated that the translation efficiency of an mRNA could be determined by sequences in the 3' untranslated region (Liebhaber and Kan 1982, Miller et al. 1984, Treisman 1985). In particular, we noted that the transcripts with the longest 3' noncoding region [the 2.3 kb class] show the most modulation throughout development. Thus, it is possible that this region contains sequences that act as signals for a rapid mRNA turnover pathway that is active during particular developmental situations. Gene transformation experiments in which 3' noncoding regions are fused to easily identifiable reporter genes is one approach that could prove useful in addressing some of these issues.

The Drosophila 5C actin gene is the second actin gene that has been reported to produce multiple transcripts. Carroll et al. [1986] have reported that the chicken α-smooth muscle (aortic) actin gene produces transcripts of four different sizes that are apparently the result of multiple polyadenylation site utilization. The developmental regulation of the use of these multiple polyadenylation sites was not investigated in this report. Primer extension experiments did not suggest the existence of 5' heterogeneity in the transcripts from this actin gene [Carroll et al. 1986]. Therefore, the study presented here is the first report of developmental specificity in the regulation of 5' initiation/splicing patterns in the generation of actin transcripts in any organism.

Materials and methods

Isolation and subcloning of 5C actin genomic DNA

5C actin genomic DNA was isolated from the D. melanogaster phage λ genomic bank of Maniatis et al. [1978] as described (Sanchez et al. 1983). The probe used was an actin plasmid [pKla] isolated from HindIII-cleaved D. melanogaster Schneider cell 2 DNA (Tobin et al. 1980). An 8.4-kb EcoRI fragment containing the entire 5C actin gene was subcloned into the EcoRI site of pBR322. The cloned DNA was identified by in situ hybridization to polytene chromosomes (Gall and Pardue 1971) as described by Tobin et al. [1980]. Isolation of DNA fragments through preparative agarose gels, subcloning into pBR322, transformation of E. coli RR1, and preparation of plasmid DNA were carried out as described [Tobin et al. 1980].

Isolation, fractionation, and blotting of RNA

RNA was isolated from developmentally staged whole organisms by a phenol extraction procedure [Lamb and Laird 1976] for embryo RNA or by the guanidinium isothiocyanate/cesium chloride procedure of Chirgwin et al. [1979] as adapted by Tobin et al. [1980]. Poly(A) RNA was selected by two successive fractionations through oligo(dT)-cellulose, electrophoresed on 1.5% formaldehyde agarose gels, and transferred with no further treatment to Biodyne nylon membranes (Pall Ultrafine Filtration Corp.) in 20 x SSC for 18–24 hr [Maniatis et al. 1982]. After transfer was completed, filters were air-dried for 1 hr and baked for 2 hr at 80°C in a vacuum oven.

Filters were prehybridized at 55°C for 4–8 hr in 50% formamide, 50 mM sodium phosphate (pH 6.5), 5 x SSC, 0.1% SDS, 5 x Denhardt’s solution [Denhardt 1966], and 200 μg/ml denatured E. coli DNA. They were then hybridized to 3–5 x 10^6 cpm/ml of probe in four parts prehybridization solution : 1 part 50% (wt/vol) dextran sulfate. Hybridizations were carried out at 55°C for 18–24 hr. Filters were washed once at room temperature for 30 min and 3–5 x for 30 min each at 65°C in 0.1 x SSC, 0.1% SDS. Filters were exposed to Kodak XAR-5 or XAR-2 film using intensifying screens. Blots that were reused were washed at 90°C for 1–3 hr in 50% formamide, 50 mM sodium phosphate (pH 6.5), 5 x SSC, and 0.1% SDS.

