THE MOLECULAR STRUCTURE OF THE *BEADEX AND HELDUP-A* LOCI OF *DROSOPHILA MELANOGASTER*

Thesis by

William Wayne Mattox

In Partial Fulfillment of the Requirements

for the Degree of

Doctor of Philosophy

California Institute of Technology

Pasadena, California

1987

(Submitted July 14, 1986)

This thesis is dedicated to the memory
of my great-grandmother, Mabel Blair,
to whom education was paramount.

Acknowledgments

First and foremost, I would like to thank Norman Davidson, my advisor, for tolerating me and my quirky ideas. It is hard to imagine a better advisor than Norman. He was helpful when I needed help, patient when I needed time, and always concerned with what was best for me.

Several others have contributed to my education at Caltech. Scott Falkenthal and Vann Parker taught me much of what I know about molecular biology and were good friends outside the lab. Ron Davis got me into "Drosophology" and encouraged me to persue my own ideas. Dave Hershey, Arthur Roach, Sandy Sharp, Mickey Hu, Al Goldin, Kath Mixter, Nevis Fregein, Chuck Rozek, Mike Snyder and others in Norm's lab were always willing to listen and give their advice when problems arose. Madeline Crosby, Sue Celniker, Sarah Smolik-Utlaut and Loring Craymer helped me understand aspects Drosophila biology that were foreign to sub-basement dwellers. Mark Garfinkel also provided advice, but more importantly, he expressed a constant interest in my work that I found to be both refreshing and encouraging.

Barry Maurer, with whom I shared many bottles of wine, provided a unique perspective on the world that often led to interesting discussions. Whether it was science, tennis or last minute corrections in my thesis; my friend, Lei Yu, was always willing to help out. My friends Bev and Greg Matthews have made my tenure at Caltech indescribably more pleasurable. I want to exspecially (sic) thank Bev for her help with this thesis and her constant encouragment. All of my friends in the Eating Group at Prufrock House provided a pleasant diversion each evening. I will sorely miss both their company and their desserts.

I was supported, for much of my stay at Caltech, by a National Research Service Award from the National Institutes of Health. Support was also provided by the Wiegle Fund.

Finally, I would like to thank the Richardson Vending Company for providing such an assortment of culinary delights, without which I would have never made it through the last days of writing this thesis.

ABSTRACT

Genetic studies indicate that excess-of-function *Beadex* mutations and loss-of-function *heldup-a* mutations affect different parts of a single bipartite genetic unit. In order to investigate the molecular nature of the *Beadex* and *heldup-a* mutations we isolated DNA from a 49 kilobase region surrounding the sites at which these mutations occur. We found gross structural alterations associated with one *heldup-a* mutation and each of the 13 different *Beadex* mutations that we examined. As expected from previous genetic studies these loci are separated by only a short molecular distance which is at most 1.5 kilobases. The structural alterations associated with the thirteen *Beadex* excess-of-function mutations examined are clustered within a three kilobase region. Several of these mutations are found to be associated with the deletion of part of an 800 bp region. This indicates that an element that normally represses gene activity is located within this region: A 200 base pair segment which is required for the function of the wild-type *heldup-a* locus is defined using a *heldup-a* mutation which results from a small deletion.

Two RNA transcripts have been found which span this 200 base pair segment. One of these two transcripts, a four kilobase RNA, is expressed during stages in which the *Beadex* structural gene product is expected to be active. The structure of this transcript is affected by one *heldup-a* mutation and each of five *Beadex* mutations that were examined. However, only one of the *Beadex* mutant alleles expresses significantly elevated levels of this RNA. Other RNA species in the region surrounding the the *Beadex* and *heldup-a* loci were not affected either in amount or size by *Beadex* mutations.

We have also reintroduced a wild-type 10.4 kilobase fragment, which includes the region in which *Beadex* and *heldup-a* mutations map, into the *Drosophila* genome by P element mediated transformation. This introduced

fragment fails both to complement loss-of-function *heldup-a* mutations and to enhance an excess-of-function phenotype. This result indicates that some sequences required for the normal *heldup-a* function must be located outside this 10.4 kilobase region.

TABLE OF CONTENTS

INTRODUCTION

GENERAL INTRODUCTION

Despite the intricate mechanisms employed by living cells to maintain and accurately transmit their entire genome, only a fraction of the genes contained within it are utilized by any single cell. This is because many of the gene products that are encoded in the genome would be superfluous or detrimental to the requirements of a particular cell encountering a particular situation. In order to assess a variety of external and internal stimuli and to activate and inactivate the appropriate genes, cells employ elaborate regulatory networks. Many of the components of these networks are themselves the products of genes. Other components are sequence elements which do not encode diffusable products but play roles as binding or recognition sites for other regulatory molecules. Any genomic component which accomplishes an identifiable regulatory function can be regarded as a "regulatory gene." On the basis of genetic criterion, regulatory genes fall into two classes. Those genes that exert regulatory effects on other genes regardless of their positions in the genome are called *trans*-acting regulatory genes, while those genes which must reside on the same chromatid (and usually in close proximity to) the genes that they regulate, in order to exert their control, are called *cis*-acting. Both types of regulatory genes may act upon other genes positively, to activate them, or negatively to repress them.

Although it is possible to imagine mechanisms by which a gene's activity could be regulated using only *trans* or only *cis* acting regulatory genes, most systems examined to date use both types. Generally, *trans*-acting genes produce products that interact in some way with *cis*-acting elements[1]. These products may do this either directly, by binding to or modifying the *cis*-acting elements

[1]There is some divergence of opinion as to the semantic issue of whether or not *cis*-acting sequences, which have an identifiable genetic function but do not encode an RNA product (e.g., a binding site) should be referred to as regulatory "genes." To avoid this I will refer to them as regulatory "elements" in the rest of this discussion.

themselves, or indirectly, via the action of other *trans*-acting regulatory factors. *Cis*-acting elements may reside within or outside the transcription units of the genes which they control, and they may function in the form of DNA, RNA, or protein (see below).

The work presented in this thesis concerns a particular *cis*-acting regulatory element in the fruit fly *Drosophila melanogaster*, the *Beadex* locus. This locus is thought to control the level of activity of a nearby gene called *heldup-a* (Lifshytz and Green, 1979, this thesis).

Two of my principal findings are that the wild-type *Beadex* locus acts negatively to reduce gene activity (presumably that of *heldup-a*) and that this repression of gene activity[2] probably occurs at the level of protein synthesis or function. Below, I discuss what is known at present about some other eucaryotic *cis*-acting negative regulatory elements, because these systems provide a useful framework for thinking about *Beadex*.

EUCARYOTIC *CIS*-ACTING NEGATIVE REGULATORY ELEMENTS

Although a variety of *cis*-acting negative regulatory elements have been studied in procaryotes, only a few such elements have been well-studied in eucaryotic systems. However, among these elements a diversity of regulatory mechanisms is displayed. I have confined the discussion below to those elements for which molecular details are available.

T antigen binding sites. One of the best understood eucaryotic regulatory systems in which a *cis*-acting negative element plays a role is the one which regulates the transcription of early RNAs during the life cycle of the simian virus 40 (SV40) (Tjian, 1981). Shortly after infection of a permissive host cell, this DNA

[2]In this thesis the term "gene activity" refers to the activity of a gene at the phenotypic level. Therefore, an increase in the amount of gene activity may or may not reflect an increase in the transcriptional activity of the gene itself.

virus expresses mRNAs which encode tumor antigens (T antigens) (Sambrook et al., 1973). The so-called "early promoter," which is located between 70 and 155 base pairs upstream of the transcription start site (Myers et al., 1981), is required in *cis* for the synthesis of these mRNAs. The large T antigen binds at three closely spaced but distinct sites which are located near the 5' end of the major early transcription unit (Reed et al., 1975; Tjian, 1978). Site I is located within the 5' end of this transcription unit. Site II overlaps the transcription initiation site while site III is located just upstream of the initiation site and may overlap the early promoter. There is a hierarchal relationship between the binding affinities of these three sites for large T antigen such that site I is bound more readily than site II, which, in turn, is bound more readily than site III.

In vitro studies have shown that the binding of sites I and II by large T antigen results in repression of early gene transcription (Rio et al., 1980). Furthermore, mutations that affect this binding cause a five-fold overproduction of early mRNA *in vivo* (Tegtmeyer et al., 1975). Thus, the T antigen binding sites I and II make up a *cis*-acting negative regulatory element which serves to autoregulate T antigen synthesis at the transcriptional level.

When T antigen binding sites are inserted within other transcription units (downstream from the initiation site) they have no effect on the length or amount of transcripts made, even when all three sites are occupied by T antigen (Rio et al., 1980; Tjian, 1981). This led Tjian (1981) to suggest that bound T antigen does not repress transcription by blocking the movement of RNA polymerase down the DNA template but rather by blocking transcription initiation. This idea is supported by the observation that the position of transcription initiation changes when only site I is occupied (Hansen et al., 1981).

In many respects the T antigen binding sites of SV40 are like the procaryotic elements which bind the lambda repressor and *cro* proteins (Ptashne et al., 1980).

In both cases there are three binding sites with hierarchal affinities for the repressor protein; the particular pattern of bound sites can determine which sites are used to initiate transcription and binding can be autoregulative. Thus, it seems possible that SV40 T antigen binding sites may play a role in repression of transcription initiation which is similar to that played by these procaryotic operator elements.

The *STE6* operator. Another eucaryotic *cis*-acting negative element that acts at the level of transcription initiation is found in *Saccharomyces cerevisiae*. In cells of the *a* mating type, a set of genes which is normally expressed in cells of the *a* mating type (the *a* specific genes) is repressed in *trans* by the action of the *MAT*α2 gene (Strathern et al., 1981; Sprague et al., 1983). Sequence comparisons between the five known *a*-specific genes revealed a common 32 base pair (bp) upstream sequence (Johnson and Herskowitz, 1985; Miller et al. 1985). Johnson and Herskowitz showed that the *MAT*α2 protein binds within a copy of this 32 bp repeat, which is located upstream of the *a*-specific gene, *STE6*. When a synthetic copy of this sequence is inserted in the upstream flanking region of the *CYC1* gene which is not normally under the control of *MAT*α2, it causes a 200-fold reduction in *CYC1* mRNA in the presence of the *MAT*α2 protein. The repression was equally effective when the 32 bp repeat was inserted in either orientation and was most effective when the repeat was placed between the upstream activator sequences (UAS) and the TATA box sequence. However, the repeat could also cause a 20-fold repression when placed 60 bp upstream of the UAS.

Because it binds a repressor (*MAT*α2) and affects a gene in *cis*, the 32 bp repeat located upstream of *STE6* can be considered to be a eucaryotic operator element (Johnson and Herskowitz, 1985) as can T antigen binding sites I and II. It should be noted, however, that these yeast elements may function when located at positions that are hundreds of base pairs from the transcription initiation site.

This is quite unlike procaryotic operators, which are found in close proximity to promoters so that bound repressor may sterically block bacterial RNA polymerase. Since the role that the UAS and other sequence elements play in RNA polymerase II recognition and transcription initiation are not yet understood, it is difficult to say whether or not the binding of *MAT*α2 blocks promoter recognition by RNA polymerase II. Johnson and Herskowitz (1985) have proposed a model wherein polymerase binds at the UAS and then "slides" down the DNA molecule to the initiation site. In this model, *MAT*α2 binding between the UAS and the initiation site might repress transcription by blocking the polymerase from sliding. Such a mechanism would explain why binding at sites between the UAS and the TATA sequence are most effective at repressing gene expression. If this model is correct, then the mechanism employed by this yeast operator to repress *a*-specific genes is quite different from that of the classical procaryotic operators.

A similar repression system is probably utilized by *a*/*a* diploids to repress haploid-specific genes (Miller et al., 1985). In this case the repressor activity requires both the *MAT*α2 and the *MATa1* gene products. Examination of the sequences upstream of several haploid-specific genes has revealed a repeated sequence that is similar but distinct from the *MAT*α2 binding site. Furthermore it has been shown by Miller et al. that several such repeats mediate the repression of the *HO* gene in *a*/α cells (Miller et al., 1985).

The yeast "silencer." *Cis*-acting elements that repress transcription over even longer distances than the *STE6* operator have been discovered. The silent mating type cassettes of *Saccharomyces cerevisiae* (*HMR* and *HML*) are kept transcriptionally quiescent via the action of the four *trans*-acting *SIR* genes and two *cis*-acting "silencer" elements (*HMRE* and *HMLE*) (Haber and George, 1979; Klar et al., 1979; Rine et al., 1979). Brand et al. (1985) have reported that *HMRE*

may repress transcription of genes within *HMR* from a distance greater than 2600 bp. Like the *STE6* operator, *HMRE* can operate in either orientation relative to the repressed gene. Unlike the *a2/a1* repression system, the establishment of *SIR* repression mediated by *HMRE* requires DNA replication (Miller and Nasmyth, 1984). This fact along with the finding that *HMRE* repression is not blocked by intervening promoters or dependent on the transcriptional orientation of the repressed gene has led to speculation that this element acts by affecting chromatin structure over a long distance (Brand et al., 1985).

Other elements that repress transcription in *cis* are found near the yeast *MAT*α1 (Siliciano and Tatchell, 1984), *ADR2* (Russell et al., 1983) and *CYC1* genes (Guarente and Mason, 1983), the human β-interferon gene (Zinn et al., 1983) and within the enhancer element of murine sarcoma virus (Gorman et al., 1985). However, these elements have not yet been studied as extensively as those described above.

The P element intron. The mechanisms by which *cis*-acting negative elements regulate gene expression are not confined to the repression of transcription. Recent studies on the P transposable element of *Drosophila melanogaster* showed that the expression of its transposase activity can be blocked at the level of RNA processing. Within the P element there are four long open reading frames (ORF0, ORF1, ORF2, and ORF3) (O'Hare and Rubin, 1983). The integrity of each is required for the transposase activity. Site specific *in vitro*-synthesized mutations within any of these four ORFs fall into a single complementation group (Karess and Rubin, 1984). Laski et al. (1986) found that a major transcript from the P element, which contains all four ORF sequences, is produced in somatic tissues even though transposase activity is confined to the germline. In this somatic transcript only ORFs 0,1 and 2 are spliced into a single continuous reading frame. ORF3 is separated from the others by an intervening

sequence which is not spliced out. When this intervening sequence was removed from the DNA *in vitro* and the modified P element introduced into the *Drosophila* genome, transposase activity was expressed in somatic as well as germline tissues. Thus, the intervening sequence normally blocks transposase activity in somatic cells. In contrast to the *cis*-acting elements controlling transcription, which are recognized by *trans*-acting factors as part of a DNA molecule, this *cis*-acting element would be recognized by *trans*-acting factors, such as splicing enzymes, when it is in the form of an RNA molecule. It is also interesting to note that if such a *trans*-acting splicing factor exists in germline cells, it would act *positively* to allow this gene's expression by removing the *cis*-acting negative element (the intervening sequence.)

The *GAL81* locus. An example of a *cis*-acting negative element that interacts with its repressor while part of a protein molecule is provided by the *GAL81* locus of yeast. This locus is closely linked to the *GAL4* gene, which encodes a *trans*-acting positive regulator of the several galactose-utilization genes (Douglas and Hawthorne, 1966). *GAL4* activity is repressed in *trans* by the product of the *GAL80* gene. This repression has been shown to occur via an interaction with the *GAL4* protein rather than with the *GAL4* gene itself (Perlman and Hopper, 1979; Oshima, 1982). Mutations at *GAL81* make the *cis GAL4* allele insensitive to repression by *GAL80* and thus cause constitutive expression of the galactose utilization enzymes (Douglas and Hawthorne, 1972; Nogi et al., 1977). High resolution mapping of 15 *GAL81* mutations has placed all 15 in a small region which is within the *GAL4* gene itself (Oshima, 1982). That *GAL81* mutations reside within the *GAL4* protein coding region is shown by the fact that the *GAL81* mutations map between *gal4* nonsense mutations (Oshima, 1982).

Therefore, despite the fact that *gal4* and *GAL81* mutations have different phenotypic effects, they both appear to alter the same polypeptide. Since *GAL81*

mutations have the effect of making *GAL4* insensitive to *GAL80*, these mutations define a domain of the *GAL4* protein that is required for the interaction with the *GAL80* protein. It has been suggested that the *GAL81* locus is a site at which the *GAL80* protein binds to the *GAL4* protein to repress its activity (Perlman and Hopper, 1979; Oshima, 1982).

Other elements. From the handful of *cis*-acting negative regulatory elements which I have discussed above, it is easy to see that eucaryotic organisms use a number of different mechanisms to repress gene activity. Many more of these elements are likely to be found since a number of systems exist in which genetic elements that have a normal function of repressing gene expression have been identified (for examples, see Baker and Ridge, 1980; Henry et al., 1984; Hodgkin, 1980; Ingham, 1984; Karch et al., 1985; Lewis, 1978; Metzenberg and Chia, 1979; Moscoso del Prado and Garcia-Bellido, 1984; Oshima, 1982).

Before I began my thesis work, genetic analysis had shown that *Beadex* (*Bx*) mutations behave as though they increase the level of wild-type gene activity. This increase in activity could be explained in one of two ways: 1) *Bx* mutations are lesions that introduce a new *cis*-acting positive element near the structural gene, which increases its level of activity. 2) *Bx* mutations inactivate sequences that normally act in *cis* to repress the activity of the structural gene.

In the following sections I will review the genetic evidence that *Bx* mutations act in *cis* to cause an increase in gene activity. Following this review I will provide an overview of my molecular analysis of the *Bx* locus. In this overview, I will describe the evidence that *Beadex* mutations affect a negative regulatory element that probably functions at the protein level.

GENETIC ANALYSIS OF THE BEADEX LOCUS

Mutations at the *Beadex* locus, position 59.4 on the X-chromosome of *Drosophila melanogaster* cause a loss of tissue from the adult wing margin. In weak alleles this "wing scalloping" phenotype affects only the posterior margins, but in strong alleles over half of the wing is missing. A clonal analysis using the Bx^J mutation has revealed that this tissue loss results from the death of a population of cells within the early third instar wing disc (Santamaria and Garcia-Bellido, 1975). Morphological observations and staining of wing discs from mutant larvae have confirmed that there are 20-40 degenerating or dead cells in an area of the disc that gives rise to the wing margin (D. Fristrom, 1969; Waddington, 1940). Therefore, wing scalloping is likely to result directly from the death of those cells that are ancestral to the cells that form the adult wing margin. Nothing is known about the biochemical events that cause this cell death; however, a good deal of information exists regarding the genetic nature of *Bx* mutations.

Both dominant and recessive *Bx* mutations have been isolated (see Table 1). All of the known mutations are fully viable and fertile, when homozygous or hemizygous. Dominant mutations do not result from a haplo-insufficiency of Bx^+ activity, since females which are heterozygous for a deficiency which removes the Bx locus ($Df(Bx)/Bx^+$ where $Df(Bx)$ is any chromosome in which the *Bx* locus is deleted) have wild-type wings. Nor are recessive mutations associated with loss of gene function since females which carry these mutations over $Df(Bx)$ are wild type.

In an elegant genetic analysis M.M. Green showed that the *Bx* phenotype may result either from an increase in gene dosage or from mutations at the *Bx* locus itself (Green, 1953a). Green observed that females of the genotype Bx^1/Bx^r (Bx^1 is a weak dominant mutant allele and Bx^r is a phenotypically recessive

Table 1

mutant alleles	phenotype[1]	level of hyperactivity[2]
Bx^{1}	slight	3X
Bx^{2}	moderate	4X
Bx^{3}	moderate	4X
Bx^{J}	severe	6X
Bx^{46}	moderate	4X
Bx^{M}	slight	3X
Bx^{9}	severe	6X
Bx^{15}	severe	6X
Bx^{MH6}	severe	6X
Bx^{MH25}	moderate	5X
Bx^{MH26}	severe	6X
Bx^{MH32}	moderate	4X
Bx^{MH48}	very slight	2X
$Bx^{3.02.1}$	severe	6X

[1]The phenotype of homozygous females or hemizygous males.

[2]As measured by doses of Bx^{+}.

mutation), which are expected to give only Bx^1/Y or Bx^r/Y male progeny, occasionally give males which have stably reverted to wild-type (Bx^{rev}) and males which have a more severe wing scalloping phenotype than either of the expected types (Bx^e). Both types of exceptional males were always associated with recombination in the region around the *Bx* locus. Through a meticulous series of experiments, Green was able to show that the Bx^r mutation is a tandem duplication that contains two wild-type *Bx* alleles (symbolized $Bx^+.....Bx^+$). This tandem duplication is large enough to include visible markers on either side of the *Bx* locus (Green, 1952). Examination of polytene chromosomes from Bx^r individuals reveals that sections 17A, B, and C are entirely duplicated (Green, 1953a; Lifshytz and Green, 1979). The *Bx* locus is located at 17C1 (this thesis). In contrast to Bx^r, the Bx^1 mutation behave not as a tandem duplication but rather as a point mutation at the *Bx* locus itself. The Bx^{rev} and the Bx^e alleles were found to be the reciprocal products of recombination between the Bx^1 and Bx^r chromosomes at points within the duplicated interval. Such events yield $Bx^1....Bx^+$/Y (Bx^e) and Bx^+/Y (Bx^{rev}) individuals.

Green's work with another recessive *Bx* mutation, Bx^{r49k}, provided more information about the nature of *Bx* mutations. Like Bx^r, this mutation is caused by a tandem duplication which includes the normal Bx^+ locus (Green, 1953b). Unlike Bx^r, however, this allele undergoes unequal crossing over to produce triplications (symbolized as $(Bx)_3$) which can, in turn, produce quadruplications (symbolized as $(Bx)_4$) by unequal crossing over with Bx^{r49k}. Both the triplications and the quadruplications produced from Bx^{r49k} are dominant to wild-type and confer a more severe wing scalloping phenotype than Bx^{r49k} itself. By examining various genotypes Green found that the relative severity of the *Bx* phenotype is correlated with the number of Bx^+ copies in the mutant chromosomes so that females homozygous for a duplication (Bx^r/Bx^r) are less severely scalloped than

females that carry a duplication and a triplication ($Bx^r/(Bx^+)_3$), which are in turn less severe than individuals which are homozygous for a triplication ($(Bx)_3/(Bx)_3$). This dosage effect can be interpreted in one of two ways: 1.) The *Bx* mutant phenotype is the result of a quantitative increase in the level of Bx^+ gene activity. In this case, more copies of the Bx^+ locus produce more than the normal amount of Bx^+ activity, and this abnormally high level of activity causes wing scalloping. 2.) The novel position of those Bx^+ genes, which are located in the duplications, triplications and quadruplications, causes a qualitative change in their function. In this case the wing scalloping phenotype is a consequence of that abnormal activity. More copies of abnormally positioned Bx^+ genes would lead to a greater level of abnormal *Bx* activity and thus to more severe wing scalloping.

These two possibilities were distinguished by comparing females bearing a triplication on one homologue and either a normal Bx^+ chromosome or a deficiency, which removes the *Bx* locus on the other. If the level of Bx^+ activity determines the severity of wing scalloping, then $(Bx^+)_3/Df(Bx)$ females will have less severely scalloped wings than $(Bx^+)_3/Bx^+$ females since the latter genotype has more doses of Bx^+. On the other hand, if the wing scalloping phenotype is the result of some qualitative change in *Bx* activity produced by those copies of Bx^+ that are in the $(Bx^+)_3$ chromosome, then the wing phenotype should be identical in these two types of females since the number of abnormally functioning *Bx* genes is the same in both cases.

Lifshytz and Green observed that $(Bx^+)_3/Df(Bx)$ individuals have wild-type wings whereas $(Bx^+)_3/Bx^+$ have scalloped wings (Lifshytz and Green, 1979). They also observed that quadruplications behave in a similar manner. $(Bx^+)_4/Df(Bx)$ flies have slightly scalloped wings, while $(Bx^+)_4/Bx^+$ flies are more severely affected. These results show that the additional activity provided by a Bx^+ gene that is in its normal position (e.g., the one on the Bx^+ chromosome) can contribute

to the wing scalloping phenotype. Therefore, elevated levels of Bx^+ activity can cause wing scalloping, and no abnormal *Bx* activity needs to be invoked to explain the mutant phenotype of flies with extra doses of Bx^+.

How much Bx^+ activity is required to cause wing scalloping? Females that are homozygous for the Bx^r duplication have scalloped wings, whereas females with only one duplication chromosome (Bx^r/Bx^+) have normal wings. Thus, in females, which normally carry two doses of Bx^+, four doses are sufficient to cause a *Bx* phenotype, whereas three doses are not. If we assume that the level of *Bx* activity is proportional to the number of gene doses, then twice the normal amount of this activity would be sufficient to cause a wing scalloping phenotype.

Although these experiments argue that the wing scalloping that is exhibited by flies with extra doses of the Bx^+ locus results from elevated levels of wild-type gene activity, they do not address the cause of the wing scalloping which occurs in flies with *Bx* point mutations. In all known cases such mutations are dominant, although some are only weakly so (i.e., Bx^{MH48}). Homozygous females are more severely scalloped than heterozygotes ($Bx^1/Bx^1 > Bx^1/Bx^+$). The reader will recall that these mutations cannot be haplo-insufficient, since $Df(Bx)/Bx^+$ females have wild-type wings. Nor do these mutations appear to alter the activity of the *Bx* gene in a qualitative sense. If this were the case, then the genotypes $Bx/Df(Bx)$ and Bx/Bx^+ (where "*Bx*" symbolizes any dominant point mutation) would have similar levels of wing scalloping. In fact, for every *Bx* point mutation that has been examined, the genotype Bx/Bx^+ is more severely affected than $Bx/Df(Bx)$ (Lifshytz and Green, 1979). Together these results imply that *Bx* point mutations are similar to *Bx* hyperploids; they cause a quantitative increase in the level of wild-type gene activity. This notion is given further support by the observation that additional doses of Bx^+ will enhance the mutant phenotype of point mutations. For instance, $Bx^1....Bx^+/Y$ is more severe in phenotype than Bx^1/Y.

Furthermore, two Bx^+ genes have a stronger enhancing effect than one does ($Bx^1/Bx^+....Bx^+$ is more severe than Bx^1/Bx^+). These observations show that the activity of the mutated *Bx* genes and the wild-type genes are additive. This is consistent with the assertion that point mutations increase the level of the wild-type gene's activity and is inconsistent with the idea that mutations qualitatively change the *Bx* gene's function.

The apparent level of hyperactivity caused by *Bx* point mutations can be estimated by comparing the phenotypes of flies carrying these mutations to those of flies bearing multiple doses of Bx^+. For example, Bx^1/Bx^+ flies are similar in phenotype to $(Bx)_3/Bx^+$ flies. From this we may say that the Bx^1 mutation produces three times the normal level of Bx^+ activity. According to this criterion, the strongest known *Bx* point-mutant alleles produce over six times the normal amount of Bx^+ activity. The phenotypes and approximate activity levels of a number of *Bx* point mutations are presented in Table 1.

The fact that *Bx* mutations behave as though they cause quantitative effects on gene activity rather than alterations in gene function suggests that these mutations do not affect the structural gene itself, but rather affect regulatory elements which function to control the activity of the structural gene. If this is true, then it should be possible to find mutations that inactivate the structural gene and lower the level of Bx^+ activity. Although it would be difficult to predict a phenotype for such mutants, they should act to suppress the wing scalloping phenotype of hypermorphic *Bx* mutations in *cis*.

Lifshytz and Green searched for mutations in the *Bx* structural gene by mutagenizing flies which carry either the Bx^3 or Bx^J mutations and then collecting any progeny in which there was no wing scalloping. In this way they were able to isolate several point mutations, which acted in cis to suppress each of these strong *Bx* mutations. Each of the suppressor mutations were associated

with defects in the *heldup-a* (*hdp-a*) gene, which is located near the *Bx* locus. Loss-of-function mutations at the *hdp-a* gene cause flies to hold their wings in an abnormal upright position. Each of these suppressor mutations exhibited this phenotype. By doing fine genetic mapping Lifshytz and Green were able to show that a particular *hdp-a* mutation maps within a very short distance (0.0045 map units) of the Bx^3 mutation. This short distance between *hdp-a* and *Bx* as well as the coincidence of *Bx* suppression with *hdp-a* mutations suggests that a single mutational event within the *hdp-a* gene also affects the *Bx* structural gene.

What does *hdp-a* have to do with *Bx*? Lifshytz and Green proposed a model wherein the *hdp-a* and *Bx* loci make up a single bipartite genetic unit. Specifically, they suggested that the $hdp\text{-}a^+$ gene *is* the *Bx* structural gene and that the *Bx* locus itself regulates $hdp\text{-}a^+$ activity in ***cis***. In this model, the wing phenotype is determined exclusively by the level of $hdp\text{-}a^+$ activity. *Bx* mutations are hypothesized to disrupt or alter the *Bx* regulatory function and thus increase the level of $hdp\text{-}a^+$ activity. In this case, it is the elevated level of $hdp\text{-}a^+$ activity that causes wing scalloping in flies with a *Bx* mutation. *hdp-a* mutations eliminate or reduce the level of $hdp\text{-}a^+$ activity. This, they propose, leads to the aberrant wing position observed in flies carrying this mutation. Flies that carry both *Bx* and *hdp-a* mutations in the ***cis*** configuration do not exhibit wing scalloping because the loss of function mutation in *hdp-a* is epistatic to the gain-of-function mutation at the *Bx* locus.

Because it is quite likely that the *hdp-a* locus comprises the *Bx* structural gene and that the *Bx* locus itself is a regulatory element for *hdp-a* I will refer to the entire genetic unit from here on as *Bx-hdp-a*. It should be noted at this point that each of the cytologically visible duplications and deficiencies that were used by Green to study the effects of *Bx* dosage also affects the dosage of the $hdp\text{-}a^+$ gene in the same way. Therefore, the interpretation of these experiments is not

changed if *hdp-a* is the structural gene for *Bx* (increased or decreased dosage of Bx^+ is always accompanied by a similar increase or decrease in the dosage of $hdp\text{-}a^+$). Nor is the interpretation of the interaction between *Bx* mutations and extra doses of Bx^+ changed, since the level of $hdp\text{-}a^+$ is also increased when these extra doses are added.

LITERATURE CITED

Baker, B.S. and K.A. Ridge. 1980. Sex and the single cell. I. On the action of major loci affecting sex determination in *Drosophila melanogaster*. Genetics 94:383-423.

Brand, A.H., L.Breeden, J. Abraham, R. Sternglanz and K. Nasmyth. 1985. Characterization of a "Silencer" in yeast: A DNA sequence with properties opposite to those of a transcriptional enhancer. Cell 41:41-48.

Brent, R. 1985. Repression of transcription in yeast. Cell 42: 3-4.

Douglas, H.C. and D.C. Hawthorne. 1966. Regulation of genes controlling synthesis of galactose pathway enzymes in yeast. Genetics 54:911-916.

Douglas, H.C. and D.C. Hawthorne. 1972. Uninducible mutants in the *gal i* locus of *Saccharomyces cerevisiae*. J. Bacteriol. 109:1139-1143.

Fristrom, D. 1969. Cellular degeneration in the production of some mutant phenotypes in *Drosophila melanogaster*. Mol. Gen. Genet. 103:363-379.

Gorman, C.M., P.W.J. Rigby and D.P. Lane. 1985. Negative regulation of viral enhancers in undifferentiated embryonic stem cells. Cell 42:519-526.

Green, M.M. 1952. The *Beadex* locus in *Drosophila melanogaster*: The genotypic constitution of Bx^r. Proc. Natl. Acad. Sci. USA 38:949-953.

Green, M.M. 1953a. The *Beadex* locus in *Drosophila melanogaster*: On the nature of the mutants Bx^r and Bx^1. Genetics 38:91-105.

Green, M.M. 1953b. The *Beadex* locus in *Drosophila melanogaster*: Genetic analysis of the mutant Bx^{r49k}. Zeitschrift fur indukt. Abstammungs- and Vererbungslehre. 85:435-449.

Guarente, L. and T. Mason. 1983. Heme regulates transcription of the *CYC1* gene of *Saccharomyces cerevisiae* via an upstream activation site. Cell 32:1279-1286.

Haber, J.E. and J.P. George. 1979. A mutation that permits the expression of normally silent copies of mating type information in *Saccharomyces cerevisiae*. Genetics 93:13-35.

Hansen, U., D.G. Tenen, D.M. Livingston and P.A. Sharp. 1981. T antigen repression of SV40 early transcription from two promoters. Cell 27:603-612.

Henry, S.A., L.S. Klig and B.S.Loewy. 1984. The genetic regulation and coordination of biosynthetic pathways in yeast: Amino acid and phospholipid synthesis. Ann. Rev. Genet. 18:207-231.

Hodgkin, J. 1980. More sex-determinatation mutants of *Ceanorhabditis elegans*. Genetics 96:649-664.

Ingham, P.W. 1984. A gene that regulates the *bithorax* complex differentially in larval and adult cells of *Drosophila*. Cell37:815-823.

Johnson, A.D. and I. Herskowitz. 1985. A repressor (MATα2 product) and its operator control expression of a set of cell type specific genes in yeast. Cell 42:237-247.

Karch, F., B. Weiffenbach, M. Peifer, W. Bender, I. Duncan, S. Celniker, M. Crosby and E.B. Lewis. 1985. The abdominal region of the *bithorax* complex. Cell 43:81-96.

Karess, R.E. and G.M. Rubin. 1984. Analysis of P transposable element functions in *Drosophila*. Cell 38:135-146.

Klar, A.J.S., S. Fogel and K. Macleod. 1979. MAR1 a regulator or HMa and HMα loci in *Saccharomyces cerevisiae*. Genetics 93, 37-50.

Laski, F., D.C. Rio and G. M. Rubin. 1986. Tissue specificity of *Drosophila* P element transposition is regulated at the level of mRNA splicing. Cell 44:7-19.

Lewis, E.B. 1978. A gene complex controlling segmentation in *Drosophila*. Nature 276:565-570.

Lifshytz, E. and M.M. Green. 1979. Genetic identification of dominant overproducing mutations: The *Beadex* gene. Mol. Gen. Genet. 171:153-159.

Metzenberg, R.L. and W. Chia. 1979. Genetic control of phosphorus assimilation in *Neurospora crassa*: Dose dependent dominance and recessiveness in constitutive mutants. Genetics 93:625-643.

Miller, A., V.L. MacKay and K.A. Nasmyth. 1985. Identification and comparison of two sequence elements that confer cell-type specific transcription in yeast. Nature 314:598-602.

Miller, A.H. and K. Nasmyth. 1984. Role of DNA replication in the repression of silent mating type loci in yeast. Nature 312:247-251.

Moscoso del Prado, J. and A. Garcia-Bellido. 1984. Genetic regulation of the *Achaete-scute* complex of *Drosophila melanogaster*. Roux's Arch. Dev. Biol. 193:242-245.

Muller, H.J. 1932. Further studies on the nature and causes of gene mutations. Proc. 6th Int. Cong. Genet. 1:213-255.

Myers,R., D.C. Rio, A.K. Robbins and R. Tjian. 1981. SV40 gene expression is modulated by the cooperative binding of T antigen to DNA. Cell 25:373-384.

Nogi, Y.K., K. Matsumoto, A. Toh-e and Y. Oshima. 1977. Interaction of super-repressible and dominant constitutive mutations for the synthesis of galactose pathway enzymes in *Saccharomyces cerevisiae*. Mol. Gen. Genet. 152:137-144.

O'Hare, K. and G.M. Rubin. 1983. Structures of P transposable elements and their sites of insertion and excision in the *Drosophila melanogaster* genome. Cell 34:25-35.

Oshima, Y. 1982. Regulatory circuits for gene expression: The metabolism of galactose and phosphate. pp. 159-180 in *The molecular biology of the yeast Saccharomyces*. J.N. Strathern, E.W. Jones and J.R. Broach eds., Cold Spring Harbor Laboratories, Cold Spring Harbor, New York.

Perlman, D. and J.E. Hopper. 1979. Constitutive synthesis of the *GAL4* protein, a galactose pathway regulator in *Saccharomyces cerevisiae*. Cell 16:89-95.

Ptashne, M., A. Jeffrey, A.D. Johnson, R. Maurer, B.J. Meyer, C.O. Pabo, T.M. Roberts and R.T. Sauer. 1980. How the λ repressor and *cro* work. Cell 19:1-11.

Reed, S.I., J. Ferguson, R.W. Davis and G.R. Stark. 1975. T antigen binds to simian virus 40 DNA at the origin of replication. Proc. Natl. Acad. Sci. USA 72:1605-1609.

Rine, J., J. Strathern, J. Hicks and I. Herskowitz. 1979. A suppressor of mating type locus mutations in *Saccharomyces cerevisiae*: Evidence for and identification of cryptic mating type loci. Genetics 93:837-901.

Rio, D., A. Robbins, R. Myers and R. Tjian. 1980. Regulation of simian virus 40 early transcription *in vitro* by purified tumor antigen. Proc. Natl. Acad. Sci. USA 77:5706-5710.

Russell, D.W., M. Smith, D. Cox, V.M. Williamson and E.T. Young. 1983. DNA Sequences of two yeast promoter-up mutations. Nature 304:652-654.

Sambrook, J., B. Sugden, W. Keller and P.A. Sharp. 1973. Transcription of Simian Virus 40.III. Mapping of "early" and "late" species of RNA. Proc. Natl. Acad. Sci. USA 70:3711-3715.

Santamaria, P. and A. Garcia-Bellido. 1975. Developmental analysis of two wing scalloping mutants ct^6 and Bx^J of *Drosophila melanogaster*. Wilhelm Roux Arch. Entwicklungsmech. Organ. 178:233-245.

Siliciano, P.G. and K. Tatchell. 1984. Transcription and regulatory signals at the mating type locus in yeast. Cell 37:969-978.

Sprague, G.F., L.C. Blair and J. Thorner. 1983. Cell interactions and regulation of cell type in the yeast *Saccaromyces cerevisiae*. Ann. Rev. Microbiol. 37:623-660.

Strathern, J., J. Hicks and I. Herskowitz. 1981. Control of cell type in yeast by the mating type locus: the α1-α2 hypothesis. J. Mol. Biol. 147:357-372.

Struhl, G. 1981. A gene product required for correct initiation of segmental determination in *Drosophila*. Nature 293:36-41.

Tegtmeyer, P., M. Schwartz, J.K. Collins and K. Rundell. 1975. Regulation of tumor antigen synthesis by simian virus 40 gene A. J.Virol. 16:168-178.

Tjian, R. 1978. The binding site on SV40 DNA for a T antigen related protein. Cell 13, 165-179.

Tjian, R. 1981. T-antigen binding and the control of SV40 gene expression. Cell 26:1-2.

Waddington, C.H. 1940. The genetic control of wing development in *Drosophila*. J. Genetics 41:75-139.

Zinn, K., D. DiMaio and T. Maniatis. 1983. Identification of two distinct regulatory regions adjecent to the human β interferon gene. Cell 34:865-879.

OVERVIEW OF THESIS

Molecular Analysis of the *Beadex* and *heldup-a* loci

Although the genetic analysis performed by Lifshytz and Green (1979) led them to form a simple model that explains the interaction between the *Bx* and *hdp-a* loci, several critical issues regarding the molecular basis of this interaction were not addressed by these investigators. For example, even though their genetic experiments show that *Bx* mutations cause an increase in *hdp-a*$^+$ gene activity, these experiments do not indicate what the normal function of the *Bx* locus is. One possibility is that the wild-type *Bx* locus functions as a negative regulatory element for the *hdp-a* gene. In this case *Bx* mutations would disrupt this regulatory function and thus relieve the *hdp-a* gene from repression. Alternatively, it is possible that *Bx* mutations result from events that insert or create a new positive regulatory element near the *hdp-a* gene and act to increase its activity. In this case the normal *Bx* locus would merely be a site at which such events occur and would not necessarily have any regulatory function in the wild-type state.

Another critical ambiguity concerns the level at which *hdp-a*$^+$ activity is increased in *Bx* mutants. It is possible that *Bx* mutations affect any one of several steps involved in the synthesis and function of a hypothetical *hdp-a*$^+$ gene product. For instance, *Bx* mutations could increase the amount of the *hdp-a*$^+$ gene product synthesized by increasing the rate at which *hdp-a*$^+$ sequences are transcribed or translated. Alternatively, *Bx* mutations might destroy a nuclease recognition site within the *hdp-a*$^+$ RNA product. This might in turn lead to a decrease in the rate at which this RNA would decay and consequently to an increase in its amount.

Instead of causing a net increase in the amount of *hdp-a*$^+$ gene product, *Bx* mutations might cause an increase in the specific activity (i.e., the activity per molecule) of the *hdp-a*$^+$ gene product. This could occur, for instance, if the *Bx* locus encodes a repressor binding site within the *hdp-a*$^+$ protein. In such a case *Bx*

mutations would disrupt repressor binding and the *hdp-a*$^{+}$ product might become constitutively active in the same manner that the *GAL4* protein is made constitutive by *GAL81* mutations (Oshima, 1982). Another way that *Bx* mutations might increase the *hdp-a*$^{+}$ product's specific activity is via structural changes that increase the efficiency with which this product functions (e.g., an amino acid substitution that results in an increase in the V_{max} of an enzyme).

These issues can be most directly addressed by isolating recombinant DNA clones containing the sequences that make up the *Bx* and *hdp-a* loci. Such genomic clones can be used to analyze both the structural changes associated with *Bx* and *hdp-a* mutations and the effects that mutations have on the quantity and structure of any gene products encoded at these loci. This thesis presents such a molecular analysis.

In the first chapter I describe the isolation of recombinant clones that contain the chromosomal DNA surrounding the *Bx*^{2} locus and a set of clones that contain homologous sequences from wild-type DNA. Using these clones I show, in Chapter 1 and Appendix 1, that each of 13 different *Bx* mutations examined have structural alterations within a small portion of the cloned region. These structural alterations are of two types: 1.) insertions that are associated with spontaneous *Bx* mutations and 2.) deletions that are associated with alleles induced by P-M hybrid dysgenesis (Engels and Preston, 1981). Examination of the insertion alleles reveals that several of the inserted DNAs have restriction maps which are similar to those found in known retrovirus-like transposable elements (copia, roo, gypsy and 3S18). In the *Bx*^{2} strain two closely spaced DNA insertions have been identified as gypsy transposons. At least one of these insertions is probably the cause of the *Bx*^{2} mutation since *Bx*^{2} is suppressed by the third chromosome mutation, ***suppressor of Hairy wing***, which is known to specifically repress gypsy insertion mutations. The DNA insertions in each of the six

spontaneous mutant lines that I examined disrupt a single 500 bp segment (in DNA from the Bx^2 strain one gypsy element is inserted in this segment and the other is inserted at a site about 1 kb away). This suggests that the integrity of sequences within this segment is critical to the expression of a normal wing phenotype.

Examination of seven hybrid dysgenesis induced mutations shows that these alleles arise via imprecise excision of a P element located near the *Bx* locus in the wild-type strain π_2. These excision events apparently remove the entire P element as well as 1200 bp to 2200 bp of sequences flanking it. The net result is a small deletion of the sequences surrounding the P element insertion site.

By mapping the positions of these deletion mutations I am able to define an 800 base pair segment in which some part of the wild-type *Bx* locus must reside. This segment is adjacent to the 500 bp segment disrupted by insertion mutations. Sequences within this 800 bp segment must have a negative regulatory function in wild-type flies, since several small deletions that remove them cause an increase in gene activity (i.e., the flies have a *Bx* mutant phenotype). I argue that, at least in the deletion mutations, the *Bx* mutant phenotype is not caused by the creation or insertion of a new *cis*-acting positive regulatory element at the *Bx* locus but rather by the loss of a *cis*-acting negative element.

From my deletion mapping results, I am also able to show that one strain that is phenotypically indistinguishable from other *hdp-a* mutations ($hdp\text{-}a^{D30r}$) and exhibits no wing scalloping has suffered a 1400 bp deletion, which must include part of the *Bx* locus as well as *hdp-a*. I interpret this to mean that the *Bx* mutation is suppressed in *cis* by loss of the $hdp\text{-}a^+$ function. This supports the model proposed by Lifshytz and Green (1979) and is consistent with my observation that this allele behaves in the same way that large deficiencies which remove both *Bx* and *hdp-a* do when heterozygous to *Bx* mutations (i.e., $Bx/hdp\text{-}a^{D30r}$ is less severely scalloped than Bx/Bx^+). Furthermore, by comparing the

hdp-a^{D30r} to other *Bx* deletion mutations I am able to conclude that some part of the *hdp-a* locus must reside inside a 200 bp segment which is closer than 1500 bp away from part of the *Bx* locus.

In Chapter 2 I identify a number of transcripts which are encoded within a 34 kb region surrounding the *Bx* locus. Two of these transcripts are found in close proximity to the region altered in *Bx* and *hdp-a* strains. These two poly(A)$^+$ transcripts are 4 kb and 2 kb in length. They are expressed at different developmental stages. The 2 kb transcript is expressed only in early embryos while the 4 kb species is first expressed in late embryonic stages and persists throughout both larval and pupal development. I find that these two transcripts are transcribed from the same DNA strand and are coextensive over a three kb region. Therefore, it seems probable that they overlap one another and encode related products.

The position of these two transcripts is consistent with the idea that they might encode the *hdp-a*$^+$ function. The main body of both transcripts is on the centromere-distal side of the *Bx* locus as is the *hdp-a* point mutation which was mapped by recombination by Lifshytz and Green (1979). Both transcripts overlap the 200 bp segment in which part of the *hdp-a* locus is located. Furthermore, the 4 kb transcript spans the 500 bp segment in which DNA insertions occur in *Bx* flies. At least some of the sequences encoded near the 3' end of this RNA originate from the 800 bp segment, defined by deletions, in which part of the *Bx* locus must reside. The 3' end of the 2 kb RNA appears to originate from sequences near to or just within this 800 bp segment.

To determine whether these or any other nearby transcripts are quantitatively or qualitatively affected by *Bx* mutations, I examined RNA isolated from several mutant strains. Among the transcripts encoded within the 34 kb region, I find that only the two transcripts described above are affected by *Bx* and

hdp-a mutations. Although little or no quantitative increase in the amount of any transcript was observed in RNA from the *Bx* mutant strains that we examined, probes that normally hybridize with the 2 kb and 4 kb RNAs were observed to hybridize with transcripts of abnormal size in almost every case. When late embryonic and third instar larval RNAs were examined, the 4 kb transcript was found to be replaced by shorter transcripts in each of five *Bx* mutants and one *hdp-a* mutant. During early embryonic stages the 2 kb transcript is replaced by abnormally sized transcripts in the *hdp-a* mutant and in four of the five *Bx* mutants examined. These aberrant early embryonic transcripts are distinct from those that replace the 4 kb transcript in later stages. In each case the abnormal size transcripts are expressed in the same developmental pattern as the wild-type transcript that they replace and in only one case is the steady-state level of the abnormal size transcript higher (two-fold) than the level of the normally size transcript expressed in wild-type strains.

Because most *Bx* mutations appears not to increase the quantity of RNA produced by any of the transcription units surrounding the *Bx* locus, it seems unlikely that the *Bx* locus plays a role in regulating the synthesis or decay of RNA. Instead, qualitative changes in the structure of transcripts that are encoded near the *Bx* locus suggests that *Bx* mutations increase gene activity by altering the gene product's structure. Structural changes could increase either the efficiency with which these RNAs are translated or the specific activity of the gene product encoded by them.

A closer examination of the structure of the mutant transcripts that replace the 4 kb RNA reveals that in each mutant the altered species is missing sequences that originate from the 800 bp region in which part of the *Bx* locus resides. As stated above, sequences from within this region are found in the wild type 4 kb transcript. In one case these sequences are absent from the mutant RNA simply

because they are deleted from the DNA. However, the other four mutant strains contain transposable DNA insertions which are located adjacent to and just upstream of the 800 bp region but do not disrupt it. Since the missing sequences originate from a site that is downstream from the sites of insertion, it seems quite likely that mutant RNA transcripts are terminated within these transposable DNA insertions. Indeed, none of the mutant transcripts are homologous to any of the sequences examined downstream from the insertion sites. Thus, transcripts from all of the deletion mutations and these four insertion mutations share a common characteristic--they lack sequences that are very near to or within the *Bx* locus.

On the basis of these results I propose that sequences encoded at the *Bx* and *hdp-a* loci are normally included within the 4 kb transcript and that *Bx* sequences function in *cis* to repress the activity encoded within this transcript at the level of protein synthesis or protein function. In *Bx* deletion mutants the *Bx* sequences are missing from the DNA and therefore cannot possibly be included in RNA or protein. In *Bx* insertion mutants, transcripts are terminated near DNA insertions which are located upstream from the *Bx* locus. Thus, both types of mutant transcripts lack sequence information from the *Bx* locus (the 200 bp region which contains part of the *hdp-a* locus is upstream of the insertions and would be contained in these transcripts). I suggest that the increased gene activity which is observed in *Bx* mutants is primarily a result of the absence of *Bx* sequences from the mutant RNA or the hypothetical protein product.

In Chapter 3 I describe germline transformation experiments designed to determine whether a 10.4 kb DNA fragment from the cloned region is capable of either complementing *hdp-a* mutations or contributing to the *Bx* phenotype of flies with multiple doses of Bx^{+}. This fragment contains the entire region which I know to hybridize with the 4 kb and 2 kb RNA species as well as over 4 kb of upstream sequences. Seven transformed lines were obtained. Of these seven, five were

shown to have acquired one copy of the intact 10.4 kb fragment inserted into the genome at new locations. However, all five insertions failed to complement a *hdp-a* mutation when they were present in a single dose. By combining various insertions I tested two, three and four doses of the 10.4 kb fragment for complementation of *hdp-a*. In no case was complementation observed. Bx^+ activity was tested in two ways: by examining the wing phenotype of Bx^+/Bx^+ females that carry 2 or more doses of the inserted fragment, and by examining the phenotype of Bx^r/Bx^+ females, which also carry one copy of the inserted fragment. Wing scalloping should be evident in either case if the inserted fragment behaves like the Bx^+ gene. In the first case no wing scalloping was observed even when as many as four doses of the inserted element were added. In the second case some wing scalloping was observed, but control crosses showed that this was only an effect of genetic background. Therefore, it appears that the 10.4 kb fragment has neither the ability to function as the $hdp\text{-}a^+$ or the Bx^+ gene (hypothetically these are the same).

There are at least two possible explanations for this unexpected finding:

1.) Neither the 4 kb nor the 2 kb transcript encode the $hdp\text{-}a^+$ function, or

2.) Some sequences necessary for expression of the 2 kb and/or the 4 kb transcripts (possibly exon sequences not detected by RNA blot hybridization) are located outside of the 10.4 kb fragment and therefore, even though it encodes $hdp\text{-}a^+$, one of these transcripts is not expressed. The results presented in Chapter Two make the former explanation seem unlikely, but a firm resolution of this question awaits a more extensive characterization of these transformant lines and the assay of more DNA sequences by transfomation.

LITERATURE CITED

Engels, W.R. and C.R. Preston. 1981. Identifying P factors in *Drosophila* by means of chromosome breakage hotspots. Cell 26:421-426.

Lifshytz, E. and M.M. Green. 1979. Genetic identification of dominant overproducing mutations: The *Beadex* gene. Mol. Gen. Genet. 171:153-159.

Oshima, Y. 1982. Regulatory circuits for gene expression: The metabolism of galactose and phosphate. pp. 159-180 in *The molecular biology of the yeast Saccharomyces*. J.N. Strathern, E.W. Jones and J.R. Broach, eds., Cold Spring Harbor Laboratories, Cold Spring Harbor, New York.

CHAPTER 1

Isolation and Characterization of the *Beadex* Locus of *Drosophila melanogaster*: a Putative *Cis*-acting Negative Regulatory Element for the *heldup-a* Gene

WILLIAM W. MATTOX AND NORMAN DAVIDSON

Molecular and Cellular Biology

4:1343-1353 (1984)

MOLECULAR AND CELLULAR BIOLOGY, July 1984, p. 1343–1353
0270-7306/84/071343-11$02.00/0

Vol. 4, No. 7

Isolation and Characterization of the *Beadex* Locus of *Drosophila melanogaster*: a Putative *cis*-acting Negative Regulatory Element for the *heldup-a* Gene

WILLIAM W. MATTOX[1] AND NORMAN DAVIDSON[2]*

Division of Biology[1] and Department of Chemistry,[2] California Institute of Technology, Pasadena, California 91125

Received 6 February 1984/Accepted 19 April 1984

We isolated recombinant lambda phage clones spanning 49 kilobases of DNA which contain the *Beadex* and *heldup-a* loci of *Drosophila melanogaster*. These cloned DNAs were used to analyze the structure of eight dominant mutant alleles of the *Beadex* locus which show increased gene activity. A region, only 700 base pairs in length, is altered in each of these mutants. Six of the mutations have DNA insertions within this segment. Most of these insertions resemble retrovirus-like transposable elements. In one case (*Beadex*^{2}) the inserted sequences are homologous to the gypsy transposon family. The other two *Beadex* alleles were induced by hybrid dysgenesis and suffered deletions which included at least part of the 700-base-pair segment. These deletions appear to have resulted from imprecise excision or deletion of a nearby P element found in the wild-type parental strain. Analysis of one *heldup-a* allele (*heldup-a*^{D30r}) indicates that a similar P element-mediated event is responsible for this lesion. In this mutant, deletion of sequences no more than 1,600 base pairs from the *Beadex* locus accompanies the loss of *heldup-a* function. The deleted sequences in *heldup-a*^{D30r} include the entire 700-base-pair segment within which at least part of the *Beadex* locus resides, yet these flies have no *Beadex* phenotype. This indicates that a functional *heldup-a* gene is necessary for expression of the *Beadex* phenotype. Together, these results suggest that the *Beadex* functional domain is contained within a short segment of DNA near the *heldup-a* gene and support the hypothesis that the *Beadex* locus functions as a *cis*-acting negative regulatory element for the *heldup-a* gene.

Very few eucaryotic loci have been identified where mutations cause abnormally high levels of gene activity and result in a visible phenotype (7, 18, 39). The 17BC segment of the *Drosophila melanogaster* X-chromosome contains such a locus. Simple duplications including this chromosomal segment cause a characteristic loss of tissue along the wing borders (wing scalloping) (16, 17). Triplications which include 17BC cause more severe wing scalloping, and flies with quadruplications are still more severely affected (16, 17, 24). Thus, the degree of wing scalloping appears to be related to the increased activity of some gene within this segment.

Although gross deficiencies which include the 17BC segment have no dominant effect on wing phenotype ($Df/+$ is wild type), several dominant wing-scalloping mutations have been mapped to the *Beadex* (*Bx*) locus which is located within this segment at map position 1-59.4 (see Table 1). The 17BC segment is not visibly altered or duplicated in any of these *Bx* alleles, suggesting that they are simple mutations at the *Bx* locus and are not associated with increased gene dosage (unpublished data). The phenotypic similarity of these *Bx* mutants to flies with multiple doses of the 17BC segment suggests that mutations at the *Beadex* locus lead to an increase in the wild-type activity of a gene within the 17BC segment. This idea is supported by the interactions of these mutations with deficiencies and duplications that include 17BC. Flies bearing a *Bx* mutation on one homologue and a duplication of the wild-type 17BC segment on the other have an enhanced *Beadex* phenotype, whereas deficiencies suppress *Bx* mutations. For instance, the weakly dominant mutant allele Bx^1 is totally suppressed by any deficiency which includes 17BC, but wing scalloping is enhanced by duplications including this segment ($Bx^1/Df < Bx^1/+ < Bx^1/Dp$) (16, 17, 24). Strong *Bx* alleles are only partially suppressed by deficiencies (24). These observations strongly suggest that the *Bx* locus has a regulatory function and that mutations at this locus lead to an increase in the normal activity of some nearby structural gene rather than a loss or qualitative alteration of the activity of that gene product.

Mutations which result in loss of *Bx* activity have been isolated as phenotypic revertants and suppressors of *Bx* mutant alleles (24). These mutations fall into two classes: gross deletions including the 17BC region and point mutations in the closely linked *heldup-a* (*hdp-a*) gene (see Table 1), which is also within 17BC. These *hdp-a* mutations completely suppress even strong *Bx* alleles in *cis* and mimic the partial suppression of *Bx* exhibited by deficiencies in *trans*. *hdp-a* mutations themselves are associated with a recessive heldup wing position phenotype (25). Lifschytz and Green (24) showed that the *Bx* and *hdp-a* loci are separated by only 0.0045 map units and that one *hdp-a* allele maps to the centromere distal side of Bx^3. They proposed a model, based on the interactions of these two loci, in which the wild-type *Bx* locus functions as a *cis*-acting negative regulatory element for the adjacent *hdp-a* structural gene (24). In this model mutations at the *Bx* locus would result in an increase in $hdp\text{-}a^+$ activity. The wing-scalloping phenotype observed in *Bx* mutants would be a result of this increased $hdp\text{-}a^+$ activity. Mutations in the *hdp-a* structural unit, which eliminate $hdp\text{-}a^+$ activity, result in a heldup wing position phenotype with no wing scalloping. In the *cis* double mutant (*hdp-a Bx*), disruption of the *hdp-a* gene would preclude any hyperactivity due to the *Bx* mutation. Thus *hdp-a* mutations suppress *Bx* alleles in *cis*.