In vitro transcription with SP6 RNA polymerase and RNase mapping

The plasmids used as templates for synthesizing single-stranded RNA probes were constructed by inserting the appropriate DNA fragment into compatible sites of the Riboprobe SP6 vectors pSP64 or pSP65 [Promega Biotec]. Plasmids were propagated in E. coli RR1 and plasmid DNA prepared by centrifugation in cesium chloride gradients [Maniatis et al. 1982]. Plasmids were linearized to completion with appropriate restriction enzymes that cut once within the cloned insert or the polylinker region. phenol-chloroform-extracted, and ethanol-precipitated twice. These linear plasmid DNAs were used as templates for in vitro transcription with SP6 RNA polymerase as described by Melton et al. [1984] with the following modifications. Synthesis was carried out for 10 min at 40°C rather than for 1 hr and the transcription mix was supplemented with 100 μM unlabeled UTP when [α32P]-UTP was used. These modifications resulted in almost 90–100% full-length probes as determined by electrophoresis on denaturing gels. After synthesis, the reaction was stopped with RQ1 DNase I [Promega Biotec] and extracted once with 1 : 1 phenol/chloroform, and ethanol-precipitated. The resulting RNA–RNA hybrids were extracted once with 1 : 1 phenol/chloroform and precipitated with ethanol; the pellets were resuspended in 3 μl of loading buffer and electrophoresed on 6% acrylamide–8 M urea sequencing gels [Maxam and Gilbert 1980]. 32P-Labeled 174-6x HaeIII DNA fragments [New England Biolabs] were used as molecular weight standards. Autoradiograms were
quantiﬁed with the use of a soft laser scanning densitometer (Biomed Instruments) and the areas under the peaks were determined with an electronic graphics calculator (Numonics Corp, North Wales, Pennsylvania).

**Primer extension analysis**

A 69-bp Sall–TaqI fragment (+32 to +100) was isolated and end-labeled by ﬁrst removing 5’ phosphates with alkaline phosphate (Boehringer Mannheim), followed by replacement with [α-32P]ATP (Amersham) using T4 polynucleotide kinase. The labeled fragment (0.1–1.0 × 10^6 cpm) was mixed with hybrid-selected RNA in a solution containing 80% formamide, 0.4 M NaCl, 10 mM PIPES (pH 6.8), and 2 mM EDTA, heated for 10 min at 90°C, and immediately transferred to 42°C overnight. After hybridization, the mixture was ethanol-precipitated and washed free of formamide with 70% ethanol. cDNA extension of RNA–DNA hybrids was carried out as described by Ingolia et al. (1980). The resulting extension products were extracted once with 1:1 phenol/chloroform and precipitated with ethanol, and the pellets were resuspended in loading buffer and electrophoresed on 6% acrylamide sequencing gels (Maxam and Gilbert 1980).

RNA homologous to the 5C actin 3’ 1.6-kb HindIII–HindIII fragment was selected by the positive-hybridization technique of Ricciardi et al. (1979). DNA was digested with HindIII (for a total of 5 μg of insert), and the fragments were separated on a 1% agarose gel and transferred to nitrocellulose (Southern 1975). The portion of the ﬁlter containing the 1.6-kb HindIII–HindIII insert was used to select homologous RNA from 30–50 μg of poly[A] RNA.

A 53-nucleotide primer from cDNA pD 11 was isolated, end-labeled (see above), and digested with restriction enzyme Hhal to remove the GC tail; the products were separated on a 10% acrylamide preparative gel. The band corresponding to the 53-nucleotide fragment was detected by autoradiography, excised from the gel, and the DNA electroeluted and processed as above. Primer extensions were done as described above except that hybridization was carried out at 45°C.

**Isolation and characterization of cDNA clones**

cDNAs pD11 and pD12 were synthesized from poly(A) RNA isolated from third instar salivary gland induced in vitro with ecdysone (Wolmer 1980). Ten clones identiﬁed as actin cDNAs were kindly provided by G. Guild and M. Wolmer. SC actin cDNAs were identiﬁed by hybridization to the 5’ and 3’ non-coding probes (see Fig. 1). Further characterization included restriction mapping. Southern blotting, and DNA sequence determination of selected regions.

A second cDNA library from imaginal disc RNA cloned into λ gt10 (kindly provided by J. Fristrom) was screened with a SC actin coding region probe using standard screening techniques (Maniatis et al. 1982). An initial screen of 5 × 10^9 plaques produced 24 positives. Nine SC actin clones were identiﬁed by hybridization to 5’ and 3’ non-coding probes. CDNA clones p1D and p3A were isolated and subcloned into pTZ vectors (Pharmacia) and characterized as described above.