A rigorous test of this model requires both quantitative and qualitative characterization of the *hdp-a* gene product in both *Bx* and *hdp-a* mutant flies. Here we report the isolation of DNA sequences corresponding to at least part of both the *Bx* and *hdp-a* loci. These cloned DNA sequences should

* Corresponding author.

allow us to identify any RNA molecules encoded by these genetic elements.

In the course of identifying these DNA sequences we studied the molecular structure of a number of *Bx* mutations and one *hdp-a* allele. We report here that all the *Bx* mutations we examined have alterations in a single 700-base-pair region. Six *Bx* alleles contain DNA insertions in this region, whereas two alleles, which were induced by hybrid dysgenesis, are the result of imprecise P element excisions or deletions which remove sequences from this segment. A single *hdp-a* mutation, which is associated with a similar hybrid dysgenic event, suggests that part of the *hdp-a* locus resides no more than 1,600 base pairs from the *Bx* locus.

MATERIALS AND METHODS

Fly stocks. Mutant fly stocks were obtained from The *Drosophila* Stock Center at the California Institute of Technology, The Mid-America *Drosophila* Stock Center at Bowling Green University, W. Baker at the University of Utah, and W. Engels at the University of Wisconsin. The *Bx* and *hdp-a* mutant strains used in this work are described in Table 1.

In situ hybridizations to polytene chromosomes. *Drosophila melanogaster* salivary gland squashes were prepared as described by Gall and Pardue (15) except that the acid treatment was omitted. Each squash was hybridized with 10^6 cpm of [^{3}H]cRNA in 50% formamide (deionized)–0.75 M sodium chloride–0.075 M sodium citrate–5 mg of calf liver tRNA per ml (Sigma Chemical Co.). Slides were washed, autoradiographed, and stained with Giemsa solution (Accra-Labs, Inc.) as described by Gall and Pardue (15).

Construction of cosmid libraries. Ten micrograms of high-molecular-weight *Drosophila* genomic DNA was digested in separate reactions with 0.005, 0.02, 0.04, 0.06, 0.10, and 0.16 U of *Mbo*I (New England Biolabs) per μg for 12 min at 37°C (60 μg total). The reaction with the lowest amount of enzyme yielded *Mbo*I fragments with an average size greater than 50 kilobases (kb), whereas the reaction containing the largest amount yielded fragments averaging ca. 10 kb in length. These digests were pooled and digested with calf intestine alkaline phosphatase and then sedimented through a 10 to 40% linear sucrose gradient (1 M sodium chloride, 0.02 M Tris-hydrochloride [pH 8.0], 0.01 M EDTA) in a Beckman SW27.1 rotor at 22,000 rpm for 24 h at 20°C. Fractions containing 36- to 42-kb *Mbo*I fragments were pooled and dialyzed against 200 mM sodium chloride–10 mM Tris-hydrochloride (pH 7.6)–1 mM EDTA. This DNA was concentrated by ethanol precipitation and ligated with an equal mass (6 μg) of pJB8 DNA, which was prepared as described by Ish-Horowicz and Burke (21). Cosmid DNA was packaged in vitro by using the protocol of Mullins et al. (30). We obtained a total of about 2×10^5 ampicillin-resistant colonies (equivalent to 46 *Drosophila* haploid genomes) when these packaged cosmids were transfected into *Escherichia coli* HB101. The inserts in 12 randomly selected clones ranged in size from 31 to 41 kb and averaged 37 kb.

We found that omission of the size selection step led to much poorer efficiencies (less than 10^2 colonies per μg of target DNA), presumably due to the molar excess of small *Mbo*I fragments which ligated with PJB8 to form molecules too small to package.

Screening of phage and cosmid genomic libraries. The Bx^2 cosmid library was screened by the colony hybridization method of Hanahan and Meselson (20) with the modifications added by Grosveld et al. (19). The Canton-S Charon 4 phage library (26) was screened by the method of Benton and Davis (3).

Isolation of DNA. Cosmid and plasmid DNA were isolated as described by Steinmetz et al. (41) except that bacteria were grown in medium containing 100 mg of ampicillin per liter rather than tetracycline. Recombinant phage for DNA isolation were grown by the PDS method of Blattner et al. (6). Phage were precipitated with polyethylene glycol and sodium chloride as described by Yamamoto et al. (44). They were then purified by CsCl equilibrium centrifugation, and DNA was isolated as described by Maniatis and co-workers (26). *Drosophila* genomic DNA was isolated from adult flies as described by Davis and Davidson (8).

DNA blot hybridizations. DNA was digested with a five-fold excess of restriction endonucleases purchased from New England Biolabs or Boehringer Mannheim. After agarose gel electrophoresis, the digested DNA was transferred to nitrocellulose (40). Blots were incubated with denatured probe, washed, and autoradiographed as described by Rozek and Davidson (34). Radioactive probes were removed from blots by washing twice for 30 s with 0.015 M NaCl–0.0015 M sodium citrate (pH 7.5) at 100°C.

In vitro labeling of nucleic acids. [^{3}H]cRNA was prepared as described by Wensink et al. (43). DNA was nick translated as outlined by Rigby and co-workers (33). For the uniform labeling of *Bam*HI fragments from p24R8.0B, this plasmid was digested to completion with *Bam*HI and labeled with [^{32}P]dGTP by using the large fragment of DNA polymerase I (Boehringer Mannheim) (27).

Electron microscopy. *Sal*I-cut cBx^{2}17C-1 (0.1 μg) was denatured in 0.1 N NaOH for 30 s at 0°C and then neutralized with Tris-hydrochloride. This was diluted to 0.5 μg/ml in a solution containing 50% formamide, 0.1 mg of cytochrome *c* per ml, 100 mM Tris-hydrochloride (pH 8.5), and 10 mM EDTA and immediately spread for electron microscopy as described by Davis et al. (9).

TABLE 1. Mutant fly strains used to isolate DNA

Strain	Wing-scalloping phenotype	Origin (reference)
Bx^1	Weak	Spontaneous (25)
Bx^2	Moderate	Spontaneous (25)
Bx^4	Moderate	Spontaneous (14)
Bx^J	Strong	Heat induced (22)
Bx^{4b}	Moderate	Unknown
Inscy w Bx^M	Weak	Spontaneous in *Inscy w* found by Muller
Bx^9	Strong	Hybrid dysgenesis induced on the π_2 X-chromosome by W. Engels
Bx^{15}	Strong	Hybrid dysgenesis induced on the π_2 X-chromosome by W. Engels
hdp-a$^{D5ar^+}$	Wild	Hybrid dysgenesis induced on the π_2 X-chromosome by W. Engels

[a] Several different genes have been called *heldup* (10, 13, 24). Here we designate the suppressor of *Beadex* as *hdp-a* and the closely linked mutations described by Engels as *hdp-b*. *hdp-a* and *hdp-b* fall into separate complementation groups (unpublished data).

RESULTS

Strategy for molecular cloning of the *Beadex* locus. To isolate recombinant clones containing DNA from the *Bx* locus, we used a modification of the transposon-tagging approach first suggested by Bingham et al. (5). The mutant allele Bx^2 is suppressed in flies homozygous for a mutation at the unlinked suppressor of Hairy-Wing [*su(Hw)*] locus (23, 25). Since a variety of *Drosophila* mutations, which are

suppressed by *su(Hw)* possess DNA from the gypsy transposon family at the mutated locus (29), we examined the possibility that Bx^2 possesses a gypsy transposon at the *Bx* locus. Bx^2 polytene chromosomes were hybridized with [^{3}H]cRNA made from λbx^{34e}-6a2 recombinant phage DNA (a gift from W. Bender) in situ. This phage contains about 7 kb of unique sequences from the *bithorax* region, as well as one copy of the 7.3-kb gypsy transposon. The chromosome in Fig. 1a shows the expected hybridization at the 89E region of the third chromosome (*bithorax*), as well as hybridization to the 17C segment of the X-chromosome (*Beadex*). No other sites of localized hybridization were observed; however, a number of grains were consistently visible over the chromocenter. We presume that these grains represent hybridization to gypsy sequences in heterochromatin. Indeed, whole genome Southern blots with Bx^2 DNA revealed 6 to 10 different sites of gypsy insertion in the Bx^2 genome (data not shown).

To isolate *Bx* sequences by transposon tagging, clones containing the gypsy transposon from 17C must be identified. We took advantage of a P element, which is within polytene bands 17C2-3 in the strain π_2, to identify such clones. Genetic experiments indicate that this P element lies near the *Bx* locus (see below). Since the 17C2-3 P element from π_2 had already been isolated (31), a probe, pS25.1 containing only unique sequences from the 17C2-3 region was available. This plasmid contains a 1.8-kb *Bam*HI fragment, from the wild-type strain Canton-S, which is homologous to the DNA immediately flanking the 17C2-3 P element of π_2 but contains no P sequences. Any gypsy homologous recombinant clones from a Bx^2 genomic library which hybridize with this fragment must contain the gypsy element residing near 17C2-3.

Bx^2 has two gypsy elements inserted in opposite orientation at 17C2-3. A cosmid library containing Bx^2 genomic sequences was constructed by using the vector pJB8 (see

FIG. 1. In situ hybridizations with [^{3}H]RNA made from: (a) λbx^{34e}-6a2 and hybridized with Bx^2 polytene chromosomes, (b) λCS17C-19 and hybridized to Canton-S chromosomes, (c) λCS17C-2 and hybridized to homozygous *In(1)D6j* chromosomes, or (d) λCS17C-9 and hybridized to homozygous *In(1)D6j* chromosomes. Arrows point to sites of consistent hybridization. cc, Chromocenter.

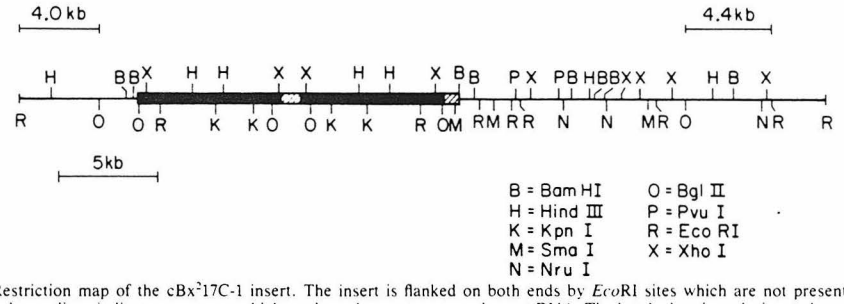


FIG. 2. Restriction map of the cBx²17C-1 insert. The insert is flanked on both ends by *Eco*RI sites which are not present in the Bx^2 genome. The heavy lines indicate sequences which are homologous to gypsy element DNA. The hatched regions designate homology to the 1.8-kb *Bam*HI fragment from pS25.1. The bars above the restriction map indicate fragments used as probes to isolate recombinant phage clones from the Canton-S Charon 4 library.

above). Three *Drosophila* genome equivalents were plated by using the *recA*⁻ host HB101. These clones were screened in parallel for homology to the 1.8-kb *Bam*HI fragment of pS25.1 and the 6.8-kb *Xho*I fragment which is internal to the gypsy element in λbx^{34e}-6a2. Since *Xho*I cuts only in the terminal direct repeats of the gypsy element, this 6.8-kb fragment contains all sequences within the transposon. Twenty-three positive clones were identified with the gypsy probe and only two with the 1.8-kb *Bam*HI fragment. Both of the latter were also among the gypsy homologous set. One of these clones was lost during subsequent rescreening; the other was isolated and designated cBx²17C-1. Cosmid DNA was prepared from this clone and digested with several restriction enzymes to generate the map shown in Fig. 2. cBx²17C-1 contains 42 kb of Bx^2 genomic DNA inserted at the *Bam*HI site of PJB8. Genomic blots with cloned Canton-S DNA from the same region as a probe confirmed that this map represents the configuration on the Bx^2 X-chromosome and is not an artifact of in vivo manipulation or growth in *E. coli* (see below).

Within cBx²17C-1 there are two regions, each ca. 7 kb in length, which have restriction patterns identical to those found in the gypsy element of λbx^{34e}-6a2. These two regions reside within 1 kb of one another and are oriented oppositely. Blot hybridization experiments with the 6.8-kb *Xho*I fragment from the gypsy element in λbx^{34e}-6a2 as a probe demonstrate that each of these two gypsy-like segments is homologous to gypsy DNA (data not shown). The regions of gypsy homology are indicated in Fig. 2. Hybridizations using the 1.8-kb *Bam*HI fragment insert of pS25.1 as a probe show that these unique sequences lie between and adjacent to the two gypsy elements found in cBx²17C-1 (data not shown). However, fragments which span the left end of the leftward gypsy element (as they are shown in Fig. 2) do not hybridize to the 1.8-kb fragment. Our interpretation is that the segment of homology within these fragments is too short to form a detectable hybrid (since the *Bam*HI site is less than 100 base pairs to the left of the left end of the gypsy element) and that both gypsy elements inserted into the 1.8-kb fragment.

To confirm the structure deduced from the blot hybridization experiments above, we linearized cBx²17C-1 with *Sal*I (the only *Sal*I site in this clone is within the vector, pJB8), melted the duplex in alkali, allowed it to self-anneal briefly at neutral pH (see above), and prepared the DNA for electron microscopy. A stem-loop structure was consistently observed. An example is shown in Fig. 3. This stem-loop structure is of the proportions predicted for the inverted repeat formed by the two gypsy elements, with a duplex stem of about 7.5 kb and a single-stranded loop of about 0.9 kb. No bubbles were observed in the duplex stem, indicating that the two gypsy elements are identical with one another at the electron microscope level of resolution.

Isolation of clones containing wild-type DNA from 17C2-3. To study the wild-type *Beadex* locus and to obtain single-copy probes for whole genome blotting experiments, we screened a random shear Charon 4 library constructed with Canton-S genomic DNA (26) by using the 4.0 and 4.4-kb *Bgl*II-*Eco*RI fragments from cBx²17C-1 (see Fig. 2). These fragments are at opposite ends of the cosmid insert, allowing isolation of clones covering a large region. From five genome equivalents screened with these ^{32}P-labeled fragments, seven different recombinant phage clones were isolated. DNA from each of these phages was mapped for restriction sites. The composite map is shown in Fig. 4a. A total of 49 kb of Canton-S DNA is represented in the seven overlapping phage inserts. Using these overlapping clones as probes in blotting experiments, we confirmed that the restriction map shown in Fig. 4a correctly represents the map of Canton-S genomic DNA (see Fig. 5 and 6). Using [^{3}H]cRNA made from λCS17C-19 as a probe in in situ hybridizations to Canton-S polytene chromosomes, we determined that these sequences localize uniquely at 17C, as expected (Fig. 1b).

Other than several simple restriction site polymorphisms, the map in Fig. 4 differs from the map of the cBx²17C-1 insert in only two respects: the absence of the gypsy transposons found in cBx²17C-1 (these would be inserted at

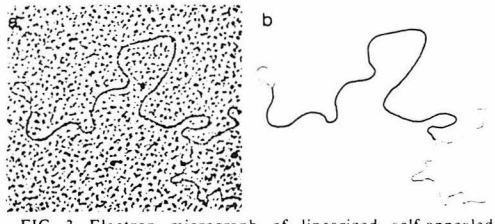


FIG. 3. Electron micrograph of linearized self-annealed cBx²17C-1 DNA. (a) Photograph. (b) schematic representation of photograph. Measurements on five such stem-loop structures were made by using pBR322 as a double-stranded size standard and ϕX174 DNA as a single-stranded standard. The average size was ca. 7.5 kb for the stem and about 0.9 kb for the single-stranded loop.

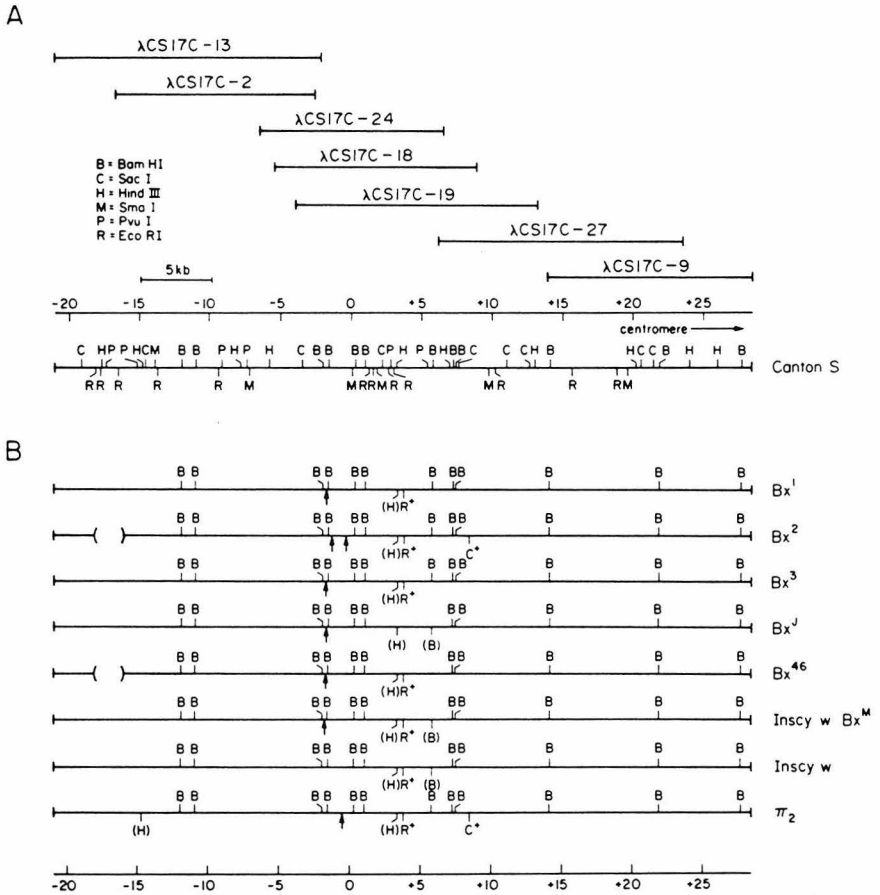


FIG. 4. Composite restriction maps of both wild-type and mutant strains from the *Beadex-heldup-a* region. (A) Restriction map of Canton-S was deduced from the maps of recombinant phage clone inserts and confirmed by blot hybridization by using Canton-S genomic DNA. The segment covered by each phage insert is shown above the restriction map. The *Sma*I site in the 1.8-kb *Bam*HI fragment was arbitrarily assigned coordinate position 0.0. Sequences on the centromere proximal side (right) of this site are given positive coordinates corresponding to their distance in kilobases from the 0.0 position. Sequences to the left are given negative coordinates in the same manner. (B) Composite maps of *Beadex* mutants and their parental strains deduced by blot hybridizations. The *Bam*HI sites are shown in each case for reference. Differences relative to Canton-S are indicated below the line in each case. Sites not present in the mutant strain are indicated with parentheses. Sites which are found in the mutant but not Canton-S are identified with a +. Maps of the insertions are not included, but their positions are indicated by the arrows. Maps for the inserted elements can be found in Fig. 2, 8, and 9. Note that the *Inscy w* and π_2 strains are both Bx^-.

coordinate positions −0.5 and −1.5 in Fig. 4) and a deletion of 50 to 100 base pairs somewhere between coordinates +1.4 and +2.0. This small deletion seems to be a peculiarity of the Canton-S strain, since whole genome blotting analysis of eight *Bx* mutant strains and two other Bx^- strains failed to reveal another example (see the 2.0-kb *Sma*I fragment in Fig. 5b).

Determination of the chromosomal orientation of cloned DNA. Genetic mapping experiments have placed the *Bx* locus a short distance to the proximal side of the *hdp-a* locus (24). Since mutations at these two loci are known to interact in *cis*, we were interested in determining the molecular structure of both *hdp-a* and *Bx* functional domains. Relating the genetic positions of these loci to our restriction map requires a knowledge of the chromosomal orientation of the cloned DNA. We determined this by using a large inversion which has one breakpoint in the cloned region. This inversion, called *In(1)D6j*, was induced by hybrid dysgenesis in the strain π_2 and includes the region between 5E3-7 and 17C2-3 (3, 12). Blotting analysis indicates that the 17C2-3 breakpoint is between coordinates −1.6 and +0.2 (W. Engels and C. Preston, personal communication). This places the endpoint within or very close to the resident P element found at −0.8 in π_2 flies (see Fig. 8) (31). Separately, we

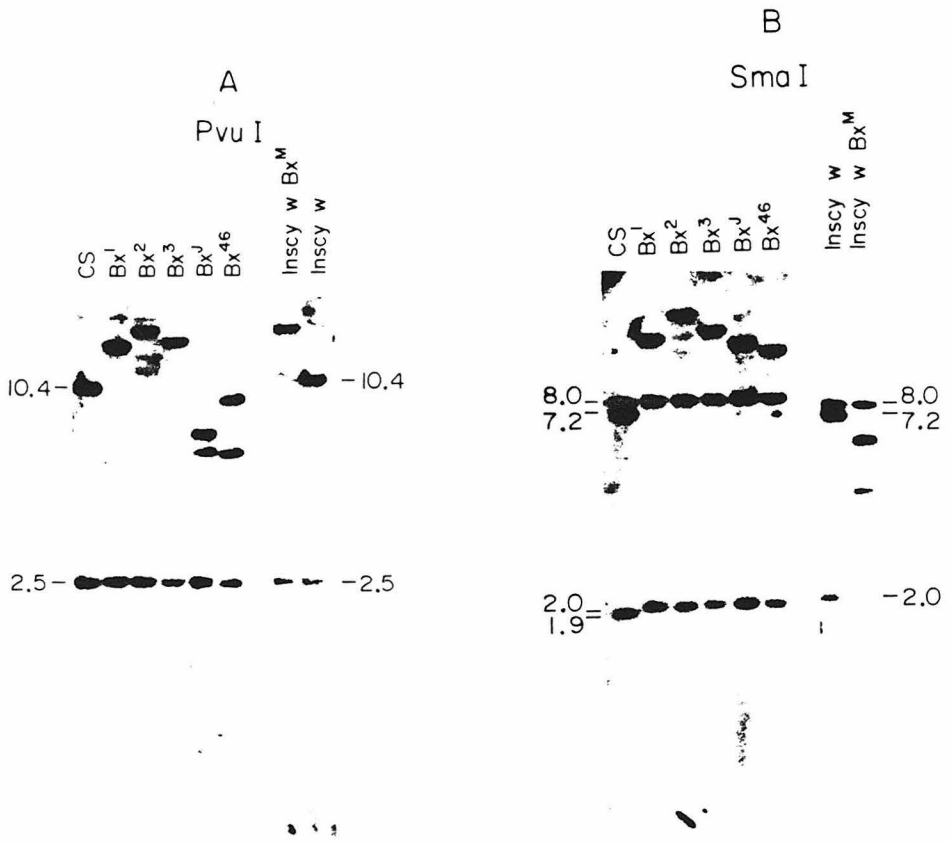


FIG. 5. Restriction-digested *Beadex* and wild-type DNAs. Each blot was probed with ^{32}P-labeled λCS17C-24 DNA. (A) *Pvu*I-digested genomic DNAs. (B) *Sma*I-digested genomic DNAs.

hybridized [^{3}H]cRNA made from λCS17C-2 and λCS17C-9 to polytene chromosomes from *In(1)D6j*. These clones contain inserts from opposite sides of the inversion breakpoint. Fig. 1c and d show hybridizations done with λCS17C-2 and λCS17C-9, respectively. λCS17C-2 hybridized only at the distal breakpoint, indicating that these sequences are within the inverted segment. λCS17C-9 hybridized only at the proximal breakpoint, indicating that the sequences cloned in this phage are outside the inversion. Taken together, these results indicate that λCS17C-2 is distal to λCS17C-9 (Fig. 4a).

Six mutant alleles of the *Beadex* locus associated with insertions in a 500-base-pair region. As reported above, the spontaneous mutation Bx^2 is associated with the insertion of two gypsy transposons at coordinates −0.5 and −1.5. To analyze the restriction maps of three other spontaneous *Bx* alleles (Bx^1, Bx^3, Bx^M), one heat-induced allele (Bx^J), and one allele of unknown origin (Bx^{46}), we probed whole genome blots of DNA from each of these strains with overlapping phage clones from the Canton-S library. The four clones that we selected, λCS17C-13, λCS17C-24, λCS17C-19, and λCS17C-9, span the entire 49-kb cloned region (Fig. 4a). DNA from Bx^2 flies, as well as Canton-S and *Inscy w* flies, were included in this analysis. We included *Inscy w* because the spontaneous allele Bx^M originated in this stock (we used the *Inscy w* Bx^M stock to isolate DNA for our analysis of the Bx^M mutation). Any differences between *Inscy w* and *Inscy w* Bx^M are almost certainly associated with the occurrence of this mutation.

Figure 5 shows representative blots which were hybridized with ^{32}P-labeled λCS17C-24 DNA. Note that this clone contains DNA from the region between coordinates −6.6 and +6.4, within which the gypsy elements are inserted in Bx^2. As expected from the composite restriction map in Fig. 4a, two *Pvu*I fragments (10.4 and 2.5 kb) and three *Sma*I fragments (8.0, 7.2, and 1.5 kb) in Canton-S DNA showed homology to λCS17C-24. *Inscy w* flies, which are also Bx^+, showed the same set of homologous fragments (the 2.0-kb *Sma*I fragment in *Inscy w* differs from the 1.95-kb fragment of Canton-S only by the 50- to 100-base-pair deletion discussed above).

When hybridized with λCS17C-24 DNA, each of the *Bx*

mutant DNA digests exhibited restriction patterns different from those observed with Canton-S DNA. As expected, the 10.4-kb *Pvu*I fragment (−7.4 to +3.0) and the 7.2-kb *Sma*I fragment (−7.2 to 0.0) were missing in the Bx^2 strain. In both cases the missing fragment was replaced by a single larger fragment which represents the insertion of two gypsy transposons (see restriction maps of Bx^2 in Fig. 4b and cBx^{2}17C-1 in Fig. 2). The other *Bx* mutants exhibited a similar pattern. In each case the 10.4-kb *Pvu*I and 7.2-kb *Sma*I fragments were replaced by fragments of novel size, indicating that a rearrangement had occurred between coordinates −7.2 and 0.0. As in Bx^2, these novel frgments can only be explained by the insertion of new DNA sequences in this segment. The restriction patterns we observed when mutant and wild-type DNAs were digested with *Eco*RI, *Sac*I, *Hin*dIII (data not shown), and *Bam*HI (see below) confirmed this conclusion and further defined the sites at which the new sequences had inserted.

Figure 6 shows a whole genome blot done with *Bam*HI-digested genomic DNAs and probed with end-labeled, *Bam*HI-cut p24R8.0B. This plasmid subclone contains the 8.0-kb *Eco*RI fragment (−6.6 to +1.4) from λCS17C-24 inserted at the *Eco*RI site of pUC8 (42). End-labeled, *Bam*HI-cut plasmid was used to give uniform labeling of different-sized genomic *Bam*HI fragments on the blot. Canton-S DNA contains five *Bam*HI fragments (9.0, 5.0, 1.8, 0.8, and 0.4 kb in length) which are homologous to this probe. Bx^J, *Inscy w*, and *Inscy w* Bx^M DNA have no *Bam*HI site at coordinate +5.8 due to a simple polymorphism, resulting in the appearance of a 6.4-kb *Bam*HI fragment instead of the 5.0-kb fragment of Canton-S. This restriction site polymorphism has no bearing on the *Bx* phenotype of these strains, since both Bx^+ and *Bx* mutant strains possess it. The 0.4-kb *Bam*HI fragment (−2.0 to −1.6) is missing in Bx^1, Bx^3, Bx^J, Bx^{46}, and *Inscy w* Bx^M. Since the 0.4-kb fragment is the only *Bam*HI fragment within the −7.2 to 0.0 segment which is altered in these *Bx* mutant strains, it must contain the sites at which the insertions occur in each of the alleles. The positions of each of the insertions are indicated on the composite restriction maps in Fig. 4b. Note that the distal gypsy element in Bx^2 is located at coordinate −1.5, very close to the 0.4-kb *Bam*HI fragment (−2.0 to −1.6). Thus, each of these six *Bx* mutations is associated with an insertion in the 500-base-pair interval between coordinates −2.0 and −1.5.

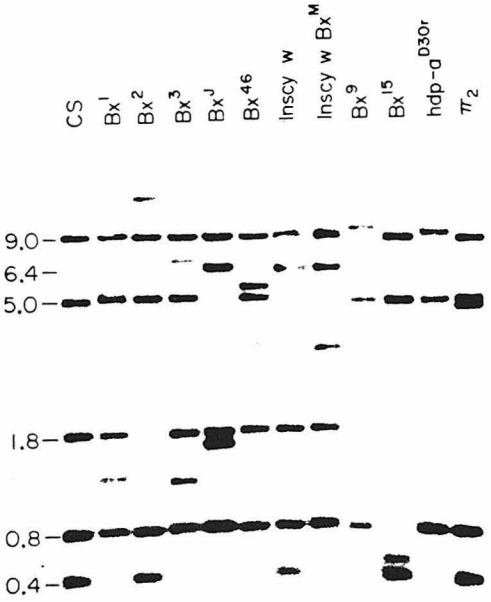


FIG. 6. *Bam*HI-cut *Beadex* mutant and wild-type DNAs were probed with *Bam*HI-digested p24R8.0B DNA which had been labeled at the *Bam*HI ends with [^{32}P]dGTP by using the Klenow fragment of DNA polymerase I. This clone contains sequences which lie between coordinates −6.6 and +1.4.

In 49 kb of DNA surrounding this 500-base-pair segment, we found few significant differences between the restriction maps of *Bx* mutant DNAs and Canton-S DNA (see Fig. 4b for composite maps.) In several instances *Bx* mutant strains were found to possess restriction sites not found in Canton-S and vice versa. However, with the exception noted below, both forms of each polymorphism were found among the three Bx^+ strains examined (Canton-S, *Inscy w*, and π_2). Therefore, none of these simple polymorphisms is related to the *Bx* phenotype of the flies that possess them.

A more complex loss of restriction sites was found in the Bx^2 and Bx^{46} strains (see Fig. 4b). In these mutants, the loss of three *Eco*RI sites, a single *Hin*dIII site, and a *Pvu*I site was due to the deletion of ca. 1.6 kb of DNA in the region between coordinates −17.8 and −15.8. Even though Bx^2 and Bx^{46} were isolated independently and possess different inserted elements (see below), these two deletions appear to be identical, as surmised by blot hybridization analysis (data not shown). The −17.8 to −15.8 interval is identical to Canton-S in each of the other *Bx* mutants examined. However, since we did not find this deletion in any of the Bx^+ strains that we examined, we cannot exclude the possibility that it plays a role in manifesting wing scalloping in these two strains.

The nature of DNA sequences inserted at the *Beadex* locus. We were interested in determining whether any of the inserted DNA sequences described above were related to one another. To do this we compared their sizes and restriction maps. Restriction enzymes which do not cut within the DNA insertions give rise to only one novel-sized fragment when genomic DNA cut with them is examined by blot hybridization. The difference in size between this insertion-bearing fragment and the wild-type fragment is equal to the length of the inserted element. Based on the blot hybridization experiments described above, we estimated the size of the inserted elements. These results are summarized in Fig. 7. The insertions ranged in size from 5.2 kb in the Bx^{46} strain to over 12 kb in the Bx^M allele.

Enzymes which have recognition sites within the inserted elements give rise to two novel-sized fragments homologous to cloned Canton-S DNA. The lengths of these fragments can be used to deduce the positions of restriction sites within the inserted elements. For instance, the Bx^{46} insertion gave rise to two novel *Pvu*I fragments which are 9.6 and 6.0 kb in length. Each of these fragments had one end within the inserted element and the other in flanking DNA. Using probes on either side of this insertion (λCS17C-13 on the

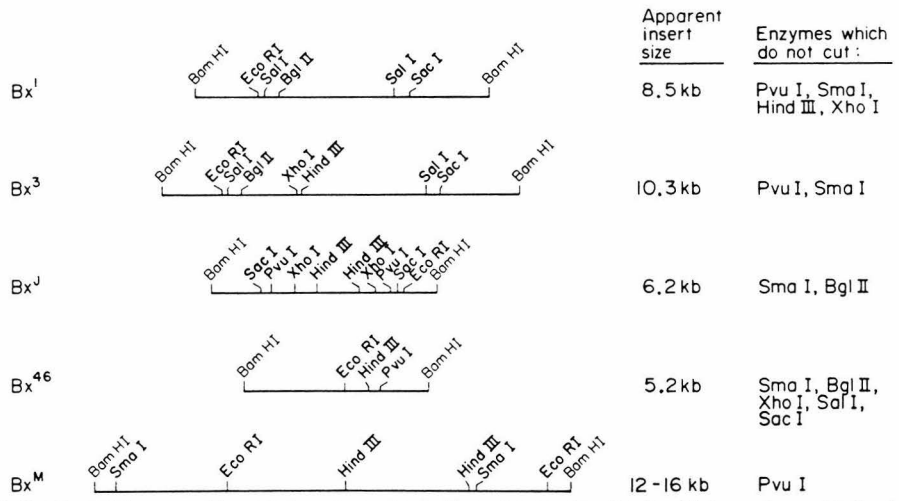


FIG. 7. Partial restriction maps of inserted elements associated with *Beadex* mutations. The *Bam*HI sites at the termini of each map are those at coordinates −1.6 and −2.0 in the Canton-S map and are not part of the inserted elements. Each map is oriented with the right end toward the centromere. Enzymes which were tested but did not cut within the insertions are shown on the right.

distal side and pS25.1 on the proximal side), we determined which novel fragment arose from each end of the insertion (data not shown). This analysis positioned the 9.6-kb fragment distally and the 6.0-kb fragment proximally and placed the single internal *Pvu*I site about 1.4 kb distal to the *Bam*HI site at coordinate −1.6 (see Fig. 7). Similar analyses with other restriction enzymes are summarized in Fig. 7. Since each inserted element possesses sequences that are not homologous to our Canton-S-derived cloned probes, we could not detect fragments which are internal to the inserted elements themselves. For this reason the restriction maps shown in Fig. 7 are incomplete. Nonetheless, it is apparent that the inserted elements are all different None of their restriction maps bears any resemblance to that of the gypsy transposons inserted in Bx^2 DNA. The Bx^1 and Bx^3 elements have five matching restriction sites and appear to differ only by an internal 1.8-kb deletion-insertion. This deletion-insertion includes the region containing the unique *Xho*I and *Hin*dIII sites found in the Bx^1 element since the Bx^1 element is not cut by these enzymes. The maps of both of these elements bear a resemblance to that of the retrovirus-like transposable element roo (28), also called B104 (36). The restriction map of the inserted element in Bx^{46} DNA is similar to published maps of the copia transposon and is approximately the same length (32). The insertion found in Bx^J has a restriction map like that of the 3S18 retrovirus-like transposon (1). We are not aware of any transposable element with a restriction map similar to that of the element inserted in *Inscy w* Bx^M DNA.

***Beadex* and *heldup-a* mutations induced by hybrid dysgenesis.** The P element inserted at 17C2-3 in the *D. melanogaster* strain π_2 has recently been cloned and sequenced (31). Its position corresponds to coordinate −0.8 on our restriction map. Since this strain has normal wings, this P element does not appear to disrupt either the *Bx* or the *hdp-a* locus. However, the π_2 X-chromosome gives rise to both *Bx* and *hdp-a* mutations at a frequency above 10^{-3} in the progeny of dysgenic hybrids (4, 11, 13, 38). This high rate is specific to the π_2 chromosome and seems to be associated with the 17C2-3 P element (13, 38). In situ hybridization of P element DNA with polytene chromosomes revealed that the 17C2-3 P element is deleted in a number of these *Bx* and *hdp-a* mutants (W. Engels and C. Preston, personal communication; 12).

To find out how such events disrupt *Bx* and *hdp-a* functional domains, we analyzed whole genome blots from two dysgenesis-induced *Bx* alleles (Bx^9 and Bx^{15}) and one *hdp-a* allele ($hdp\text{-}a^{D30v}$) which was induced in the same way. DNA from each of these mutants was digested with *Bam*HI, *Sac*I, *Eco*RI, *Hin*dIII, *Sma*I, and *Pvu*I in separate reactions and was compared to Canton-S and π_2 genomic DNA cut with the same enzymes. Figure 6 shows *Bam*HI digests of each of these DNAs, which were hybridized with a ^{32}P-end-labeled *Bam*HI digest of p24R8.0B DNA (see above). As a result of the 2.9-kb P element insertion, the 1.8-kb band (−1.6 to +0.2) of Canton-S was replaced by a 4.7-kb band in π_2. All three hybrid dysgenesis-induced mutants lacked this 4.7-kb band, as expected from the in situ hybridizations. Precise excision of the P element would result in reconstitution of a 1.8-kb fragment comigrating with the Canton-S fragment. No such fragment was found in any of the three mutant strains, indicating that each P element excision or deletion was imprecise. Since each mutant is missing *Bam*HI fragments which flank the 4.7-kb fragment of π_2 (for instance Bx^9 is missing both the 0.4- and 9.0-kb bands), the deletions include not only the P element but also flanking DNA. Similar blot hybridization analyses, using genomic DNA cut with the other five restriction enzymes mentioned above (data not shown), confirmed that each of the three mutants has lost most, if not all, of the P element (see below), as well as some flanking DNA.

Figure 8 shows the segments that were deleted. Bx^{15} is missing sequences on both sides of the P element insertion site. This 2.0-kb deletion certainly includes the sequences between coordinates −1.3 and +0.4. The Bx^9 deletion includes 1.2 kb of flanking DNA and overlaps the Bx^{15}

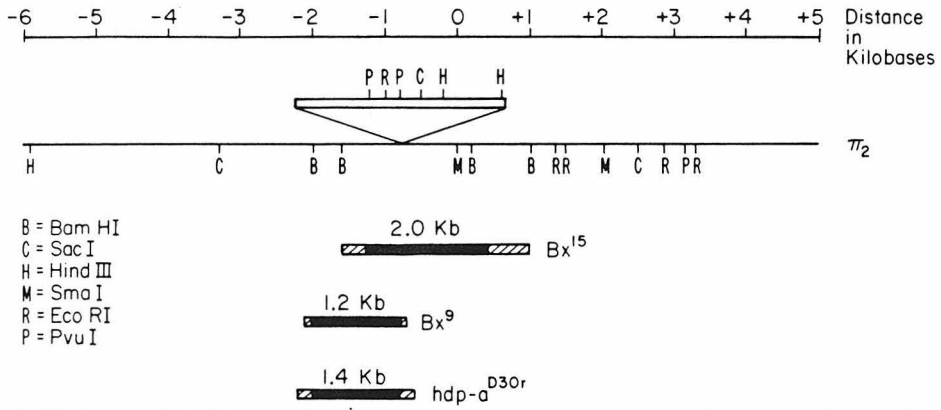


FIG. 8. Deletion maps of hybrid dysgenesis-induced mutants. The restriction map shown is that of DNA from π_2, the parental strain to each of the mutants shown. The coordinates are the same as those in Fig. 5. The P element is represented as an insertion at coordinate −0.8. Dark bars indicate the regions known to be deleted in each of the mutants. The hatched segments indicate regions that may or may not be deleted. The size of each deletion, exclusive of the 2.9 kb of P element DNA that was lost, is indicated above the bar in each case. All sites outside of the deleted segment are identical to π_2 in each of the mutant strains. No restriction sites from the P element itself remain in any of the mutants.

deletion. In this mutant the proximal (right) boundary of the missing segment is within 100 base pairs of the P element insertion site at −0.8 and may be coincident with it. The entire −1.5 to −2.0 segment, in which insertions occurred in six *Beadex* alleles described above, is deleted in Bx^{9}.

The *hdp-a*^{D30r} deletion has a right boundary indistinguishable from that of the Bx^{9} deletion, in this case with about 200 base pairs of uncertainty. Despite the fact that the *hdp-a*^{D30r} mutant has no wing scalloping, this deletion appears to include the entire segment removed in Bx^{9}. This is not surprising since genetic studies indicate that *hdp-a* mutations are potent suppressors of *Bx* in *cis* (24). Thus, the phenotype of *hdp-a*^{D30r} is consistent with that of a double mutation affecting both *hdp-a* and *Bx*. Note that the *hdp-a* function is not disrupted in Bx^{9} and Bx^{15} since these alleles have no heldup wing phenotype and loss of the *hdp-a*$^{+}$ function would suppress the wing-scalloping phenotype (both Bx^{9} and Bx^{15} mutants exhibit severe wing scalloping).

Blot hybridizations on Bx^{9}, Bx^{15}, and *hdp-a*^{D30r} with probes covering the remainder of the 49-kb cloned region failed to reveal any other differences between these strains and π_2.

DISCUSSION

The evidence presented above indicates that we isolated DNA sequences from both the *Beadex* and *heldup-a* loci of wild-type *Drosophila melanogaster*. We confirmed the identity of these sequences by blot hybridization analysis of one *heldup-a* allele and eight *Beadex* alleles, all of which behave genetically as simple mutations at the gene locus. When compared with Canton-S DNA, each of the *Beadex* alleles exhibited alterations (insertions or deletions) within a 700-base-pair segment which lies between coordinates −2.0 and −1.3. No alterations relative to Canton-S were observed in this region in either of two Bx^{+} strains examined. Thus, disruption of this short segment is highly correlated with mutations in the *Beadex* locus. Examination of both the banding patterns on polytene chromosomes and quantitative blot hybridizations failed to detect any increase in the copy number of the *Bx-hdp-a* region in the mutant strains (unpublished data). Therefore, we think it is unlikely that any increases in gene dosage, which might contribute to the *Beadex* phenotype, accompanied these disruptions.

In several cases (*Inscy w* Bx^{M}, Bx^{9}, and Bx^{15}), we were able to compare mutants with the parental strains from which they were isolated. Gross changes, coincident with the mutation event, should be detected as differences in the restriction maps of the mutant and parental strains. In each case we found that, over the entire 49-kb interval, the mutant strains differed from their parental strains only by the insertions and deletions mentioned above.

Although the parental strain for the mutant allele Bx^{2} is not available, it is known that this mutation behaves genetically as a gypsy transposon insertion in that it is suppressed by the *su(Hw)* mutation (23, 25, 29). Our molecular analysis detected gypsy transposon insertions in this strain at coordinates −0.5 and −1.5. The −1.5 insertion is within the segment deleted in the Bx^{9} mutation and thus is near or within the *Beadex* functional domain. On the basis of these facts we suggest that the gypsy insertion at −1.5 is responsible for the Bx^{2} mutation. Since gypsy elements are known to disrupt gene function at a distance without physically disrupting the structural locus itself, it is possible that the +0.5 element also plays a role in the Bx^{2} mutation (2).

No parental strains are available for comparison with the mutant alleles Bx^{1}, Bx^{3}, Bx^{4}, or Bx^{46}. Each of these strains differs from Canton-S by DNA insertions within the −2.0 to −1.6 segment. The close proximity of these insertions to one another and to the *Beadex* locus itself (as defined by deletions) strongly suggests that they are within sequences required for the regulatory function of the *Beadex* locus. However, it is conceivable that these insertions do not physically disrupt the *Beadex* locus but rather are outside the regulatory domain and act at a distance to disrupt the normal *Beadex* function. Also, we cannot exclude the possibility that the true lesions in these strains are too small to detect by our blotting analysis and that these insertions, although close to the *Beadex* locus, do not affect it.

Genetic analysis has indicated that *Beadex* mutations result in an increase in gene activity (24). Mutations in the nearby *heldup-a* gene suppress the *cis* allele of the *Beadex* locus, suggesting that *heldup-a* encodes the structural gene activated by *Bx* mutations. We showed that deletion of the segment of DNA between coordinates −2.1 and −0.7 (the Bx^9 strain) is sufficient to elicit a strong dominant *Beadex* phenotype. This suggests that the normal function of these sequences is to repress *heldup-a*$^+$ activity. Lifschytz and Green found that large chromosomal deficiencies which include both *Beadex* and *heldup-a* exhibit no dominant *Beadex* phenotype (24). The dominant wing scalloping found in Bx^9 and the normal wing position of these flies strongly suggest that the *hdp-a*$^+$ locus is not disrupted by this deletion. The Bx^{15} deletion, which removes 2.0 kb of DNA between +1.0 and −1.6, has a phenotype very similar to Bx^9, indicating that it also leaves *hdp-a*$^+$ intact.

In contrast to Bx^9 and Bx^{15} the *hdp-a*^{D30r} mutant exhibits a fully penetrant wing position abnormality and no wing scalloping. Our blot hybridization analysis of this mutant indicates that it is deficient for the entire region lost in the Bx^9 deletion. However, the *hdp-a*^{D30r} deletion is about 200 base pairs larger than that of Bx^9. We suggest that this deletion removes sequences from both the *Beadex* and *heldup-a* domains and that the loss of *heldup-a*$^+$ function precludes the expression of the *Beadex* phenotype. An alternative interpretation of this mutant would place the true *hdp-a* mutation outside the deleted segment. In this case the occurrence of the imprecise P element excision-deletion would be merely coincidental. Since deletion events like the one in *hdp-a*^{D30r} occur at a frequency of only 10^{-3}, we think it unlikely that such an event coincided with an independent *hdp-a* mutation.

Given the above caveat, the segment of the *hdp-a*$^+$ functional domain removed in *hdp-a*^{D30r} must be within DNA which is not deleted in Bx^9 or Bx^{15} (since both of these mutants are *hdp-a*$^-$). This segment includes DNA between coordinates −2.0 and −2.3. This suggests that sequences only a short molecular distance from the *Beadex* locus (less than 1,600 base pairs) are necessary for manifestation of the *Beadex* mutant phenotype. This position agrees well with the recombinational mapping (24) which placed an *hdp-a* mutation 0.0045 map units distal to *Beadex*.

The close molecular proximity of these two loci supports the model proposed by Lifschytz and Green (24), wherein the wild-type *Beadex* locus functions as a *cis*-acting negative regulatory element for the *heldup-a* gene. This model predicts that deletion of *Beadex* sequences would increase the amount of *heldup-a*$^+$ activity which would in turn result in a wing-scalloping phenotype. Deletions including *heldup-a*, such as *hdp-a*^{D30r}, would obviate the possibility of hyperactivity and thus suppress *Beadex* mutations in *cis*.

Although it is clear from this analysis that some portion of the *Beadex* locus lies between coordinates −2.1 and −0.7, this does not place any limits on the extent of the *Beadex* functional domain. The inversion *In(1)D6j*, which was derived from the π_2 strain, is informative in this regard. The proximal breakpoint of this inversion is probably located within the P element inserted at coordinate −0.8 and undoubtedly lies between coordinates −1.6 and +0.2 (W. Engels and C. Preston, personal communication). This breakpoint disrupts the *heldup-b* gene but not *heldup-a* or *Beadex* (the *heldup-b* gene is closely linked to *heldup-a*, and these mutants have a similar phenotype; however, *heldup-b* mutations are complemented by *heldup-a* and fail to suppress *Beadex*) (13; unpublished observations). This means that the *Beadex* and *heldup-a* functional domains can extend no farther rightward than +0.2. If we assume that *Beadex* sequences are found only to the right of *heldup-a*, the *Beadex* locus is entirely confined to the region between coordinates −2.3 and +0.2. Of course, we have not proven that portions of the *Beadex* functional domain do not lie both to the right and left of *heldup-a*.

Our analysis of *Beadex* and *heldup-a* alleles induced by hybrid dysgenesis revealed that these mutations result from imprecise excision or deletion of a nearby P element. Previous studies on the excision of P elements detected precise excision events (31, 35, 37) and imprecise events which left behind a portion of the P element and did not remove any flanking sequences (12, 37). We showed that the events associated with the Bx^9, Bx^{15}, and *hdp-a*^{D30r} mutations removed sequences which were adjacent to the P element in the parental strain. In one case (Bx^{15}) sequences from both sides of the P element were clearly deleted. We believe that each of these mutants lost the entire P element at −0.8. No restriction sites from the P element were detected in the mutant DNAs (including a *Hin*dIII site only 39 base pairs from the terminus), and in situ hybridizations of labeled P element DNA to these mutant chromosomes failed to detect any homology at 17C (W. Engels and C. Preston, personal communication). Such imprecise excision events provide a novel means for isolating both *Beadex* and *heldup-a* mutations at a high frequency (11, 13, 38). Molecular analysis of mRNA encoded by the *Bx-hdp-a* region of these mutants should be useful in determining the structural relationship between the *Beadex* and *heldup-a* loci. In particular, we hope to determine whether *Beadex* mutations affect sequences within or outside transcribed regions and whether the quantity or size of transcripts are affected. Resolution of these questions will lead to greater insight into the molecular role of the *Beadex* locus and help us to distinguish between transcriptional, post-transcriptional, and translational mechanisms for its effect on the level of *hdp-a*$^+$ activity.

ACKNOWLEDGMENTS

We thank I. S. Margitich for expert technical assistance with electron microscopy. We are indebted to W. Bender, K. O'Hare, and G. Rubin for generously providing cloned DNA. W. R. Engels has been exceptionally helpful; we thank him for communicating results before publication and for giving us mutant fly stocks. We thank B. J. Bond, M. D. Garfinkel, E. Meyerowitz, S. Sharp, and B. Wold for their helpful comments on this manuscript. We thank all of our colleagues at Caltech for their comments and insights.

This work was supported by a National Research Service Award (GM-07616) and a Public Health Service grant (GM-10991) from the National Institutes of Health.

LITERATURE CITED

1. **Bell, J. R., T. Schmidt, A. M. Bogardus, and M. Pelligrini.** 1984. A new copia-like transposable element found in an rDNA gene unit. J. Cell. Biochem. **8**:164.
2. **Bender, W., M. Akam, F. Karch, P. A. Beachy, M. Peifer, P. Spierer, E. B. Lewis, and D. S. Hogness.** 1983. Molecular genetics of the bithorax complex in *Drosophila melanogaster*. Science **221**:23–29.
3. **Benton, W. D., and R. W. Davis.** 1977. Screening λgt recombinant clones by hybridization of single plaques in situ. Science **196**:180–182.
4. **Berg, R., W. R. Engels, and R. A. Kreber.** 1980. Site-specific X-chromosome rearrangements from hybrid dysgenesis in *Drosophila melanogaster*. Science **210**:427–429.
5. **Bingham, P. M., R. Levis, and G. M. Rubin.** 1981. Cloning of DNA sequences from the *white* locus of *D. melanogaster* by a novel and general method. Cell **25**:693–704.
6. **Blattner, F. R., B. G. Williams, A. E. Blechl, K. Denniston-**

Thompson, H. E. Faber, L. A. Furlong, D. J. Grunwald, D. O. Kieffer, D. D. Moore, E. L. Sheldon, and O. Smithies. 1977. Charon phages: safer derivatives of bacteriophage lambda for DNA cloning. Science **196**:161–169.

7. **Cline, T. W.** 1978. Two closely linked mutations in *Drosophila melanogaster* that are lethal to opposite sexes and interact with daughterless. Genetics **90**:683–698.
8. **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.
9. **Davis, R. W., M. Simon, and N. Davidson.** 1971. Electron microscope heteroduplex for mapping regions of base sequence homology. Methods Enzymol. **21**:413–427.
10. **Deak, I. I.** 1977. Mutations of *Drosophila melanogaster* that affect muscles. J. Embryol. Exp. Morphol. **40**:35–63.
11. **Engels, W. R.** 1979. Extrachromosomal control of mutability in *Drosophila melanogaster*. Proc. Natl. Acad. Sci. U.S.A. **76**:4011–4015.
12. **Engels, W. R.** 1983. The P family of transposable elements in *Drosophila*. Annu. Rev. Genet. **17**:315–344.
13. **Engels, W. R., and C. R. Preston.** 1981. Identifying P factors in *Drosophila* by means of chromosome breakage hotspots. Cell **26**:421–426.
14. **Gaissinovitch, A., and S. Gerschenson.** 1928. Ein neues Allelmorph von Beadex bei *Drosophila melanogaster*. Biol. Zentralbl. **48**:385–387.
15. **Gall, J. G., and M. L. Pardue.** 1971. Nucleic acid hybridization in cytological preparations. Methods Enzymol. **21**:470–480.
16. **Green, M. M.** 1953. The Beadex locus in *Drosophila melanogaster*: genetic analysis of the mutant Bx^{reb}. Z. Indukt. Abstammungs Vererbungsl. **85**:435–449.
17. **Green, M. M.** 1953. The Beadex locus in *Drosophila melanogaster*: on the nature of the mutants Bx^{1} and Bx^{3}. Genetics **38**:91–105.
18. **Greenwald, I. S., P. W. Sternberg, and H. R. Horvitz.** 1983. The lin-12 locus specifies cell fates in *Caenorhabditis elegans*. Cell **34**:435–444.
19. **Grosveld, F. G., H. M. Dahl, E. deBoer, and R. Flavell.** 1981. Isolation of β-globin-related genes from a human cosmid library. Gene **13**:227–237.
20. **Hanahan, D., and M. Meselson.** 1980. Plasmid screening at high colony density. Gene **10**:63–67.
21. **Ish-Horowicz, D., and J. F. Burke.** 1981. Rapid and efficient cosmid cloning. Nucleic Acids Res. **9**:2989–2998.
22. **Jollos, V.** 1933. Die Übereinstimmung der bei *Drosophila melanogaster* nach Hitzeeinwirkung entstehenden Modifikation und Mutationen. Naturwissenschaften **21**:831–834.
23. **Lewis, E. B.** 1981. Developmental genetics of the bithorax complex in *Drosophila*, p. 189–208. *In* D. D. Brown and C. R. Fox (ed.), Developmental biology using purified genes. Academic Press, Inc., New York.
24. **Lifschytz, E., and M. M. Green.** 1979. Genetic identification of dominant overproducing mutations: the *Beadex* gene. Mol. Gen. Genet. **171**:153–159.
25. **Lindsley, D. L., and E. H. Grell.** 1968. Genetic variations of *Drosophila melanogaster*. Publ. no. 627. Carnegie Institute, Washington, D.C.
26. **Maniatis, T., R. C. Hardison, E. Lacy, J. Lauer, C. O'Connell, D. Quon, G. K. Sim, and A. Efstratiadis.** 1978. The isolation of structural genes from libraries of eukaryotic DNA. Cell **15**:687–701.
27. **Maxam, A. M., and W. Gilbert.** 1980. Sequencing end-labeled DNA with base-specific chemical cleavages. Methods Enzymol. **65**:449–560.
28. **Meyerowitz, E. M., and D. S. Hogness.** 1982. Molecular organization of a *Drosophila* puff site that responds to ecdysone. Cell **28**:165–176.
29. **Modolell, J., W. Bender, and M. Meselson.** 1983. *Drosophila melanogaster* mutations suppressible by the suppressor of *Hairy-wing* are insertions of a 7.3 kilobase mobile element. Proc. Natl. Acad. Sci. U.S.A. **80**:1678–1682.
30. **Mullins, J. I., J. W. Casey, M. O. Nicholson, K. B. Burck, and N. Davidson.** 1981. Sequence arrangement and biological activity of cloned feline leukemia virus proviruses from a virus-productive human cell line. J. Virol. **38**:688–703.
31. **O'Hare, K., and G. M. Rubin.** 1983. Structures of P transposable elements and their sites of insertion and excision in the *Drosophila melanogaster* genomes. Cell **34**:25–35.
32. **Potter, S. S., W. J. Brorein, Jr., P. Dunsmuir, and G. M. Rubin.** 1979. Transposition of elements of the 412, copia and 297 dispersed repeated gene families in *Drosophila*. Cell **17**:415–427.
33. **Rigby, P. W. J., M. Dieckmann, C. Rhodes, and P. Berg.** 1977. Labeling deoxyribonucleic acid to high specific activity in vitro by nick translation with DNA polymerase I. J. Mol. Biol. **113**:237–251.
34. **Rozek, C. R., and N. Davidson.** 1983. *Drosophila* has one myosin heavy-chain gene with three developmentally regulated transcripts. Cell **32**:23–34.
35. **Rubin, G. M., M. G. Kidwell, and P. M. Bingham.** 1982. The molecular basis of P-M hybrid dysgenesis: the nature of induced mutations. Cell **29**:987–994.
36. **Scherer, G., C. Tschudi, J. Perera, H. Delius, and V. Pirrota.** 1982. B104, a new dispersed repeated gene family in *Drosophila melanogaster* and its analogies with retroviruses. J. Mol. Biol. **157**:435–451.
37. **Searles, L. L., R. S. Jokerst, P. M. Bingham, R. A. Voelker, and A. L. Greenleaf.** 1982. Molecular cloning of sequences from a *Drosophila* RNA polymerase II locus by P element transposon tagging. Cell **31**:585–592.
38. **Simmons, M. J., J. D. Raymond, N. A. Johnson, and T. M. Fahey.** 1984. A comparison of mutation rates for specific loci and chromosome regions in dysgenic hybrid males of *Drosophila melanogaster*. Genetics **106**:85–94.
39. **Smith, R. F., and R. J. Konopka.** 1982. Effects of dosage alterations at the *per* locus on the period of the circadian clock of *Drosophila*. Mol. Gen. Genet. **185**:30–36.
40. **Southern, E. M.** 1975. Detection of specific sequences among DNA fragments separated by gel electrophoresis. J. Mol. Biol. **98**:503–518.
41. **Steinmetz, M., A. Winoto, K. Minard, and L. Hood.** 1982. Clusters of genes encoding mouse transplantation antigens. Cell **28**:489–498.
42. **Vierira, J., and J. Messing.** 1982. The PUC plasmids, an M13mp7-derived system for insertion mutagenesis and sequencing with synthetic universal primers. Gene **19**:259–268.
43. **Wensink, P. C., D. J. Finnegan, J. E. Donelson, and D. S. Hogness.** 1974. A system for mapping DNA sequences in the chromosomes of *Drosophila melanogaster*. Cell **3**:315–325.
44. **Yamamoto, K. R., B. M. Alberts, R. Benzinger, L. Lawhorne, and G. Treiber.** 1970. Rapid bacteriophage sedimentation in the presence of polyethylene glycol and its application to large-scale virus purification. Virology **40**:734–744.