**DNA sequence analysis**

DNA sequencing was performed by the dideoxy chain-termina-
tion procedure of Sanger et al. (1977) using a commercial M13 cloning kit (Amersham). Restriction fragments were cloned directly into compatible sites of the vectors M13mp18 and M13mp19 following the procedures of Messing (1983). Overlapping sequence data were obtained by digesting the appropriate restriction fragment separately with Sau3A, TaqI, and Bal31. Progressive shortening of target DNA by Bal31 was done by the procedure of Guo et al. (1983). pBR322 plasmids containing the 1.6-kb HindIII–HindIII genomic fragment and the 1.9-kb HindIII–HindIII genomic fragment were linearized with EcoRI. After treatment with Bal31, linear DNA molecules were blunt-end-repaired with Klenow polymerase (Maniatis et al. 1982) and digested with HindIII and BamHI. Fragments were then cloned into HindIII–SmaI-cut M13mp19. For cDNA clones pD11 and pD12 (whose inserts are in the PstI site of pBR322), plasmid DNA was linearized with PvuI, followed by Bal31 treatment, Klenow repair, and digestion with PstI and BamHI. Fragments were cloned into PstI–SmaI-cut m13mp19. DNA sequence analyses were analyzed using the computer programs GEL and SEQ developed by Intelligenetics (Palo Alto, California).

**Acknowledgments**

We would like to thank Greg Guild and Mariana Wolmer for the gift of actin cDNA clones and James Fristrom for providing the imaginal disc cDNA library. We appreciate the technical assistance of Patricia Cook in screening the cDNA library. We would also like to thank Olivia Hanson-Painton and Jay Hanas for reviewing this manuscript and Olivia Hanson-Painton for her help in the DNA sequence analysis. The assistance provided by Bruce Roe, Elaine Johnston, and Ricky Wong with the DNA sequence is also appreciated. We are grateful to Toni Johnson and Joyce Anderson for their patience in typing this manuscript. Sequence data were processed using the Bionet computer resource (Grant 1 U41 RR-01685-03). This work was supported by a grant from the National Institutes of Health (R01-GM32588) to S.T.

**Note**

Sequence data described in this paper have been submitted to the EMBL/GenBank Data Libraries under the accession number Y00212.

**References**


*Multiple Drosophila SC actin transcripts*
Vigoreaux and Tobin

Brock, M.L. and D.J. Shapiro. 1983. Estrogen stabilizes vitello-
Caplan, A.I., M.Y. Fiszman, and H.M. Epplenberger. 1983. Mo-
lecular and cell isoforms during development. Science 221: 921–927.
Cox, K.H., L.M. Angerer, J.J. Lee, E.H. Davidson, and R.C. An-
Firtel, R.A. 1981. Multigene families encoding actin and tubu-
Fischer, J.A. and T. Maniatis. 1986. Regulatory elements in-
volved in Drosophila Adh gene expression are conserved in divergent species and separate elements mediate expression in different tissues. EMBO J. 5: 1275–1289.
Fyrb erg, E.A., B.J. Bond, N.D. Hershey, K.S. Mixter, and N. Da-
Hultmark, D., R. Kle menz, and W.J. Gehring. 1986. Transla-
tional and transcriptional control elements in the untrans-
Leff, S.E., M.G. Rosenfeld, and R.M. Evans. 1986. Complex transcriptional units: diversity in gene expression by alter-
Liehbbaer, S.A. and Y. Kan. 1982. Different rates of mRNA translation balance the expression of the two human α-
Multiple Drosophila SC actin transcripts


Stage-specific selection of alternative transcriptional initiation sites from the 5C actin gene of Drosophila melanogaster.

J O Vigoreaux and S L Tobin

*Genes Dev.* 1987 1: 1161-1171
Access the most recent version at doi:10.1101/gad.1.10.1161

---

**References**

This article cites 77 articles, 21 of which can be accessed free at:
[http://genesdev.cshlp.org/content/1/10/1161.full.html#ref-list-1](http://genesdev.cshlp.org/content/1/10/1161.full.html#ref-list-1)

**Email Alerting Service**

Receive free email alerts when new articles cite this article - sign up in the box at the top right corner of the article or [click here](http://genesdev.cshlp.org/subscriptions).

---

To subscribe to *Genes & Development* go to:
[http://genesdev.cshlp.org/subscriptions](http://genesdev.cshlp.org/subscriptions)

Copyright © Cold Spring Harbor Laboratory Press