CHAPTER 2

Beadex Mutations Have Qualitative and Quantitative Effects Upon Transcripts Which Span part of the *heldup-a* Locus

WILLIAM W. MATTOX AND NORMAN DAVIDSON

SUMMARY

Mutations at the *Beadex* locus behave genetically as though they cause an excess in the activity of the closely linked *heldup-a* gene. These mutations are thought to remove or disrupt a *cis*-acting negative regulatory element which is involved in the repression of *heldup-a*$^{+}$ activity. Here we identify a number of transcripts which are homologous to a 34 kilobase region which includes DNA from the *Beadex* and *heldup-a* loci. Of these, we find only two that span part of the *heldup-a* locus as it is defined by mutations. These two transcripts, which are 4.0 kilobases and 2.0 kilobases in length, are oriented so that their 3' ends are nearest to the the *heldup-a* locus. In addition the 4.0 kilobase transcript spans the sites at which several *Beadex* mutant alleles have transposable element insertions. Developmental analysis of these transcripts indicates that only the 4.0 kilobase transcript is expressed at stages when the *Beadex-heldup-a* product is expected to be active. We find that this transcript is altered in structure in one *heldup-a* mutation and each of five *Beadex* mutations that we examined. Although the abnormally structured transcript from one *Beadex* mutant strain was expressed at slightly elevated levels as compared to the wild-type 4.0 kilobase transcript, the amount of RNA expressed in four other *Beadex* mutants was not significantly elevated. The qualitative and quantitative effects of *Beadex* and *heldup-a* mutations on the 4.0 kilobase transcript and its possible identity as a product of the *heldup-a* gene are discussed.

INTRODUCTION

Recent molecular studies on several eucaryotic genes have identified *cis*-acting sequence elements which are involved in the repression of gene expression (Johnson and Herskowitz, 1986; Miller et al., 1985; Brand et al., 1985; Laski et al., 1986; Oshima, 1982). These elements act at a variety of levels including RNA transcription, RNA splicing and protein function.

There are several reasons to believe that the *Beadex* (*Bx*) locus of *Drosophila melanogaster* is also such a *cis*-acting negative element. Mutations at this locus behave genetically as though they result from gene hyperactivity (Lifshytz and Green, 1979). Indeed, the wing scalloping phenotype, which is associated with *Bx* mutations, can be mimicked by simply increasing the dosage of a chromosomal region that contains the wild-type *Bx* locus (Green, 1953a; 1953b). Furthermore, *cis*-acting suppressors of dominant *Bx* alleles have been isolated. All of these suppressor mutations are coincident with mutations at the closely linked *heldup-a* (*hdp-a*) gene (Lifshytz and Green, 1979). Therefore, it appears that an intact $hdp\text{-}a^{+}$ gene is required in *cis* for the expression of the *Bx* mutant phenotype. This fact suggests that there is a *cis*-regulatory interaction between *Bx* and *hdp-a*. Lifshytz and Green (1979) have proposed that *Bx* and *hdp-a* mutations affect a single bipartite genetic unit in which the *hdp-a* gene is a structural element that is under the control of the *Bx* locus. They suggest that *Bx* mutations cause an increase in the activity of the $hdp\text{-}a^{+}$ gene and that this increase in $hdp\text{-}a^{+}$ activity is the cause of the wing scalloping phenotype associated with *Bx* mutations. Such a close relationship between these two loci is supported by their observation that one *Bx* mutant allele (Bx^{3}) and an *hdp-a* mutation ($hdp\text{-}a^{RBE1}$) map within a very short genetic distance, 0.0045 map units, of one another (Lifshytz and Green, 1979).

The analysis of DNA from the *Bx* and *hdp-a* loci has confirmed that they are closely associated with one another (Chapter 1). DNA structural alterations (insertions and deletions) that are associated with *Bx* mutations are found clustered in a small part of the cloned region. An 800 bp segment which must contain at least part of the *Bx* locus is defined by several *Bx* deletion mutations and a chromosome inversion breakpoint which does not affect the *Bx* locus (Chapter 1 and Appendix 1). The analysis of one *hdp-a* deletion mutation indicates that part of the *hdp-a* locus lies within a 200 bp segment which is itself located only 500 bp away from the 800 bp segment in which part of the *Bx* locus is located.

The existence of excess-of-function *Bx* mutations that result from deletions indicates that the normal function of the *Bx* locus is to repress gene activity. The *cis*-interaction of the *hdp-a* gene with *Bx* mutations makes it the most likely candidate for the structural gene which is repressed by the wild-type *Bx* locus. In order to understand the mechanisms by which the *Bx* locus might repress the *hdp-a* gene it is necessary to identify gene products that are encoded at the *hdp-a* locus and to determine the effects of *Bx* mutations on the quantity and/or structure of these products. Here we identify RNA transcripts that span part of the *hdp-a* locus and examine the effects of *Bx* and *hdp-a* mutations on them.

MATERIALS AND METHODS

Construction of isogenic strains. To minimize any quantitative differences in RNA expression, which might result from differences in genetic background, the mutants Bx^1, Bx^2, Bx^{46}, and Bx^J were made isogenic to Canton-S. This was done by first crossing *Bx* males with virgin Canton-S females. The virgin heterozygous F1 females were then backcrossed to Canton-S males. This was repeated for six generations. Finally, *Bx* male progeny were crossed with heterozygous virgins to

obtain homozygous females for construction of the isogenic homozygous stocks. The strains *Bx*^{15}, *hdp-a*^{D30r}, and *In(1)D6,y cho hdp-b* did not require such crosses since they were always compared directly to their parental strain π_2.

Isolation of Drosophila Poly A$^+$ RNA. Poly(A)$^+$ RNA was isolated by using guanidinium thiocyanate as described by Chirgwin et al. (1979) with the modifications employed by Bond and Davidson (1986).

RNA blot hybridization. RNA was fractionated using formaldehyde-agarose gels as described by Rozek and Davidson (1983). It was then transferred directly to either Hybond-N (Amersham) or Biodyne-A (Pall) nylon membranes by capillary blotting with 20X SSC. RNA was then irreversibly bound by wrapping the dried blot in plastic wrap (Springfield) and placing it directly upon a UV transilluminator for one to two minutes.

RNA blots were prehybridized in 50% formamide, 0.5 M Nacl, 0.1 M PIPES (pH 7.0), 0.2% polyvinylpyrolidine, 0.2% bovine serum albumin, 0.2% ficoll, 0.2% SDS and 250 mg/liter yeast RNA (type III, sigma) for a period of 3 hours. When ^{32}P-RNA probes were used 10^6cpm/ml were added and the hybridization was performed for 16 hours at 68^{o}C and then for 2 hours more at 75^{o}C. The blots were then washed with 0.2X SSC and 0.01% SDS at 68^{o}C. When nick translated ^{32}P-DNA probes were used, the prehybridization and hybridization were done in the above solution with 100 mg/l of calf thymus DNA and 5% dextran sulfate at 50^{o}C. Blots were then washed in 0.2X SSC, 0.01% SDS at 55^{o}C.

Synthesis of ^{32}P-RNA probes. Restriction fragments from the *Bx-hdp-a* region were subcloned into either pSP64 (Promega Biotech.), pSP65 (Promega Biotech.), pSP62PL or pIBI76 (International Biotechnologies). To prepare these plasmids for *in vitro* synthesis of runoff transcripts these plasmid subclones were linearized using restriction enzymes that do not leave a 3' overhang. High specific activity transcripts were made by using either T7 or SP6 RNA polymerase under

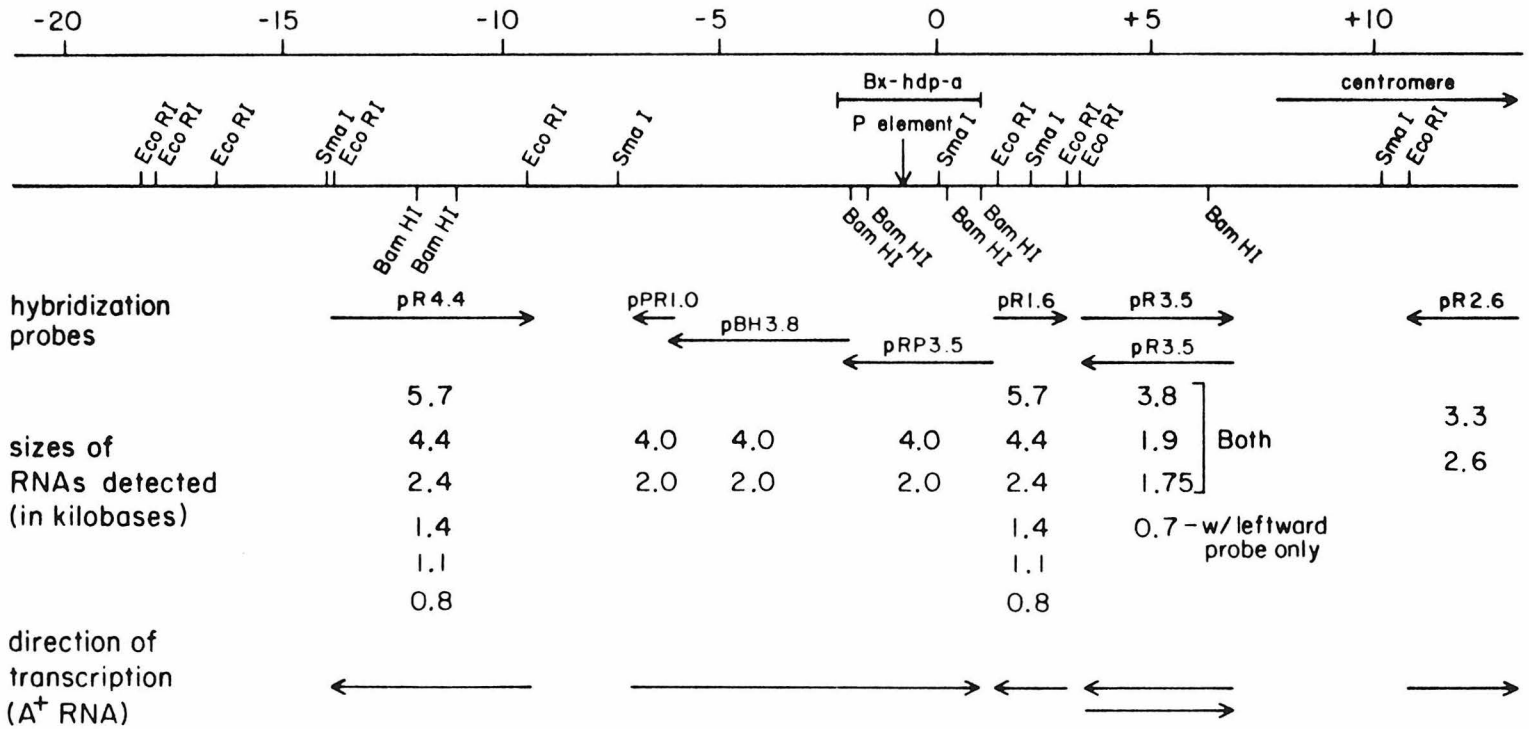
the conditions described by Zinn et al. (1983) except that only 200 nanograms of template were used in a 50 microliter reaction. Runoff products were checked by formaldehyde gel electrophoresis and autoradiography to assure that some fraction of the transcripts were of full length. We found that by limiting the enzyme and template concentrations the fraction of full length molecules could be increased at the expense of decreased levels of isotope incorporation. Trace labeled runoff transcripts for positive and negative controls were made as above except that 0.5 mM ^{32}P-CTP at a specific activity of 20 Ci/mmole was used.

RESULTS

Molecular structure of *Beadex* and *heldup-a* mutations. At the top of Figure 1 we have indicated the segment, located between coordinates -2.3 and +1.0, in which all structural alterations associated with *Bx* and *hdp-a* mutations are located (a more detailed map of this region can be found in Figure 5). In all, 13 different *Bx* mutations which disturb this region have been mapped. By mapping several small *Bx* deletion mutations we have established that at least the segment between -1.6 and -0.8 is required for the normal repressing function of the *Bx* locus. An inversion breakpoint at -0.8, in the *In(1)D6, y cho hdp-b* strain, does not disrupt the normal *Bx* function; therefore, it is likely that all sequences relevant to this function lie to the left of -0.8. Also, we note that several *Bx* mutations are associated with transposon insertions in the -1.6 to -2.0 segment. Analysis of one *hdp-a* deletion mutation reveals that it is missing sequences from between -0.8 and -2.3. Because *Bx* deletion mutations (which are *hdp-a*$^{+}$) are known to be missing the region between -0.8 and -2.1, sequences necessary for the *hdp-a*$^{+}$ function must be located between -2.3 and -2.1.

Our goal in the analysis presented here is to find transcripts which are affected, in either amount or structure, by the above mutations. To do this we examined a 34 kb region that surrounds the -2.3 to +1.0 segment.

Figure 1. Transcript map of the *Bx-hdp-a* region. The restriction map of the Canton-S strain is shown at the top of the figure with coordinates given in kilobases as in Chapter 1. The bar above the restriction map indicates the region between -2.3 and +1.0 in which structural alterations occur in all mapped *Bx* and *hdp-a* mutations. The top set of arrows represents only those probes which detected at least one Drosophila RNA species. The bottom set of arrows indicates the direction in which these RNAs are transcribed.


-20
-15
-10
-5
0
+5
+10
Bx-hdp-a
P element
centromere
Eco RI
Sma I
Bam HI
hybridization probes
pR4.4
pPR1.0
pBH3.8
pRP3.5
pR1.6
pR3.5
pR3.5
pR2.6
sizes of RNAs detected (in kilobases)
5.7
4.4
2.4
1.4
1.1
0.8
4.0
2.0
4.0
2.0
4.0
2.0
5.7
4.4
2.4
1.4
1.1
0.8
3.8
1.9
1.75
Both
0.7 - w/ leftward probe only
3.3
2.6
direction of transcription (A+ RNA)

Identification of transcripts in the *Bx-hdp-a* region. To search for transcripts that are encoded near the *Bx* and *hdp-a* loci, we have used ^{32}P-RNA probes synthesized from the SP6 and T7 phage promoters. Twelve different restriction fragments, which encompass the 34 kilobases (kb) region between -21.2 and +13.4 (see Table 1), were inserted into plasmid vectors that contain these phage promoters (see Materials and Methods). Note that although most of the 34 kb region is included within these twelve fragments, there are several small segments which were not contained in any of them. To make probes for blot hybridization experiments we synthesized high specific activity (approx. $6X10^8$ dpm/microgram) ^{32}P-RNA runoff transcripts from both strands of each of the twelve fragments. Each of these 24 probes was then hybridized with a separate filter on which gel fractionated poly(A)$^+$ RNA from 0-4 hr embryos, 8-20 hr embryos and 60 to 72 hr larvae (late second instar) was bound. In this way we were able to determine the locations and transcriptional orientation of a number of RNAs that are homologous to sequences within the 34 kb region. Subsequent higher resolution mapping of *Drosophila* transcripts was achieved by transcribing subsegments of the probes in Table 1.

For use as positive and negative controls in these hybridization experiments we synthesized trace labeled runoff transcripts in both orientations from each of the twelve fragments. Gel lanes containing either 5 or 50 picograms (pg) of each of these runoffs were then blotted and hybridized in the same container with the blotted *Drosophila* poly(A)$^+$ RNAs. Because the specific activity of these RNAs is very low (approx. 3×10^4 dpm/microgram), autoradiographs of the unhybridized controls did not show any visible bands after a 1 week exposure. By including one of these "control blots," on which both the probe strand and the probe-complementary strand were bound, we were able to: 1.) determine that our hybridizations had proceeded normally; 2.) determine whether or not our probes

TABLE 1

BLOT HYBRIDIZATION SURVEY OF THE *BX-HDP-A* REGION

PLASMID	COORDINATES	LEFTWARD PROBE		RIGHTWARD PROBE	
		transcripts detected (kb)	sensitivity limit (picograms)	transcripts detected (kb)	sensitivity limit (picograms)
pIR3.0	-21.2 to -18.2	none	<0.5	none	<0.5
pIR1.4	-17.9 to -16.5	none	<5. 0	none	<5.0
pIR2.7	-16.5 to -13.8	none	<0.5	none	<5.0
pIR 4.4	-13.8 to -9.4	none	<5.0	5.7, 4.4, 2.4 1.4, 1.1, 0.8	<0.5
pIMR2.1	-9.4 to -7.3	none	<5.0	none	<5.0
pSPR1.0	-6.9 to -5.9	4.0, 2.0	<0.5	none	<0.5
pSBH3.8	-6.1 to -2.0	4.0, 2.0	<0.5	none	<0.5
pSRP3.5	-2.1 to +1.4	4.0, 2.0	<0.5	none	<5.0
pIR1.6	+1.6 to +3.0	none	<0.5	5.7, 4.4, 2.4, 1.4, 1.1, 0.8	<0.5
pIR3.5	+3.3 to +6.8	3.8, 1.9, 1.75, 0.7	<1.0	3.8, 1.9, 1.75	<1.0
pIHR3.6	+7.2 to +10.8	none	<10.0	none	<5.0
pIR2.6	+10.8 to +13.4	3.3, 2.6	<0.5	none	<0.5

were strand specific; 3.) measure the approximate level of sensitivity of each hybridization; and 4.) estimate the concentration of each of the transcripts that we detected in whole fly RNA preparations.

Autoradiographs of the blotted *Drosophila* poly(A)$^+$ RNAs are shown in Figure 2. Only those blots on which hybridizing *Drosophila* RNA transcripts were detected are shown. A corresponding set of control blots is shown in Figure 3. The results obtained from both the control and experimental blot hybridizations are summarized in Table 1. In most cases 5 pg of the complementary runoff transcript were readily visible after a 24 hr exposure (see fig 3B for an example). In all but one case the 5 pg band was visible after longer exposures. We estimate that our limit of detection is less than 0.5 pg of RNA with many of the probes that we used (see Table 1). In all cases the hybridization observed on the control blots was strand specific, although on long exposures it is sometimes possible to see a band in the lane that contains 50 pg of runoff RNA from the same strand as the probe (see Figure 3A,H).

Several of the transcripts that we detected in *Drosophila* RNA, specifically, those detected by pIR3.5L (+6.8 to +3.3), pIR3.5R (+3.3 to +6.8) and pIR2.6L (+13.4 to +10.8), are located exclusively to the right of the inversion breakpoint, mentioned above, that is located near coordinate position -0.8 and defines the right limit of the *Bx* and *hdp-a* genes (Figure 1 and Figure 2F,G,H). It therefore seems unlikely that these transcripts are the products of the *hdp-a* gene.

Two transcripts hybridize with a probe (pSRP3.5L) that covers the short segment in which the structural alterations associated with *Bx* and *hdp-a* mutations are found (coordinates -2.3 to -0.8; see Figure 1). These RNAs are transcribed from left to right on the map in Figure 1 and are 2.0 kb and 4.0 kb in length (Figure 2B,C,D). In 0-4 hr embryos only the 2.0 kb species is expressed whereas 8-20 hr embryos and 60-72 hr larvae express only the 4 kb species. Both

Figure 2. Blot hybridization survey of the 34 kb region. Only those blots on which *Drosophila* RNA transcripts were detected are shown. Ten micrograms of poly(A)$^+$ RNA from 0-4 hr, 8-20 hr and 60-72 hr Canton-S flies were electrophoresed, blotted and hybridized with the indicated probes as described in Materials and Methods. The two marks beside each blot indicate the positions at which mouse 28S (5.09 kilobases) and 18S (1.76 kilobases) ribosomal RNAs migrated. The probes that were used are indicated below the blot both by plasmid name and coordinates covered. Probe names ending with an "L" denote leftward probes and those ending with an "R" denote rightward probes. The 60-72 hr RNAs shown on blots B and C are from a longer exposure than the 0-4 hr and 8-20 hr samples.

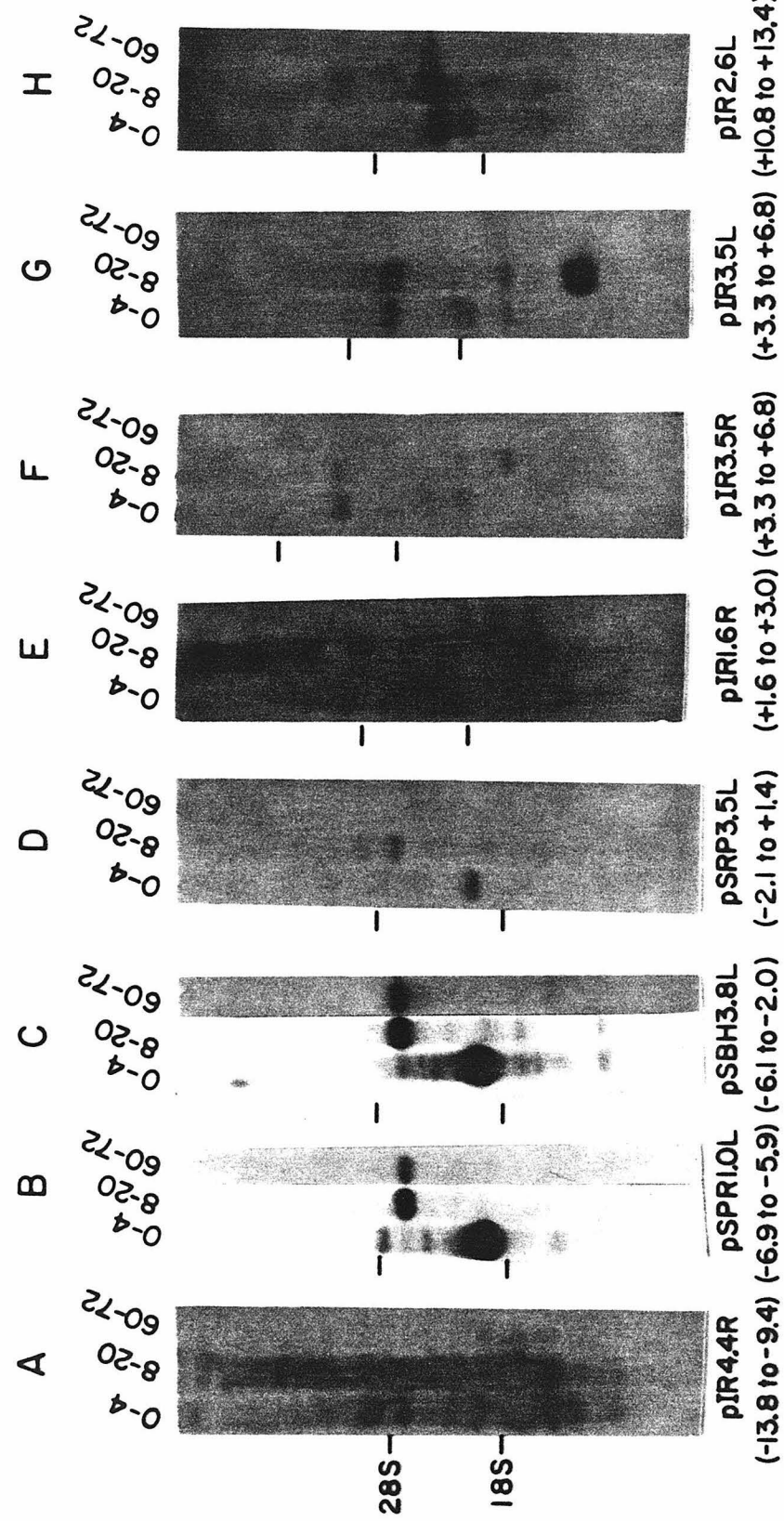
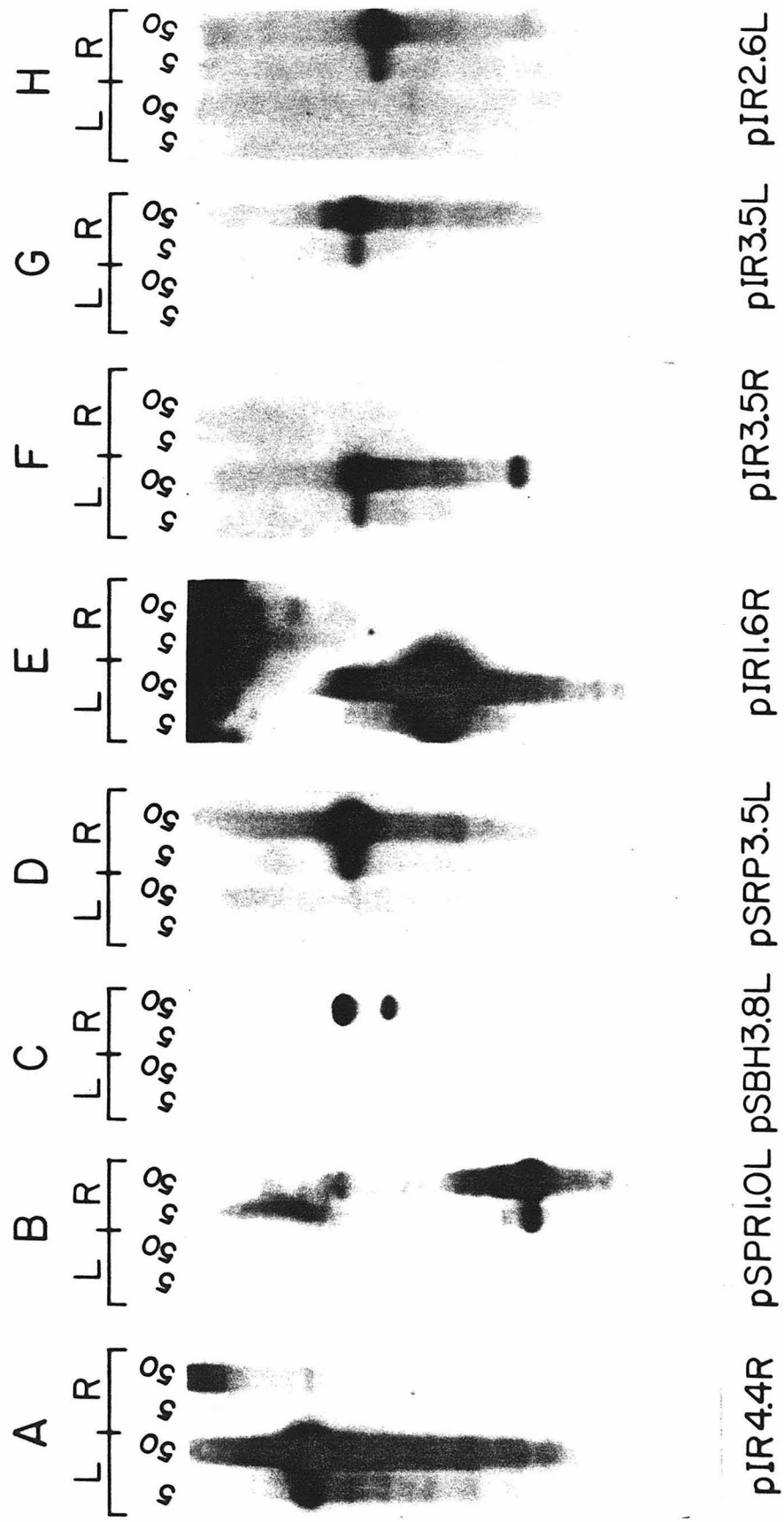

A
B
C
D
E
F
G
H
0-4
8-20
60-72
28S
18S
pIR4.4R (-13.8 to -9.4)
pSPR1.0L (-6.9 to -5.9)
pSBH3.8L (-6.1 to -2.0)
pSRP3.5L (-2.1 to +1.4)
pIR1.6R (+1.6 to +3.0)
pIR3.5R (+3.3 to +6.8)
pIR3.5L (+3.3 to +6.8)
pIR2.6L (+10.8 to +13.4)

Figure 3. Control blot hybridizations. Only the control blots corresponding to the probes in figure 2 are shown. From left to right the lanes on each blot contain: 5 pg of the leftward transcript from the same template as the probe, 50 pg of the leftward transcript, 5 pg of the rightward transcript, 50 pg of the rightward transcript. These blots were made in parallel and were hybridized in the same container with those in Figure 2. In each case the exposure shown is the same as the exposure for the corresponding blot in Figure 2.


A
B
C
D
E
F
G
H
L
R
pIR4.4R
pSPR1.0L
pSBH3.8L
pSRP3.5L
pIR1.6R
pIR3.5R
pIR3.5L
pIR2.6L

of these transcripts also hybridize with two other leftward probes (pSBH3.8 and pSPR1.0) which span the regions between coordinates -6.9 and -2.0 (see Figure 1). Therefore, these transcripts must span the -2.3 to -2.1 segment in which part of the *hdp-α* gene has been mapped (see above and Chapter 1).

Higher resolution mapping with probes containing sequences from the region between coordinates -3.3 and +1.4 (Figure 4) indicate that the apparent 3' end of the 2.0 kb transcript lies between coordinates -2.1 and -1.6 (Figure 4B,D,E) while the 3' end of the 4.0 kb transcript apparently lies between -1.6 and -1.0 (Figures 4D,E and 5). Thus, in addition to the *hdp-α* locus, the 4.0 kb transcript overlaps at least part of the 800 bp segment between -1.6 and -0.8 in which sequences needed for the Bx function have been mapped (see Figure 5). Our results also show that the 2.0 kb and 4.0 kb transcripts are coextensive over the region between -5.9 and -2.1 (Figures 2B and 4B). It is therefore possible that they share sequence content.

Other transcripts that may be associated with the *Bx* and *hdp-α* loci are detected by probes on either side of the region encompassed by the 4.0 kb and 2.0 kb RNAs described above. On the left side, a rightward runoff probe (i.e., detecting RNAs that are transcribed from right to left) made from the plasmid pIR4.4 (-13.8 to -9.4) detects at least six different RNAs which range in size from 5.7 to 0.8 kb (Figure 2A). On the right side, a rightward probe made from pIR1.6 (+1.6 to +3.0) hybridizes with an identical set of transcripts (Figure 2E). Since both probes originate from the same strand, it is possible that a single large transcription unit, which produces numerous processed RNAs, spans the region between them. Because probes from the +1.6 and -9.4 region do not detect these transcripts, such a transcription unit would have to contain a large intron encompassing the entire region in which we detected the 2.0 kb and 4.0 kb transcripts as well as at least part of the *Bx* and *hdp-α* loci (Figure 1).

Figure 4. Mapping of the 3' ends of the 2 kb and 4 kb transcripts. Each blot was made with 2 micrograms of 0-4 hr and 10 micrograms of 60-72 hr poly(A)$^+$ RNAs. Probe coordinates are indicated under the blots. All probes are leftward runoff transcripts.

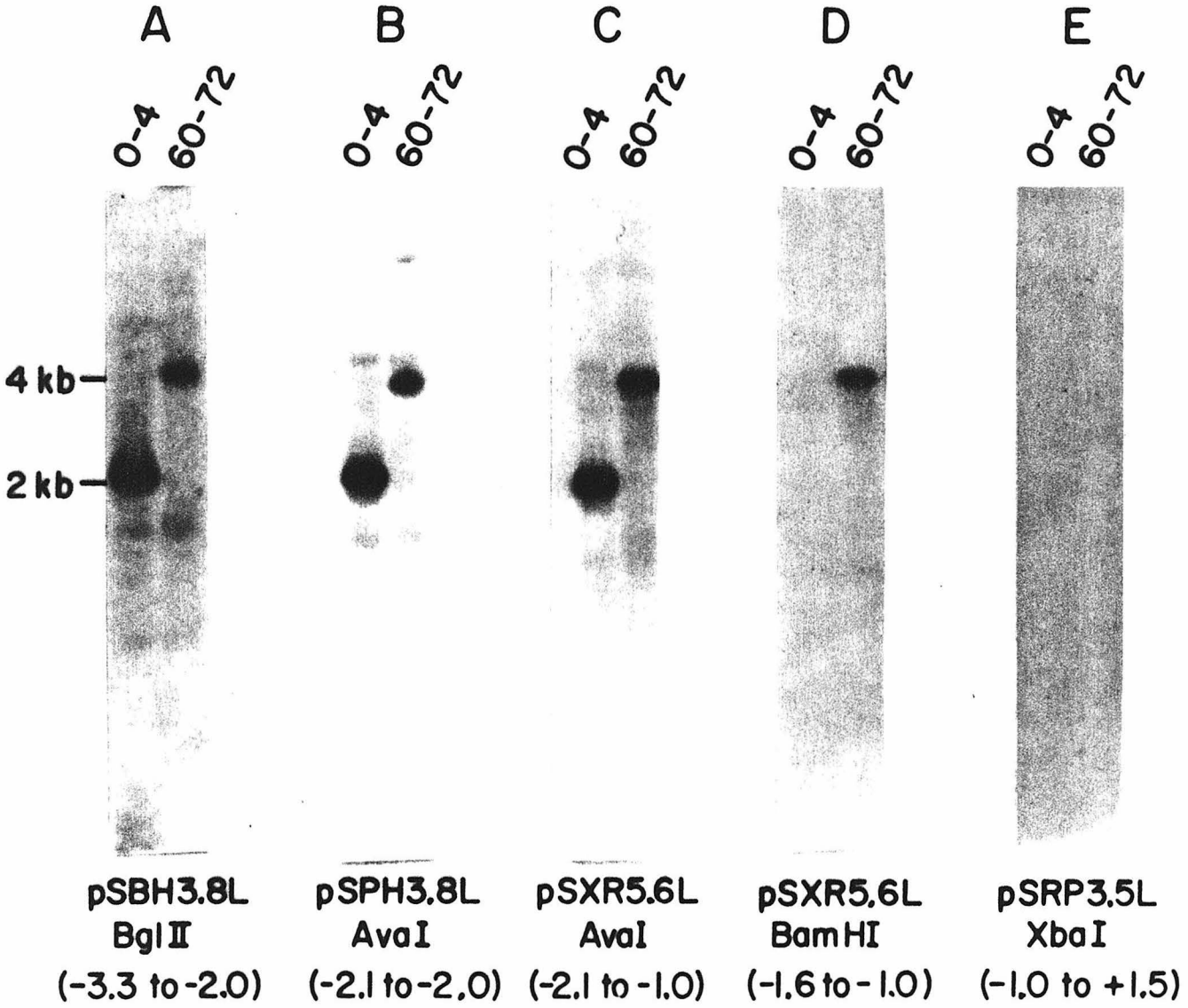
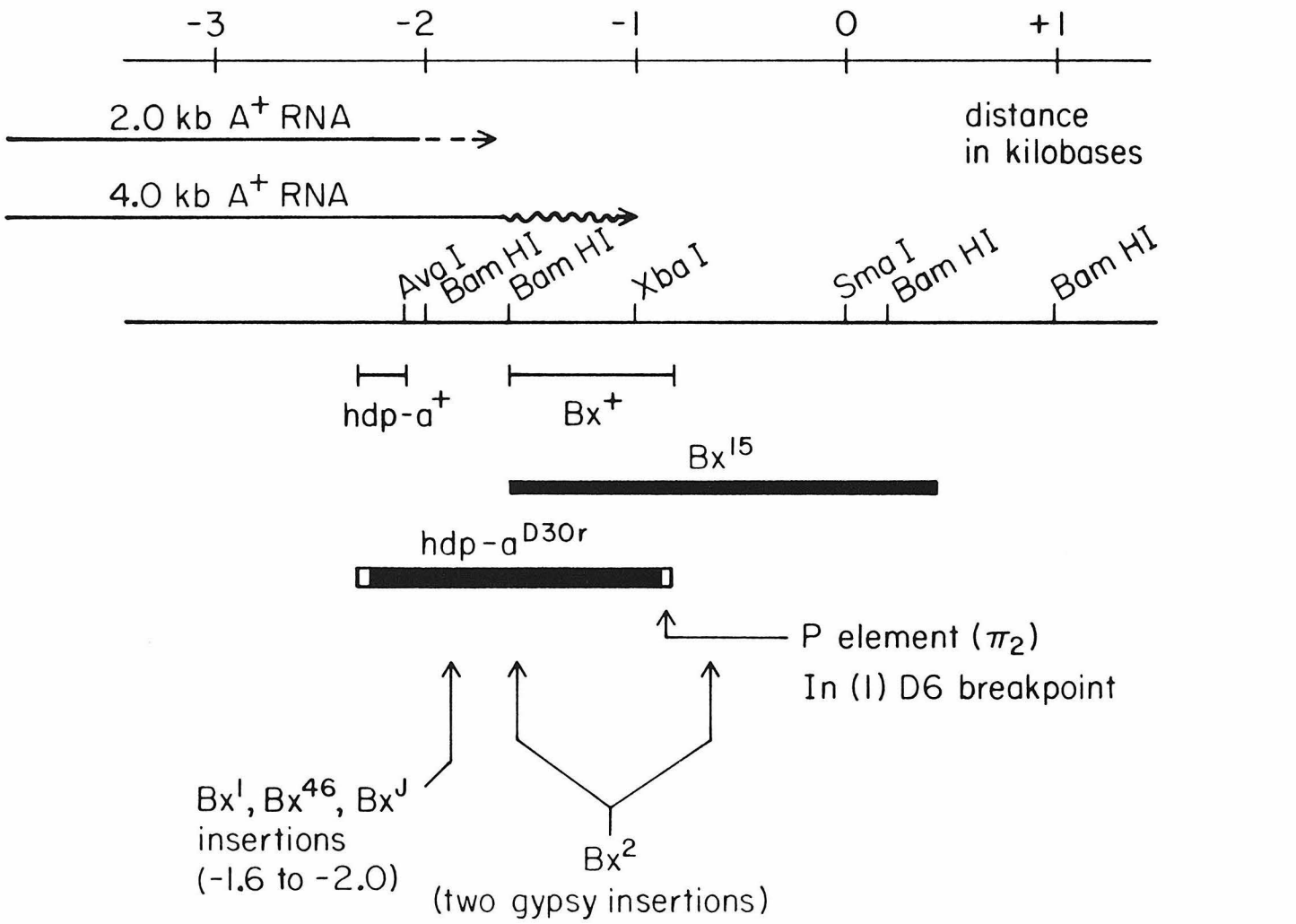

A
B
C
D
E
0-4
60-72
4 kb
2 kb
pSBH3.8L
BglII
(-3.3 to -2.0)
pSPH3.8L
AvaI
(-2.1 to -2.0)
pSXR5.6L
AvaI
(-2.1 to -1.0)
pSXR5.6L
BamHI
(-1.6 to -1.0)
pSRP3.5L
XbaI
(-1.0 to +1.5)

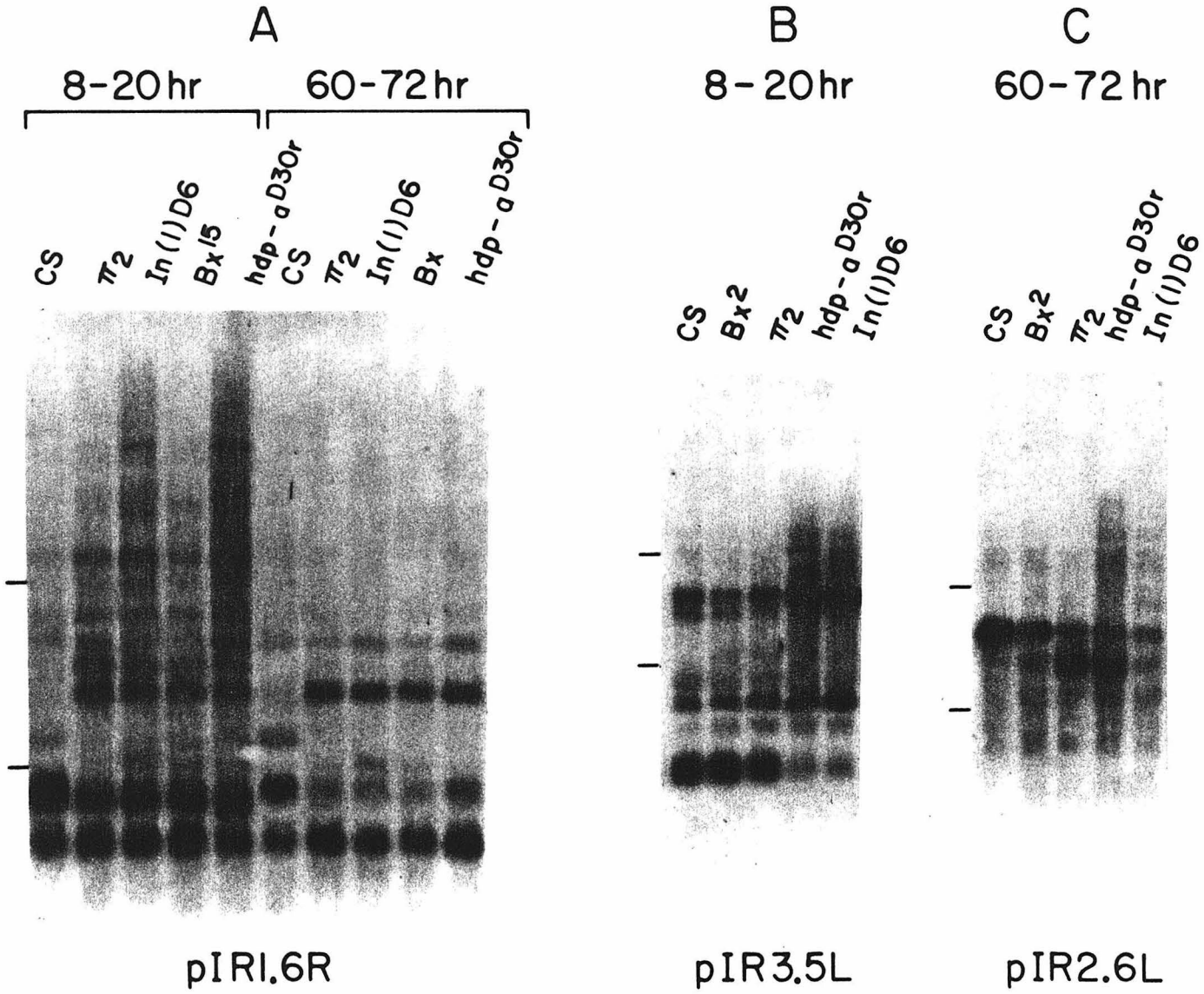
Figure 5. Positions of *Bx-hdp-a* mutations and transcripts. The arrows above the restricton map indicate the region which is homologous to the 3' segments of the 2 kb and 4 kb transcripts. The dashed part of the 2 kb arrow indicates the the region in which its 3' end is located. The wavy part of the 4 kb arrow indicates the region in which its 3' end is located. Below the restriction map the lines labeled $hdp\text{-}a^{+}$ and Bx^{+} denote the segments in which some of the sequences essential to these functions lie. The dark bars indicate the regions deleted in Bx^{15} and $hdp\text{-}a^{D30r}$ the unfilled portion of the hdp-a^{D30r} bar indicates the uncertainty about the precise endpoints of this deletion. The transposon insertion sites for each of the four *Bx* insertion mutants that we examined RNA from is indicated.


-3
-2
-1
0
+1
2.0 kb A+ RNA
distance
in kilobases
4.0 kb A+ RNA
Ava I
Bam HI
Bam HI
Xba I
Sma I
Bam HI
Bam HI
hdp-a+
Bx+
Bx15
hdp-aD30r
P element (π2)
In (1) D6 breakpoint
Bx1, Bx46, BxJ
insertions
(-1.6 to -2.0)
Bx2
(two gypsy insertions)

To test this possibility we investigated the structure of these RNAs in the *In(1)D6, y cho hdp-b* strain. The inversion breakpoint located near coordinate -0.8 in this strain is between the positions of the two probes which detect the set of transcripts in question. If a large transcription unit spans the site of this breakpoint it would be disrupted by this inversion and we would expect the RNAs transcribed such a unit to be altered in size in inversion bearing flies. Shown in Figure 6A is a blot made with RNA from the *In(1)D6, y cho hdp-b* strain which has been hybridized with the rightward pIR1.6 runoff probe. In both 8-20 hr embryos and 60-72 hr larvae no differences are observed between RNAs isolated from *In(1)D6, y cho hdp-b* and its parental wild-type strain, π_2. Therefore, it seems quite unlikely that these transcripts actually span the site of this breakpoint. We conclude that the transcripts detected by these probes originate either from two separate transcription units or from transcription unit(s) that are located elsewhere in the genome but share sequence homology with the probes used here.

Only the 2.0 kb and 4.0 kb transcripts are qualitatively altered by *Bx* and *hdp-a* mutations. Because several different transcripts are found near the *hdp-a* and *Bx* loci, it is unclear, from the above analysis, which, if any, of them encodes the *hdp-a*$^+$ gene product. For this reason we examined the effects of *Bx* and *hdp-a* mutations upon all of the transcripts which are homologous to the 34 kb region. In each case we compared mutant strains only with a wild-type strain of a similar genetic background. Strains which are isogenic to Canton-S were used to prepare RNA from the mutants Bx^1, Bx^2, Bx^{46} and Bx^J (see Materials and Methods for details about the construction of isogenic strains). RNAs isolated from the mutants Bx^{15}, hdp-a^{D30r}, and *In(1)D6, y cho hdp-b* were compared with RNA from π_2 (the wild-type strain from which these mutants were isolated). By doing this we hope to minimize differences in RNA structure and amount that originate from genetic variables located outside of the *Bx-hdp-a* region.

Figure 6. Blot hybridization analysis of RNAs from the 34 kb region in *Bx* and *hdp-a* mutations. Six micrograms of poly(A)$^+$ RNA from each of the indicated strains was hybridized with: (A) a rightward runoff transcript from pIR1.6 (+1.6 to 3.0), (B) a leftward probe from pIR3.5 (+3.3 to +6.8) and (C) a leftward probe from pIR2.6L. The two prominent bands in Canton-S (CS) RNA in part A are the 1.1 kb and 0.8 kb species. 18S and 28S rRNA markers are indicated at the left side of each blot.


A
8-20hr
60-72hr
CS
π2
In(1)D6
Bx15
hdp-aD30r
CS
π2
In(1)D6
Bx
hdp-aD30r
pIR1.6R
B
8-20hr
CS
Bx2
π2
hdp-aD30r
In(1)D6
pIR3.5L
C
60-72hr
CS
Bx2
π2
hdp-aD30r
In(1)D6
pIR2.6L

In Figure 6A we compare RNA from Bx^{15} and $hdp\text{-}a^{D30r}$ to that of their parental strain, π_2, using the rightward probe from pIR 1.6. Neither the *Bx* nor the *hdp-a* mutation appears to have any effect on the transcripts that are homologous to this probe in either 8-20 hr embryos or 60-72 hr larvae. We note, however, that several of the bands detected in strains with a π_2 genetic background differ from those found in Canton-S. Because both the π_2 and Canton-S strains are wild type, these differences must not result in any visibly abnormal phenotype.

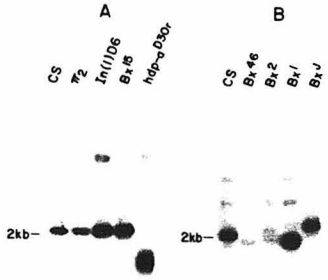
Despite its location to the right of the *In(1)D6, y cho hdp-b* inversion breakpoint, a 0.7 kb transcript, which is detected by the leftward probe from pIR3.5 (Figure 6B) was noticeably decreased in amount in the $hdp\text{-}a^{D30r}$ strain as compared to π_2. However, we do not believe that this transcript is encoded by the *hdp-a* gene because it is also reduced in amount in *In(1)D6,y cho hdp-b* flies, which are $hdp\text{-}a^{+}$. Also a quantitative comparison of the strains Bx^2 and Canton-S failed to detect any quantitative increase in the amount of this transcript in the Bx^2 strain (see Table 2).

Examination of the 3.3 kb and 2.6 kb transcripts that are homologous to the leftward probe from pIR2.6 did not reveal any substantial differences between Bx^2 and Canton-S or between $hdp\text{-}a^{D30r}$ and π_2. However, again we note that there are differences between Canton-S and π_2.

In contrast, the 2.0 kb and 4.0 kb RNAs, which were shown above to overlap part of the *hdp-a* locus, are consistently affected both by *hdp-a* and by *Bx* mutations. Figure 7 shows blot hybridizations using a runoff probe made from pSPR1.0 (-6.9 to -5.9). This probe is homologous to both the 2.0 kb and 4.0 kb RNAs of Canton-S (Figure 2B). 0-4 hr RNAs from several *Bx* mutations, which are associated with DNA insertions in the region between -1.5 and -2.0 (Chapter 1 and Figure 5), are examined in Figure 7B. In each of these strains the 2.0 kb transcript is replaced by a different size RNA species. In some cases (Bx^J and

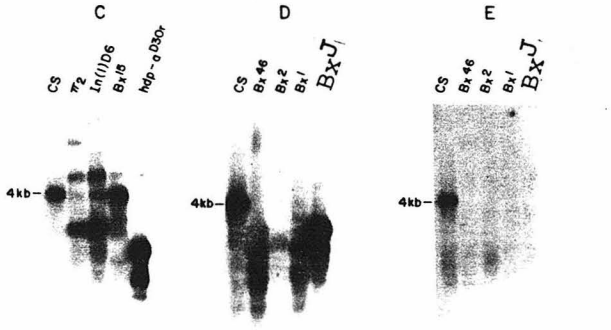
Figure 7. Effects of *Bx* and *hdp-a* mutations on the 2 kb and 4 kb transcripts. (A,B) Five micrograms of 0-4 hr poly(A)$^+$ RNA from the indicated strains. (C,D,E) Ten micrograms of 8-20 poly(A)$^+$ RNA from the same group of strains as in A and B. (F,G,H) Ten micrograms of 60-72 hr poly(A)$^+$ RNA. The runoff probes that were used are indicated under each blot.

0-4 hr



pSRP 1.0L
(-6.9 to -5.9)

8-20 hr

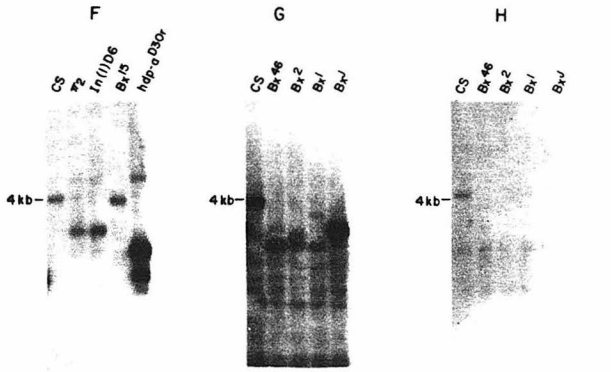


pSRP 1.0L
(-6.9 to -5.9)

pSRP 1.0L
(-6.9 to -5.9)

pSXR5.6L BamHI
+
pSRP3.5L XbaI
(-1.6 to +1.4)

60-72 hr



pSRP 1.0L
(-6.9 to -5.9)

pSRP 1.0L
(-6.9 to -5.9)

pSXR5.6L BamHI
+
pSRP3.5L XbaI
(-1.6 to +1.4)

Bx^{2}) aberrant transcripts that are larger than the wild-type 2.0 kb species are found, while in others *Bx*^{1} and *Bx*^{46} they are smaller than 2.0 kb (see Table 2 for transcript sizes.

A 2.0 kb RNA which comigrates with that of Canton-S was detected in 0-4 hr embryos from the π_2 strain (Figure 7A). This transcript is not altered in size in the *Bx*^{15} or *In(1)D6,y cho hdp-b* strains (Figure 7A), but it is replaced by two shorter transcripts in the *hdp-a*^{D30r} strain. This change in size is not surprising since part of the 2.0 kb transcript originates from sequences that are located between -2.1 and -2.0 (see Figures 4C and 5), a region which is deleted in *hdp-a*^{D30r} (Chapter 1).

None of the abnormal size transcripts found in mutant 0-4 hr embryos were found in either 8-20 hr embryos or in 60-72 hr larvae (see Figure 7C,D,F,G and Table 2). We interpret this to mean that the expression of these abnormal transcripts is developmentally controlled in the same manner as is the 2.0 kb wild-type transcript.

During these two later stages different abnormal size transcript(s) appear in each of the *Bx* mutant strains examined (see Figure 7C,D,F,G). In these mutants the 4.0 kb species which is normally present (in Canton-S) is reduced or missing (see Table 2 for sizes of abnormal transcripts). In each of the strains examined the abnormal transcript(s) appear to be the same in both the 8-20 hr and the 60-72 hr stages (compare Figure 7C,D to 7F,G). As is the case with the wild-type 4.0 kb transcript, these abnormal RNAs are expressed at uniformly lower levels during the 60-72 hr larval stage than they are during the 8-20 hr stage and they are not detected at all in 0-4 hr embryos. Therefore, it appears that each of these transcripts is developmentally regulated in the same way as is the 4.0 kb species in Canton-S.

We have noted above that the 4.0 kb transcript from *Canton-S* spans the region between -5.9 and -1.6. Since a number of Bx mutations are associated with transposon insertions located inside the region encompassed by the 4.0 kb transcript (see Figure 5), we wondered whether the mutant transcripts might result from the truncation of the the normal 4.0 kb transcript within the inserted transposons. To test this we hybridized both Canton-S and the mutant RNAs with a combination of two probes which cover the region between coordinates -1.6 and +1.4. This region is entirely downstream from the sites at which the transposons are inserted in Bx^1, Bx^{46} and Bx^J (see Chapter 1). The Bx^2 strain has two gypsy transposons inserted within the region covered by these probes. One is located between -1.5 and -1.6 while the other is near -0.6. The blot in figure 7E shows that although the 4.0 kb transcript from Canton-S 8-20 hr embryos clearly hybridizes with this downstream probe, none of the abnormal transcripts expressed in the four insertion bearing strains did so. Therefore, it appears that these insertion mutations cause premature termination of the 4.0 kb RNA at a site upstream to or within the inserted elements. The sequences which are missing from the mutant transcripts, but are present in the wild-type 4 kb transcript, include part of the 800 bp segment between -1.6 and -0.8 in which the Bx^+ locus maps.

We have also examined the expression of these downstream sequences in insertion mutant RNAs isolated from 60-72 hr larvae (Figure 7H). Again, the wild-type 4.0 kb species was detected with this probe in Canton-S, but no transcripts were detected in *Bx* mutant strains. However, the sensitivity of this experiment was limited by the low abundance of these transcripts (estimated to be less than 10^{-5} of poly(A)$^+$ RNA from Figures 2 and 3) and the relatively weak signal obtained with probes from the region between +1.5 and -2.1 (compare Figures 2D and 3D to Figures 2B and 3B).

Analysis of 8-20 hr and 60-72 hr transcripts expressed in strains with a π_2 genetic background was complicated by the fact that the π_2 strain itself expresses only a low level of the 4.0 kb transcript (Figure 7C,F). This wild-type strain also expresses, at a higher level, a 2.5 kb transcript which is not seen in Canton-S. An obvious difference between the DNA of the π_2 and Canton-S strains, which might account for this anomaly, is the presence of a P element inserted at -0.8 in the π_2 strain. Although this site is downstream of the apparent 3' end of the 4.0 kb transcript, it is possible that the P element affects this transcript at a distance or that the true 3' end of the 4.0 kb transcript is located farther downstream, outside the 34 kb region that we examined.

Examination of 8-20 hr and 60-72 hr RNAs from *hdp-a*^{D30r}, a strain which bears a deletion mutation that removes 1500 bp of DNA between -0.8 and -2.3 (as well as the entire P element present in π_2) reveals that it expresses two short transcripts of lengths 1.8 kb and 1.4 kb (Figure 7C,F). To account for the size differences between these two transcripts and the 4.0 kb wild-type transcript, other sequences, which are outside of the deleted region, must also be missing from the mutant transcripts in this strain.

The *Bx*^{15} strain, which also results from a deletion mutation that removes the P element and flanking sequences, expresses a 4 kb transcript. Although similar in size to the wild-type RNA, this transcript must have a different structure, since the entire region between -1.0 and -1.6 is missing in the *Bx*^{15}train (see Appendix 1 and Figure 5).

In summary, we find that the 2.0 kb and 4.0 kb wild-type transcripts are altered in length or in structure in all mutants studied with the single exception of the 2.0 kb transcript in *Bx*^{15}. The latter fact is not surprising because the *Bx*^{15} mutation is a deletion extending rightwards from -1.6 (i.e., just beyond the 3' end of the 2.0 kb transcript). We also note that the π_2 wild-type strain expresses only

low levels of the 4.0 kb transcript. The majority of transcripts made during 8-20 hr and 60 to 72 hr stages in this strain are 2.5 kb in length.

Quantitative effects of *Bx* and *hdp-a* mutations. Because *Bx* mutations cause an excess of gene function, we wanted to determine whether the aberrant transcripts which we found in mutant strains are expressed at higher levels than are wild-type transcripts. To quantify the amount of radioactive probe that hybridizes with transcripts on the blots shown in Figure 7 we scanned our autoradiographs using a densitometer and then integrated the peak areas. After hybridization with pSPR1.0L each of our blots was hybridized with a nick-translated probe specifically homologous to the 3' end of the 1.75 kb RNA transcript from the actin 57A gene (Not shown) (Fyrberg, et al. 1983). Densitometric values were obtained from the actin bands and used to correct for differences in the actual amount of poly(A)$^+$ RNA present on each blot lane (see Table 2). For this analysis we have assumed that the amount of this actin transcript is not affected by *Bx* and *hdp-a* mutations.

In Table 2 we show the corrected relative values for the major transcript(s) which are detected in each of the strains examined. In each case we have normalized these numbers so that the reference wild-type strain (same genetic background) has a value of 1.0. By examining the integrated values obtained from only those areas of the blots which contain an identifiable band we might be missing RNAs of heterogeneous size, which are expressed in the mutants. It is conceivable that such a heterogeneous set of RNAs could make up the majority of those transcripts which are detected by a particular probe and yet appear only as a diffuse smear on our blots. For this reason we measured the total signal in each lane by integrating all areas which showed a density higher than that of the film background. These values, which are shown in Table 3, were corrected and normalized to the wild-type values in the same manner as were the major peak values were.

Table 2
Hybridization to Major Band Only

stage	strain	wing phene.	RNA size	actin[1]	Bx-hdp-a[2]
0-4hr	CS	wild-type	2.0kb	0.7	0.8
	π_2	wild-type	2.0	0.6	1.0
	In(1)D6	heldup[3]	2.0	1.6	0.8
	Bx^{15}	severe	2.0	0.5	2.5
	hdp-a^{D30r}	heldup	1.45	0.7	1.4
	CS	wild-type	2.0	0.7	1.0
	Bx^{46}	moderate	1.75	0.6	0.4
	Bx^{2}	moderate	1.9,2.2	0.5	0.8
	Bx^{1}	weak	1.8	0.9	1.2
	Bx^{J}	severe	2.4	0.8	1.1
8-20hr	CS		4.0	1.0	1.2
	π_2		4.0	1.7	0.3
			2.5	1.7	0.7
	In(1)D6		4.0	1.9	0.4
			2.5	1.9	1.4
	Bx^{15}		4.0	1.4	1.5
	hdp-a^{D30r}		1.8	0.7	4.7
			1.4	0.7	2.1
	CS		4.0	1.6	1.0
	Bx^{46}		2.3	1.2	1.0
	Bx^{2}		2.4,2.8	0.6	0.6
	Bx^{1}		2.3	1.8	0.5
	Bx^{J}		2.7	1.0	1.9
60-72hr	CS		4.0	2.1	1.1
	π_2		2.5	2.7	1.0
	In(1)D6		2.5	2.8	1.4
	Bx^{15}		4.0	2.1	1.4
	hdp-a^{D30r}		1.8	3.4	4.2
			1.4	3.4	0.5
	CS		4.0	5.4	1.0
	Bx^{46}		2.3	4.3	1.2
	Bx^{2}		2.4,2.8	4.0	1.2
	Bx^{1}		2.3	4.4	0.7
	Bx^{J}		2.7	4.9	2.3

[1]Actual integrated density from actin 1.75 kb band in O.D.xm.m.[2]. These values were used to correct for differences in the amount of RNA loaded.

[2]These values are corrected and normalized such that the wild-type strain has a value of 1.0. In the case of 8-20 hr embryos derived from the pi_2 strain the values were normalized so that the sum of the 4 kb and the 2.5 kb bands would be 1.0.

[3]This is a result of the heldup-b mutation carried by this strain.

Table 3
Hybridization to Entire Lane

stage	strain	wing phene.	RNA size	actin[1]	Bx-hdp-a[2]
0-4hr	CS	wild-type	2.0kb	0.7	0.9
	TT_2	wild-type	2.0	0.6	1.0
	In(1)D6	heldup[3]	2.0	1.6	0.9
	Bx^{15}	severe	2.0	0.5	2.2
	hdp-a^{D30r}	heldup	1.45	0.7	1.7
	CS	wild-type	2.0	0.7	1.0
	Bx^{46}	moderate	1.75	0.6	0.6
	Bx^{2}	moderate	1.9,2.2	0.5	0.9
	Bx^{1}	weak	1.8	0.9	1.1
	Bx^{J}	severe	2.4	0.8	1.1
8-20hr	CS		4.0	1.0	1.1
	TT_2		4.0	1.7	1.0
			2.5		
	In(1)D6		4.0	1.9	0.4
			2.5		
	Bx^{15}		4.0	1.4	2.1
	hdp-a^{D30r}		1.8	0.7	4.9
			1.4		
	CS		4.0	1.6	1.0
	Bx^{46}		2.3	1.2	1.3
	Bx^{2}		2.4,2.8	0.6	1.0
	Bx^{1}		2.3	1.8	0.7
	Bx^{J}		2.7	1.0	2.1
60-72hr	CS		4.0	2.1	1.1
	TT_2		2.5	2.7	1.0
	In(1)D6		2.5	2.8	1.4
	Bx^{15}		4.0	2.1	1.4
	hdp-a^{D30r}		1.8	3.4	2.2
			1.4		
	CS		4.0	5.4	1.0
	Bx^{46}		2.3	4.3	1.1
	Bx^{2}		2.4,2.8	4.0	1.1
	Bx^{1}		2.3	4.4	0.9
	Bx^{J}		2.7	4.9	1.4

[1]Actual integrated density from actin 1.75 kb band in O.D.xm.m.[2]. These values were used to correct for differences in the amount of RNA loaded.

[2]These values are corrected and normalized such that the wild-type strain has a value of 1.0.

[3]This is a result of the heldup-b mutation carried by this strain.

In several cases we observed that the abnormal transcripts of mutants were present in significantly higher amounts than were the normal size transcripts from wild-type stains. In 0-4 hr embryos this was most noticeable in the Bx^{15} mutant. The 2.0 kb transcript, which is present in Bx^{15} at this stage, hybridized about 2.5 times as strongly as the 2.0 kb transcript from the π_2 strain. However, in most of the *Bx* mutants that we examined the major 0-4 hr transcripts were expressed at near normal or below normal levels (Bx^{46}, Bx^{2}, Bx^{1} and Bx^{J}). Normalized values from mutant lanes were not significantly higher than those obtained from integrating only the peaks originating from the major transcripts. This indicates that at this stage the mutant strains do not express a significantly higher amount of heterogeneous size transcripts than do wild-type strains.

In 8-20 hr embryos the major Bx^{15} transcript (4 kb) is apparently expressed at only 1.5 times the level of the two major transcripts (combined) from the π_2 strain. The 2.8 kb transcript found in the Bx^{J} strain hybridized about twice as strongly as did the wild type 4.0 kb RNA species from Canton-S. As was true in the 0-4 hr embryos the 8-20 hr embryos from Bx^{46}, Bx^{2} and Bx^{1} did not appear to overproduce any RNA species which is homologous to the pSPR1.0L probe. The most dramatic difference in RNA levels that we saw in this experiment was observed when we compared *hdp-a*^{D30r} to its parental strain, π_2. Relative to the sum of the two major species detected in π_2 RNA the 1.8 kb transcript expressed by this mutant hybridized to over four times as much probe. A less abundant 1.4 kb species hybridized to twice as much probe as π_2. The increased level of transcripts in *hdp-a*^{D30r} is apparent in Figure 7C even though the *hdp-a*^{D30r} lane has less than half as much poly(A)$^{+}$ RNA loaded onto it as π_2 (as measured by hybridization with the actin RNA probe, Table 2).

Higher levels relative to wild type were obtained for each of the *Bx* mutant RNAs when the integrated density of each entire lane was determined. This

indicates that, in addition to the major transcripts that we have identified in the *Bx* mutant strains, a number of minor transcripts are present during the 8-20 hr stage which are not found in wild-type. These transcripts are probably contained in some of the the faint bands and heterogeneous "smears" which are seen in the mutant lanes of the blots in Figures 7C and 7D.

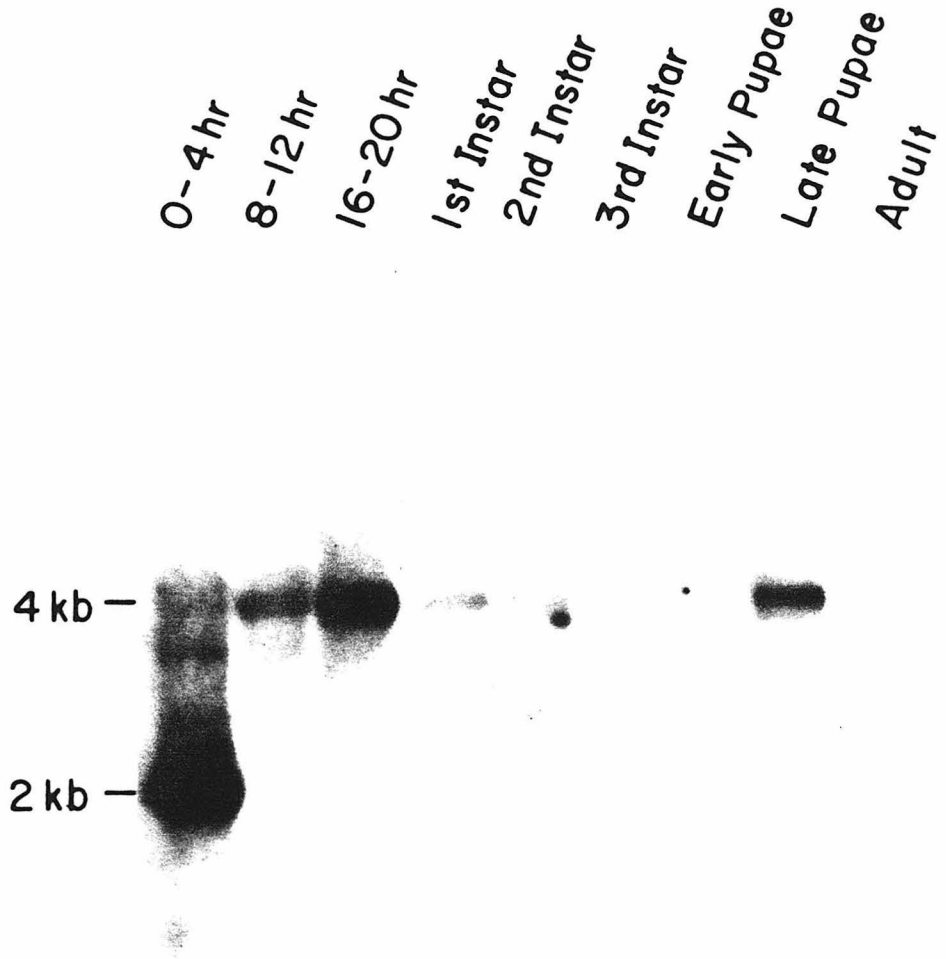
We were particularly interested in measuring transcript levels in mutant 60-72 hr larvae. Morphological studies of larval wing discs (Fristrom, 1969; Waddington, 1940) as well as clonal analysis of genetic mosaics (Santamaria and Garcia-Bellido, 1975) indicate that the wing scalloping phenotype observed in *Bx* individuals is a consequence of cell death which occurs shortly after this period. Therefore we might expect that any increase in RNA levels, which contributes to the *Bx* phenotype, would be apparent at this time. As was true in the 8-20 hr stage, only the Bx^J and $hdp\text{-}a^{D30r}$ strains expressed more than twice the wild-type levels of their major transcripts. In this case the ratio between mutant and wild-type levels was lower when the entire gel lane was integrated, but this is probably due only to the higher noise levels which are apparent on these blots.

In summary, we find that most of the excess-of-function *Bx* mutants that we examined showed no significant increase in band intensity for either the 2.0 kb or the 4.0 kb related transcripts. However, in flies bearing the Bx^{15} mutation the 2.0 kb band is increased as much as 2.5X in 0-4 hr embryo RNA. For the 4.0 kb related transcripts, the most significant change in intensity was observed in the Bx^J individuals. During the 60-72 hr stage the 2.7 kb band present in this strain hybridizes 2.3X as intensively as the 4.0 kb band in Canton-S. We also found that the 1.8 kb RNA expressed in $hdp\text{-}a^{D30r}$ during this stage had a hybridization intensity that was 4.2X the wild-type level.

Developmental analysis of the 2.0 kb and 4.0 kb transcripts. As mentioned above, transcripts encoding those gene products that are controlled by the *Bx*

Figure 8. Developmental RNA blot. Eight micrograms of poly(A)$^+$ RNA from the indicated stages were hybridized with nick-translated ^{32}P-p24R8.0B. This probe covers the region between -6.9 and +1.4. The "first instar" stage is 32-36 hrs, the "second instar" stage is 52-56 hrs, "third instar" is 80-90 hrs, "early pupae" are 6 to 10 hrs post-puparium formation and "late pupae" are from 70-80 hrs post-puparium formation.

p24R8.0B


0-4 hr
8-12 hr
16-20 hr
1st Instar
2nd Instar
3rd Instar
Early Pupae
Late Pupae
Adult
4 kb
2 kb

locus and responsible for the wing scalloping phenotype are expected to be present in 72 hr larvae or at some period prior to this stage. We have presented data indicating that the 4.0 kb transcript that overlaps the *Bx* and *hdp-a* loci is expressed during the 8-20 hr and 60-72 hr periods. To investigate the expression of this transcript during other developmental stages we examined poly(A)$^+$ RNA from several stages by blot hybridization. The blot shown in Figure 8 was hybridized with a nick-translated plasmid which covers the region between -6.9 and +1.4. It shows that the 4.0 kb species is first expressed at some time between 4 and 12 hrs and persists through the remainder of embryonic development. Although this transcript is expressed in all larval stages examined, its level is substantially lower than during embryonic stages. The 4.0 kb RNA was also detected in a late pupal stage but not in early pupae or adults. The 2.0 kb RNA was found to be expressed in 0-4 hr embryos but was not detected in any other stage that we examined.

DISCUSSION

In order to determine whether or not the *Bx* locus represses the *hdp-a* gene's activity it is first necessary to identify the *hdp-a* gene product. Ideally, we would identify this product on the basis of a biochemical function that it accomplishes; however, nothing is known about the biochemical characteristics of the *hdp-a* gene product. Therefore, we have chosen to use the available cloned DNA sequences from the *Bx-hdp-a* region to search for RNA transcripts that might be encoded by the *hdp-a* gene. By performing blot hybridizations with single stranded probes derived from a 34 kb region near the *Bx* and *hdp-a* loci, we are able to detect a number of different transcripts which are homologous to these sequences (Figure 1). By observing the effects of *Bx* and *hdp-a* mutations on each of these transcripts, we were able to identify at least one RNA species that is likely to be encoded by the *hdp-a* gene.

Location of transcripts relative to the *hdp-a* and *Bx* loci. Our previous molecular analysis of the DNA structure of 13 different *Bx* mutants and one *hdp-a* mutant identified a small region, between coordinates -2.3 and +1.0, which is altered in each of the mutant strains (Chapter 1). By examining overlapping small deletion mutations, we were able to identify a small segment (-1.6 to -0.8) which is required for the normal function of the *Bx* locus. Another segment (-2.3 to -2.1) is required for normal function of the hdp-a locus. Although these segments probably do not represent the entire extent of either the *hdp-a* or *Bx* functional domains, they do identify the positions of at least part of both of these loci. We also found that an inversion breakpoint that is located near -0.8 in the strain *In(1)D6,y cho hdp-b* does not affect the function of either *Bx* or *hdp-a* and therefore provides a right boundary on the extent of these genes (Engels and Preston, 1981; also see Chapter 1).

Our hybridization studies have identified two particular transcripts of lengths 2.0 and 4.0 kb, which are located in the same region as we would expect to find the *hdp-a*$^+$ gene. These two transcripts, which are 4.0 kb and 2.0 kb in length, both hybridize only to sequences on the centromere-distal (left) side of the *Bx* locus. This is the same side on which Lifshytz and Green (1979) placed a *hdp-a* point mutation by recombinational analysis. These RNAs are also to the left of the above-mentioned inversion breakpoint, which defines the right limit of the *hdp-a* locus. Both transcripts overlap the 200 bp segment (-2.3 to -2.1) in which part of the *hdp-a* locus lies and are oriented such that this segment is encoded near their 3' end. The apparent 3' end of the 2.0 kb transcript is a short distance downstream between -2.0 and -1.6, while the 3' end of the 4.0 kb transcript is farther downstream between -1.6 and -1.0. This means that the 4.0 kb transcript extends into the -1.6 to -0.8 region, in which part of the *Bx* locus maps, and that the 2.0 kb RNA does not. Although at present it is impossible to say whether or

not the 4.0 kb transcript contains those particular sequences within the -1.6 to -0.8 region which are required for *Bx* function, we note that the 4.0 kb transcript also spans the sites at which DNA insertions occur in several *Bx* mutant strains.

Developmental expression of transcripts near the *Bx* locus. Somatic genetic analysis and morphological observations of flies carrying *Bx* mutations have shown that the wing scalloping phenotype caused by these mutations is a consequence of cell death in the larval wing disc after about 72 hrs of development (Fristrom, 1969; Waddington, 1940; Santamaria and Garcia-Bellido, 1975). Analysis of the developmental expression of the 2.0 kb and 4.0 kb transcripts indicates that only the 4.0 kb transcript is expressed during the larval stages preceding the 72 hr stage when we would expect the *Bx-hdp-a* product to be active. For this reason we think that the 4.0 kb transcript is more likely to encode the $hdp\text{-}a^{+}$ gene product than is the 2.0 kb RNA.

Qualitative effects of mutations at the *Bx* locus on the 4.0 kb transcript. Since all of the *Bx* and *hdp-a* mutations that have been studied at the molecular level affect sequences which are within or very close to the transcription unit that encodes the 4.0 kb RNA, we expected this transcript to be altered by many, if not all, of these mutations. Our analysis of the RNAs produced by five different *Bx* mutant strains and one *hdp-a* mutant has confirmed this expectation. In all but one of the mutant strains examined, the 4.0 kb transcript is missing or greatly reduced in amount and is replaced by a smaller transcript. In the one exception, Bx^{15}, we know that sequences located between -1.6 and -1.0 have been deleted. Because these sequences are normally included in the Canton-S 4.0 kb transcript, the Bx^{15} transcript must have a different structure.

Four of the *Bx* mutants that we examined (Bx^{46}, Bx^{2}, Bx^{1} and Bx^{J}) are associated with transposon insertions in the -2.0 to -1.5 segment, which is located just upstream to the segment where we found sequences necessary for the *Bx*

function (Chapter 1). Each of these mutants produces a transcript that is smaller than the wild-type 4.0 kb species but is developmentally expressed in the same pattern as is the 4.0 kb transcript in Canton-S. Since probes made from sequences that are downstream from the inserted elements fail to hybridize with mutant RNAs, these shorter transcripts appear to be truncated near the sites of the transposon insertions.

Some of the transposons that are inserted in these Bx alleles have also been implicated in the truncation of transcripts at other gene loci in Drosophila. A gypsy transposon insertion like that in Bx^2 causes a transcript to be truncated at the *Hairy-Wing* locus (Campuzzano et al., 1986); a copia element (Bx^{46}) is known to truncate transcripts from the *white* locus (Levis et al., 1984; Pirrota and Brockl, 1984); and a roo-B104 element (Bx^1) apparently has a similar effect on RNA transcribed from the *Glued* locus (Swaroop et al. 1985).

We would like to point out what may be a significant parallel in the structure of transcripts from *Bx* insertion and *Bx* deletion mutations. In each of the 7 deletion mutants examined to date, most or all of the sequences between -1.6 and -0.8 are missing from the DNA. Therefore, these sequences cannot be included in the RNA transcripts expressed in these mutants. RNAs from each of the four insertion mutants examined above also lack these sequences due to apparent termination upstream. Since at least part of the *Bx* locus lies in the -1.6 to -0.8 region, it seems plausible that the principal defect in *Bx* mutants is the absence of sequences from the *Bx* locus itself in these transcripts. Whether this is true or not, the qualitative changes in the 4.0 kb transcript that are associated with *Bx* mutations and the lack of substantial quantitative changes in transcripts from the 34 kb region (see below) suggest that *Bx* mutations increase gene activity by altering the structure of transcripts rather than by affecting the amount of RNA produced.

Quantitative effects of *Bx* and *hdp-a* mutations. Genetic analysis of *Bx* mutant strains indicates that the wing scalloping phenotype results from an increase in *Bx-hdp-a*$^{+}$ gene activity. There are a variety of ways in which such an increase might come about. For instance structural changes in the RNA, like those we have observed in the 4.0 kb transcript, might be responsible for increasing the rate at which this RNA is translated or the rate at which the hypothetical encoded protein product functions. Another possibility is that the structural changes which we observe are merely coincidental and that *Bx* mutations also cause an increase in the amount of RNA present by affecting its transcription.

To determine whether *Bx* mutations cause an increase in the level of any of the transcripts that we studied, we quantified the signals on our autoradiographs and compared the amount of material which hybridized to mutant and wild-type transcripts at similar developmental stages. We found that among *Bx* mutants only the most severe alleles, *Bx*^{15} and *Bx*^{J} caused an elevation in the amount of any of the RNAs in the 34 kb region. In the *Bx*^{15} mutation we measured a 2.5 fold increase in the amount of 2.0 kb transcript present in 0-4 hr embryos. However, the mutant transcript, like the wild-type transcript, was not expressed at stages when the *Bx-hdp-a* product is expected to be active. The abnormal size transcript that replaces the 4.0 kb transcript in the *Bx*^{J} mutant strain is expressed at about twice the level of the wild-type transcript in Canton-S. This is true during both the 8-20 hr and the 60-72 hr stage. Comparisons of the wing scalloping phenotypes observed in *Bx*^{J} individuals to those observed in flies with multiple doses of the *Bx*$^{+}$ locus suggest that the *Bx*^{J} mutation causes at least a 5 fold increase in the level of gene activity (Lifshytz and Green, 1979). Thus, the increase that we measured is below the expected value. Other, less severe, *Bx* alleles had near normal or slightly below normal levels of hybridization as compared to wild type.

Because our quantitative analysis was done with whole fly RNA, it is possible that we are missing a substantial increase in transcript levels which occurs only in a small fraction of the cells that produce the *Bx* product. In this regard it would be informative to measure the level of transcript present in wing discs from mutant larvae.

We also examined the levels of two aberrant transcripts that are present in the *hdp-a*^{D30r} strain. We found that these transcripts are expressed at substantially higher levels than those expressed in either the π_2 or the Canton-S strains. This was true in both the 8-20 hr and the 60-72 hr stage. Although this mutation presumably results from the loss of sequences essential for the *hdp-a*$^{+}$ function, mapping of the deletion associated with it reveals that part of the *Bx* locus is also missing (presumably the *Bx* phenotype is not observed in *hdp-a*^{D30r} because it is suppressed in *cis* by the *hdp-a* defect). Therefore, it is possible that the elevated levels of these aberrant transcripts result from disruption of the repressive function of the *Bx* locus. Whether this is true or not, the *hdp-a*^{D30r} allele is a recessive loss-of-function mutation and therefore cannot result directly from an increase in the level of *hdp-a*$^{+}$ product. If the mutant transcripts are encoded by the *hdp-a* gene, then it seems likely that they are defective.

We would like to point out that the quantitative effects (both up and down) that we observe in mutant strains may be a consequence of the structural changes. For instance, the loss of sequences from mutant RNA transcripts may result in their increased stability and thus to a higher steady state level. In this regard it is interesting to note that a gypsy transposon induced RNA truncation has been found in the *Hairy-Wing*^{1} mutation, which causes a 5-20 fold increase in the amount of the truncated RNA as compared to the wild-type transcript (Campuzzano et al., 1986).

Because the 4.0 kb transcript is closely associated with the the *Bx* and *hdp-a* loci and is affected by all of the *Bx* and *hdp-a* mutations that have been examined, we are inclined to believe that this transcript is encoded by the hdp-a^+ gene itself. However, it is theoretically possible that *hdp-a* mutations affect other transcripts that are located some distance away. If this were so, then these distant transcripts might actually encode the *hdp-*a^+ gene product (for examples of such long range effects of mutations see: Brand et al., 1985; Campuzzano et al., 1985; Levis and Rubin, 1982; Zachar and Bingham, 1982). For this reason it was important to examine the effects of *Bx* and *hdp-a* mutations upon all of the transcripts that are encoded near the positions of of *Bx* and *hdp-a* mutations.

Among the transcripts found in the 34 kb region only one, aside from the 2.0 kb and 4.0 kb transcripts mentioned above, seems to be affected by the *hdp-*a^{D30r} mutation. This 0.7 kb transcript is located between coordinates +3.3 and +6.8 and is notably reduced in the *hdp-*a^{D30r} strain relative to π_2. However, this region is entirely to the right of the *In(1)D6, y cho hdp-b* breakpoint and is therefore outside of the region in which we would expect to find the *hdp-a* gene. Also, we note that this transcript is similarly reduced in RNA from the *In(1)D6, y cho hdp-b* strain. Since this inversion bearing chromosome complements *hdp-*a^+, it is doubtful that the observed reduction in the level of the 0.7 kb transcript causes any reduction in *hdp-*a^+ activity.

CONCLUSION

We conclude that the 4.0 kb transcript is the most likely candidate, among the transcripts that we have examined, to be encoded by the *hdp-a* gene. In some *Bx* mutants we have observed quantitative changes in the steady-state levels of this RNA that may contribute to the *Bx* phenotype. However, a more common characteristic of *Bx* mutations is their qualitative effect on the structure of this transcript. This suggests that changes in the structure rather than the amount of the *hdp-a* gene product lead to the excess of gene function in *Bx* mutants.

ACKNOWLEDGMENTS

We are grateful to Beverley Bond Matthews, Sue Celniker, Nevis Fregein and Mark Garfinkel for useful discussions. Lei Yu provided both helpful advice and plasmid DNA. This work was supported by a National Research Service Award (GM-07616) and a grant from the National Institutes of Health (GM-10991).

LITERATURE CITED

Bond, B.J. and N. Davidson. 1986. The *Drosophila melanogaster* actin 5C gene uses two transcription initiation sites and three polyadenylation sites to express multiple mRNA species. Mol. Cell. Biol. 6:2080-2088.

Brand, A.H., L.Breeden, J. Abraham, R. Sternglanz and K. Nasmyth. 1985. Characterization of a "Silencer" in yeast: A DNA sequence with properties opposite to those of a transcriptional enhancer. Cell 41:41-48.

Campuzano, S., L. Balcells, R. Villares, L. Carramolino, L. Garcia-Alonso and J. Modolell. 1986. Excess function *Hairy Wing* mutations caused by gypsy and copia insertions within structural genes of the *achaete-scute* locus of *Drosophila*. Cell 44:303-312.

Campuzano, S., L. Carramolino, C.V. Cabrera, M. Ruiz-Gomez, R. Villares, A. Boronat and J. Modolell. 1985. Molecular genetics of the *achaete-scute* gene complex of *D. melanogaster*. Cell 40:327-338.

Chirgwin, J.M., A.E. Przybyla, R.J. MacDonald and W.J. Rutter. 1979. Isolation of Biologically active ribonucleic acid from sources enriched in ribonuclease. Biochemistry 18:5294-5299.

Engels, W. R. and C. R. Preston. 1981. Identifying P factors in *Drosophila* by means of chromosome breakage hotspots. Cell 26:421-426.

Fristrom, D. 1969. Cellular degeneration in the production of some mutant phenotypes in *Drosophila melanogaster*. Mol. Gen. Genet. 103:363-379.

Fyrberg, E.F., J. Mahaffey, B.J. Bond and N. Davidson. 1983. Transcripts of the six *Drosophila* actin genes accumulate in a stage- and tissue-specific manner. Cell 33:115-123.

Green, M.M. 1953a. The *Beadex* locus in *Drosophila melanogaster*: On the nature of the mutants Bx^r and Bx^1. Genetics 38:91-105.

Green, M.M. 1953b. The *Beadex* locus in *Drosophila melanogaster*: Genetic analysis of the mutant Bx^{r49k}. Zeitschrift fur indukt. Abstammungs- and Vererbungslehre. 85:435-449.

Johnson, A.D. and I. Herskowitz. 1985. A repressor (MATα2 product) and its operator control expression of a set of cell-type specific genes in yeast. Cell 42:237-247.

Laski, F., D. C. Rio and G. M. Rubin. 1986. Tissue specificity of *Drosophila* P element transposition is regulated at the level of mRNA splicing. Cell 44:7-19.

Levis, R., K. O'Hare and G.M. Rubin. 1984. Effects of transposable element insertions on RNA encoded by the *white* gene fo *Drosophila*. Cell 38:471-481.

Levis, R. and G.M. Rubin. 1982. The unstable w^{DZL} mutation of *Drosophila* is caused by a 13 kilobase insertion that is imprecisely excised in phenotypic revertants. Cell 30:543-550.

Lifshytz, E. and M.M. Green. 1979. Genetic identification of dominant overproducing mutations: The *Beadex* gene. Mol. Gen. Genet. 171:153-159.

Miller, A., V.L. MacKay and K.A. Nasmyth. 1985. Identification and comparison of two sequence elements that confer cell-type specific transcription in yeast. Nature 314:598-602.

Oshima, Y. 1982. Regulatory circuits for gene expression: The metabolism of galactose and phosphate. pp. 159-180 in *The molecular biology of the yeast Saccharomyces*. J.N. Strathern, E.W. Jones and J.R. Broach, eds., Cold Spring Harbor Laboratories, Cold Spring Harbor, New York.

Pirrotta, V., and Ch. Brockl. 1984. Transcription of the *Drosophila white* locus and some of its mutants. EMBO 3:563-568.

Rozek, C.E. and N. Davidson. 1983. *Drosophila* has one myosin heavy-chain gene with three developmentally regulated transcripts. Cell 32:23-34.

Santamaria, P. and A. Garcia-Bellido. 1975. Developmental analysis of two wing scalloping mutants ct^6 and Bx^J of *Drosophila melanogaster*. Wilhelm Roux Arch. Entwicklungsmech. Organ. 178:233-245.

Swaroop, A., M.L. Paco-Larson and A. Garen. 1985. Molecular genetics of a transposon-induced dominant mutation in the *Drosophila* locus *Glued*. Proc. Natl. Acad. Sci. USA 82:1751-1755.

Waddington, C.H. 1940. The genetic control of wing development in *Drosophila*. J. Genetics 41:75-139.

Zachar, Z. and P.M. Bingham. 1982. Regulation of *white* locus expression: The structure of mutant alleles at the *white* locus of *Drosophila melanogaster*. Cell 30:529-541.

Zinn, K., D. DiMaio and T. Maniatis. 1983. Identification of two distinct regulatory regions adjacent to the human β interferon gene. Cell 34:865-879.

CHAPTER 3

P element-Mediated Transformation of *Drosophila melanogaster* with a 10.4 Kilobase Fragment from the *Beadex-heldup-a* region

WILLIAM W. MATTOX AND NORMAN DAVIDSON

SUMMARY

We have reintroduced a 10.4 kilobase DNA fragment from the *Bx-hdp-a* region into the *Drosophila* genome by germ line transformation. This fragment contains sequences that are essential to the wild-type *hdp-a* and *Bx* functions. It also includes all sequences known to hybridize with a 4 kb transcript, which is thought to be encoded by the $hdp\text{-}a^{+}$ gene. Despite this, we find that five independent insertions of this fragment fail to complement the loss-of-function $hdp\text{-}a^{D30r}$ mutation. This fragment also failed to enhance the excess-of-function wing scalloping phenotype of flies with multiple doses of the *Bx-hdp-a* region. We discuss the relevance of these findings to the structure and function of the 4 kb transcript.

INTRODUCTION

Dominant mutations at the *Beadex* locus of *Drosophila melanogaster* behave genetically as though they cause an excess of gene activity. This is evidenced by the observation that the scalloped wing phenotype that is associated with *Bx* mutations is mimicked in strains that carry extra doses of the wild-type *Bx* locus (Green, 1953a,b). This excess-of-function phenotype is *cis*-suppressed by recessesive loss-of-function mutations in the closely linked *heldup-a* (*hdp-a*) gene (Lifshytz and Green, 1979) (*hdp-a* mutations are themselves associated with a recessive abnormal wing position phenotype, but no wing scalloping is observed in any known *hdp-a* strain). This suppression suggests that *Bx* mutations induce the wing scalloping phenotype via the action of the $hdp\text{-}a^{+}$ gene. Because these loci are separated by only a very short genetic distance and interact in *cis*, Lifshytz and Green (1979) have proposed that these loci make up a single genetic unit in which the *Bx* locus is a regulatory element that controls the activity of the *hdp-a* structural gene. They have further suggested that *Bx* mutations cause an excess

of *hdp-a*$^{+}$ activity and that it is this increased *hdp-a*$^{+}$ activity which leads to the wing scalloping phenotype.

In the previous two chapters of this thesis we have shown that the molecular structure of *Bx* and *hdp-a* mutations is consistent with the idea that these loci are both parts of the same genetic unit. We have found that sequences essential for the *hdp-a*$^{+}$ function are located less than 1500 base pairs (bp) away from sequences that are essential to the repressive function of the *Bx*$^{+}$ locus. One mutation that removes DNA from both of these regions has a *hdp-a* phenotype and thus confirms that *hdp-a* mutations suppress *Bx* mutations in *cis*.

We have identified two transcripts that span at least some of the sequences which are required for the *hdp-a*$^{+}$ function. Of these transcripts only one, a 4 kb species, appears to be expressed during the developmental stages in which the *Bx-hdp-a* gene product is expected to be active. Each of the five *Bx* mutants and one *hdp-a* mutant that we examined express a structurally altered transcript in the place of the 4 kb wild-type transcript. The amount of altered transcript produced by one severe *Bx* allele (*Bx*^{J}) is twice that of the 4 kb transcript in Canton-S. Other transcripts which we detected in the 34 kb region did not appear be affected by *Bx* mutations. These results lead us to believe that the 4 kb transcript is likely to encode the *hdp-a*$^{+}$ activity.

Here we examine the ability of a 10.4 kb DNA fragment that contains all sequences which we found to be homologous with the 4 kb transcript, to complement the wing position defect of flies with an *hdp-a* mutation We also examine the ability of this fragment to enhance the wing scalloping phenotype in flies that are hyperploid for *Bx* and *hdp-a* loci.

MATERIALS AND METHODS

Fly Stocks. Mutant fly strains were provided by the Drosophila Stock Center at Caltech and by Madeline Crosby. Wild-type flies were spontaneously generated from old bananas ("Bob" the produce man at Ralphs, personal communication).

Germline transformation of *Drosophila* embryos. *Adh*^{fn23} *pr cn* embryos were microinjected with a solution containing 500 micrograms/ml of pNh10.4A and 100 micrograms/ml of phsπ (Steller and Pirrotta, 1985) by the method of Spradling and Rubin (1982). Embryos were dechorionated, dessicated, injected and allowed to develop to hatching at 18°C. They were then reared at room temperature. No heat shock was used. The flies that survived to adulthood were individually mated with *Adh*^{fn23} *pr cn*. F1 progeny were collected at three day intervals and examined for wing scalloping. All of these F1 flies were aged for five days and then selected on 7% ethanol for 16 to 20 hours (Vigue and Sofer, 1976). Ethanol-resistant individuals were then individually crossed to Adh^{fn23} pr cn to make several sublines from each transformed injectee. The Adh^{+} marker of each subline was mapped to a chromosome by making appropriate crosses with Adh null stocks carrying marked balancer chromosomes and observing the segregation of markers among the ethanol-resistant progeny. The chromosomes bearing inserted P elements were then made homozygous. All transformant lines were maintained in the presence of the *Adh*^{fn23} *pr cn* chromosome.

Genomic DNA blot hybridizations. DNA blots and hybridizations were done as described in Chapter 1 except that DNA was transferred to a nylon membrane (Hybond-N) and bound by placing the membrane on a UV transilluminator for one minute.

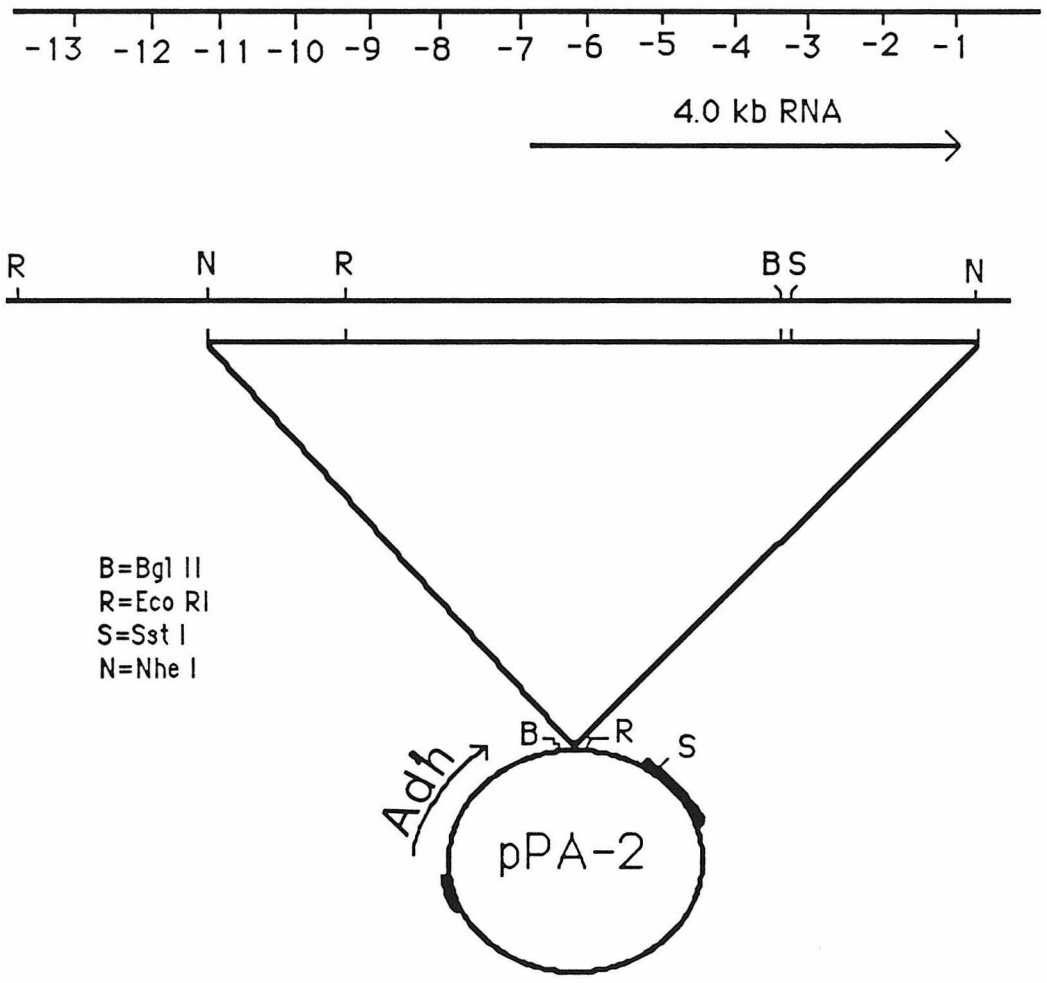
RESULTS

Plasmid construction and germ line transformation. Although our previous molecular studies on the DNA from *Bx* and *hdp-a* mutant strains defined a rightward (centromere-proximal) boundary to the extent of these loci, no leftward boundary was established. Therefore, in a formal genetic sense, we do not know how far leftward these loci extend. For this reason we used other criteria for selecting the particular restriction fragment that we would use for transformation experiments. The fragment that we selected, a 10.4 kb Nhe I fragment from Canton-S, includes the region extending from -11.2 to a point just to the left of -0.8 on the coordinate map in Figure 1. It therefore includes the 200 bp segment that is essential for the $hdp\text{-}a^{+}$ function (-2.3 to -2.1) and most of the 800 base pair segment in which at least part of the *Bx* locus lies (-1.6 to -0.8). The right end of this fragment is located about 30 bp to the left of the right limit of the *Bx* locus as defined by an inversion breakpoint which is Bx^{+} (K. O'Hare, personal communication). Also contained in this fragment is the entire region that we found to be homologous with the 4 kb transcript identified in Chapter 2. At least 4.5 kb of sequences residing upstream of this homology region are present in the 10.4 kb fragment.

We inserted this fragment into the XbaI site of the plasmid pPA-2 (provided by J. Posakony) by taking advantage of the complementary single stranded ends produced by Nhe I and Xba I cleavage. In the resulting plasmid, pNh10.4A, both the 10.4 kb Nhe I fragment and a 4.2 kb fragment that contains the *Drosophila* Adh^{+} gene are located between two P element termini. The total length of the hybrid P element is about 16 kb.

Embryos that carry a null mutation in the *Adh* gene (Adh^{fn23} *pr cn*) and wild-type alleles at both the *Bx* and *hdp-a* loci were coinjected with pNh10.4A and a helper plasmid, phsπ. This helper plasmid contains a defective P element that has

Figure 1. Restriction map of pNH10.4A. Coordinates are shown above the Canton-S restriction map. The 10.4 kb Nhe I fragment used is indicated below the restriction map. The plasmid map (circle) is that of pPA2. The segments corresponding to the P element termini are filled. The direction of transcription from the inserted 10.4 kb fragment and from the Adh gene are indicated by the arrows.


-13 -12 -11 -10 -9 -8 -7 -6 -5 -4 -3 -2 -1
4.0 kb RNA
R N R B S N
B=Bgl II
R=Eco RI
S=Sst I
N=Nhe I
B
R
S
Adh
pPA-2

a transposase gene which is transcribed from the hsp70 promoter and only one P element terminal repeat. It therefore provides transposase activity but does not itself integrate into the *Drosophila* genome. Of 578 injected embryos 41 developed into fertile adults. Six of these survivors produced Adh^{+} progeny when they were mated with Adh^{fn23} *pr cn* individuals. In all cases this ethanol resistance phenotype was transmitted faithfully to subsequent generations.

Before they were selected on ethanol, the wings of the F1 progeny were examined. In nine of the 41 lines a few of the F1 individuals had scalloped wings. In most cases a small amount of tissue was missing from only one wing, but in several individuals the scalloping was quite severe and affected both of the wings. We noticed that in almost all cases the distribution of lost tissue was similar to that of weak *Bx* mutant alleles.

To examine this further we mated these F1 individuals with Adh^{fn23} *pr cn* flies and then, after several days, selected them on ethanol. None of the *Bx*-like individuals that we selected survived on ethanol. Therefore, it seems unlikely that they express the Adh^{+} gene. The wing phenotype of these flies appears not to be transmissible since all of the F2 progeny had normal wings. When F2 brothers and sisters were mated, no scalloped wing progeny were observed. We conclude that the wing scalloping that we observed in the F1 progeny of injected individuals does not result from the stable integration of the P element from pNh10.4A into the *Drosophila* genome. Furthermore, the six stable ethanol-resistant transformant lines that we isolated do not appear to be associated with any excess of *Bx-hdp-a* gene activity, since no wing defect is evident in these lines.

By observing the segregation of the inserted Adh^{+} gene with respect to dominant markers on balancer chromosomes, we mapped the insertions of each of the six transformants to one of the two major *Drosophila* autosomes (Table 1). In each line only one chromosome was associated with resistance. Using balancer

TABLE 1

ANALYSIS OF TRANSFORMED LINES

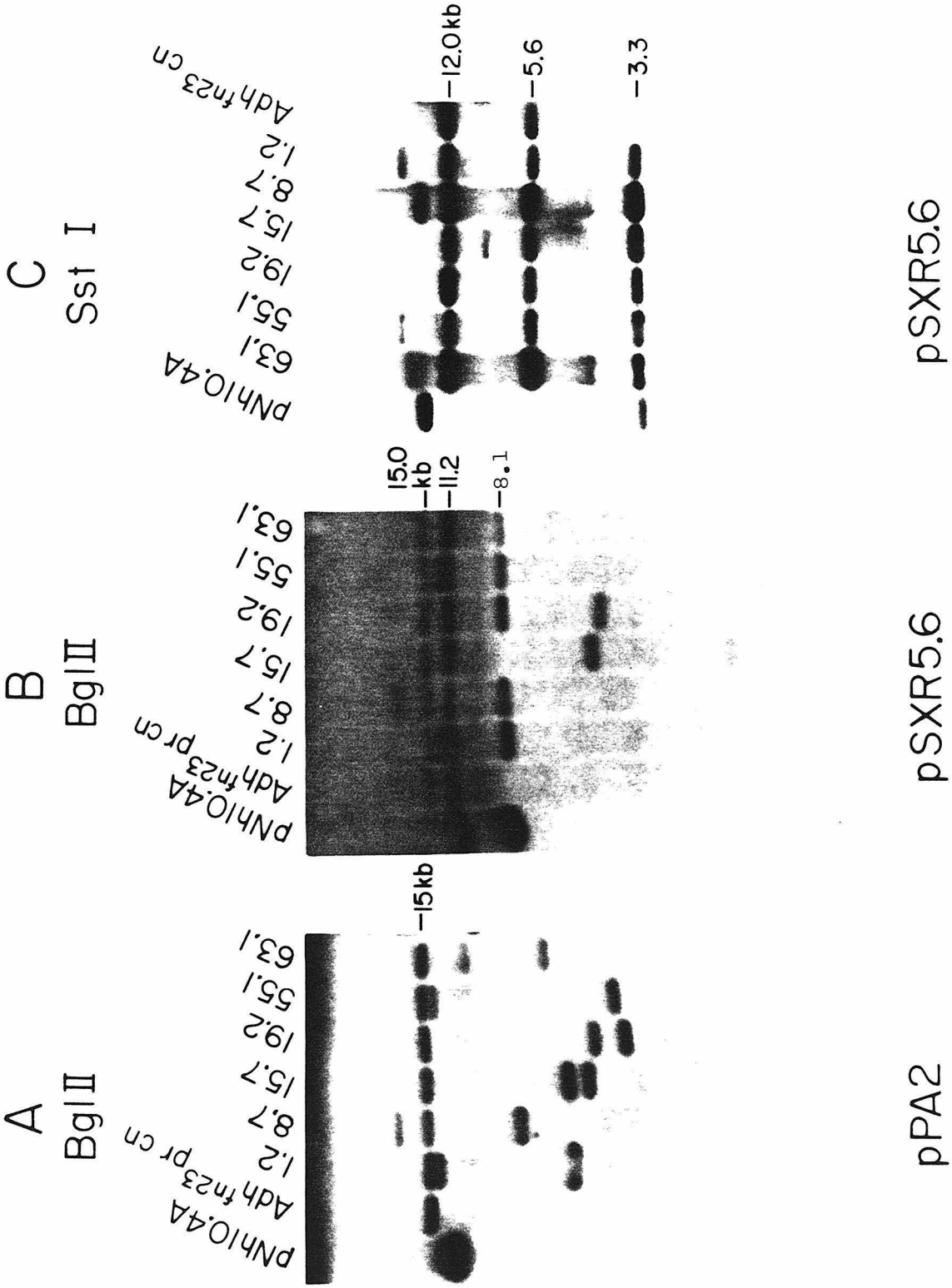
line	chromosome	hdp-a^{D30r} complement.	fraction of 3 dose females with wing scalloping	
			Adh^{+}	Adh^{-}
1.2	2nd	none	30% (30/100)	31% (28/89)
8.7	3rd	none	35% (34/98)	32% (28/88)
15.7	3rd	none	not done	not done
19.2	2nd	none	26% (16/61)	30% (35/116)
55.1	2nd	none	27% (22/82)	29% (49/169)
63.1	3rd	none	29% (22/75)	32% (24/74)
Adh^{fn23}prcn		none	--------	37% (59/161)
Double homozygote stocks				
1.2,8.7	2nd,3rd	none	not done	not done
1.2,15.7	2nd,3rd	none	not done	not done
19.2,8.7	2nd,3rd	none	not done	not done
55.1,8.7	2nd,3rd	none	not done	not done

chromosomes we made stocks that are homozygous for each of the chromosomes carrying ethanol resistance.

Analysis of genomic DNA from transformant lines. To determine whether or not each of the ethanol resistant lines has acquired an intact copy of the hybrid P element from pNh10.4A, we isolated genomic DNA from each of these lines and analyzed it by blot hybridization. ^{32}P-labeled pPA-2 DNA was hybridized with a blot of Bgl II digested genomic DNA from each of the transformed lines as well as *Adh*^{fn23} *pr cn* (Figure 2A). A 15 kb band is evident in all of the lanes containing *Drosophila* genomic DNA. This band originates from the defective *Adh* gene which is present in the *Adh*^{fn23} *pr cn* strain (the *Adh*^{fn23} mutation results from a 34 bp deletion in the protein-coding region; C. Benyajati, et al., 1983). Since each of the transformant lines carries the *Adh*^{fn23} allele, they contain this fragment as well. From the restriction map of pNh10.4A (Figure 1), we would expect that pPA2 would also hybridize with a Bgl II fragment from each end of the inserted element. Since these end fragments would extend into adjacent sequences at the insertion sites, we would expect each insertion to give rise to a unique pair of bands. Figure 2A shows that two fragments are present in each of the transformant DNAs that are not present in *Adh*^{fn23} *pr cn* DNA. This indicates that each of these lines has acquired one P element in its genome. Each of these insertions occurs at a different position in the genome because each pair of transformant specific bands is of different sizes.

From the map of pNh10.4A in figure 1 it can be seen that there is a single internal Bgl II fragment in the hybrid P element which is about 8.1 kb in length. This fragment extends from the left end of the inserted 10.4 kb fragment (-11.6) to a site at about coordinate -3.5. Because one of the Bgl II sites that define this fragment is located within the pPA2 polylinker, this 8.1 kb fragment does not normally exist in the *Drosophila* genome (see *Adh*^{fn23} *pr cn* digest in Figure 2B).

Figure 2. Blot hybridization analysis of genomic DNA from transformant lines. Two micrograms of genomic DNA from the indicated transformant lines digested with either Bgl II (A,B) or Sst I (C). Blots were hybridized with either pPA2 or pSXR5.6A (-6.9 to -1.0) nick-translated DNA as indicated below the blot. As a size marker for the 8.1 kb band in (B) and the 3.3 kb band in (C), 50 pg of pNh10.4A were digested with the same enzymes and blotted together with the genomic DNAs.


A
Bgl II
pNh10.4A
Adh fn23 pr cn
1.2
8.7
15.7
19.2
55.1
63.1
15kb
B
Bgl II
pNh10.4A
Adh fn23 pr cn
1.2
8.7
15.7
19.2
55.1
63.1
15.0
kb
11.2
8.1
C
Sst I
pNh10.4A
63.1
55.1
19.2
15.7
8.7
1.2
Adh fn23 cn
12.0kb
5.6
3.3
pPA2
pSXR5.6
pSXR5.6
pSXR5.6

A blot similar to the one used above was hybridized with ^{32}P-labeled pSXR5.6 DNA. This probe contains sequences from the -6.9 to -1.0 region and therefore hybridizes with the 8.1 kb Bgl II fragment. In Figure 2B it can be seen that the 8.1 kb band is present in five of the transformed lines but is missing from the 15.7 line. This indicates that the hybrid P element inserted in this line has a defect in the -11.6 to -3.5 region.

Figure 7C shows a similar analysis of the 3.3 kb Sst I fragment which covers the remainder of the inserted 10.4 kb fragment (-3.4 to -0.8) as well as part of one of the P element termini. In all six lines this fragment is of the expected size, indicating that it has not suffered any gross alterations. We conclude that five of our six transformed lines appear to have acquired an intact 10.4 kb Nhe I fragment.

Complementation analysis of transformed lines. To determine whether the 10.4 kb Nhe I fragment encodes the *hdp-a*$^{+}$ gene we crossed *hdp-a*^{D30r};*Adh*^{fn23} *pr cn* females with males from each of the six homozygous transformant stocks (*hdp-a*^{D30r} is a recessive loss-of-function mutation which results from a 1400 bp deletion). The male progeny from this cross inherit the *hdp-a*^{D30r} mutation on their X-chromosome and one copy of the inserted P element on an autosome. All six crosses yielded only males with the *hdp-a* mutant wing position phenotype. This indicates that a single dose of the inserted 10.4 kb fragment is not sufficient to provide the level of *hdp-a*$^{+}$ activity required to complement *hdp-a*^{D30r}.

One possible explanation for this result is that the *hdp-a*$^{+}$ gene, though included within the 10.4 kb fragment, is expressed at low levels in the transformed flies because of position effects. For this reason we have constructed strains that are doubly homozygous for each of two different P element insertions (see Table 1). This allowed us to easily assess the ability of two doses of the 10.4 kb fragment to complement *hdp-a*^{D30r}. We did this by crossing males from each of

several double homozygous strains with *hdp-a*^{D30r};*Adh*^{fn23} *pr cn* females. Although the *hdp-a*^{D30r} male progeny from this cross must carry two copies of the inserted 10.4 kb fragment, no complementation of *hdp-a*^{D30r} was observed (Table 1).

Analysis of the wing scalloping activity of the 10.4 kb fragment. Green (1953a,b) showed that females that carry four doses of the wild-type *Bx-hdp-a* region exhibit a weak wing scalloping phenotype, whereas females carrying three doses of this region do not. We were interested in determining whether the 10.4 kb fragment could contribute to this excess of activity phenotype in the same way that doses of the wild type *Bx-hdp-a* region do. We had hoped to do this by simply examining the wing phenotype of females that carry a copy of the reintroduced 10.4 kb fragment as well as three doses of the wild-type *Bx-hdp* region. Unfortunately, we found that any strain that is derived from *Adh*^{fn23} *pr cn* has a considerable enhancing effect on the wing scalloping phenotype of individuals with three doses of *Bx-hdp-a*, regardless of whether or not the 10.4 kb fragment is present. Up to 50% of the three dose-bearing females resulting from a cross to *Adh*^{fn23} *pr cn* may have wing scalloping (as compared to 1% for crosses involving the Canton-S or π_2 strains).

For this reason we crossed *Adh*^{fn23} *pr cn* females, which are heterozygous for the inserted 10.4 kb fragment, to *B Bx*^{r} *car*;*Adh*^{fn23} *pr cn* males (*Bx*^{r} males carry a tandem duplication of the wild-type *Bx-hdp-a* region on the X-chromosome; Green, 1953a). All of the female progeny from this cross have three copies of the wild-type *Bx-hdp-a* region, but only half of them inherit an autosome with the inserted 10.4 kb fragment. If the 10.4 kb fragment contributes to the wing scalloping phenotype, we would expect that the frequency of scalloped individuals would be higher among the ethanol resistant (P element-bearing) females than among ethanol sensitive females. It is clear from the results in

Table 1 that this is not the case. In each of the five lines tested the presence of the 10.4 kb fragment had no significant effect on the frequency of wing scalloping. These results suggest that the 10.4 kb fragment does not induce wing scalloping in females with three doses of the wild-type *Bx-hdp-a* region even when these individuals are close to the threshold at which scalloping occurs.

These results are supported by the observation that Bx^{+} $hdp\text{-}a^{+}$ individuals with as many as four extra copies of the 10.4 kb fragment do not exhibit a wing scalloping phenotype.

DISCUSSION

Our previous molecular studies have identified a 4 kb transcript that is closely associated with the *Bx* and *hdp-a* loci. This transcript has several characteristics which suggest that it may be encoded by the $hdp\text{-}a^{+}$ gene: It spans a region that is necessary for the normal function of the $hdp\text{-}a^{+}$ gene. It is structurally altered by the $hdp\text{-}a^{D30r}$ mutation. At least five different *Bx* mutations affect its structure. And finally, it is expressed during developmental stages in which the $Bx\text{-}hdp\text{-}a^{+}$ product is thought to be active.

By using sensitive single stranded probes in RNA blot hybridization experiments, we found a 5.9 kb region to which this transcript is homologous. No other sequences within a 34 kb region surrounding the *Bx* and *hdp-a* loci hybridized with the 4 kb RNA. Despite these facts, we find that a 10.4 kb fragment, which contains this 5.9 kb region, fails to complement a *hdp-a* loss-of-function mutation and also fails to enhance the wing scalloping phenotype of flies with excess *Bx-hdp-a* gene activity.

We think that the most likely explanation for these results is that the 10.4 kb fragment does not contain all sequences required for the expression of the 4 kb RNA transcript. We have no evidence that the 4 kb transcript is confined to the

34 kb region that we examined by hybridization. Even if one assumes that this transcript encodes the *hdp-a*$^+$ function, the situation is not improved, because a left (centromere-distal) limit on the *hdp-a* functional domain has not been defined in molecular terms. Therefore, it is possible that the 4 kb transcript is encoded by the hdp-a$^+$ gene and includes an exon which is located outside of the 34 kb region that we examined by blot hybridization.

Another consideration is the practical sensitivity limit of blot hybridization experiments. A very small exon, like those found in the *Ubx* region of the *bithorax* complex (Beachy et al., 1985) would probably be missed by our hybridization experiments. Such an exon could exist at a position which is outside of the 10.4 kb Nhe I fragment that we used. This problem is particularly worrisome when the transcripts in question are present only at very low levels as the 4 kb transcript is (10^{-5} to 10^{-6}; Chapter 2).

If the region encoding the 4 kb transcript is entirely included within the 10.4 kb fragment, there are several possible explanations for its apparent lack of function. It could be that this transcript is not expressed when inserted at novel positions in the genome. Such position effects have been observed in germ line transformation experiments with other *Drosophila* genes (Hazelrigg et al., 1984; Spradling and Rubin, 1983; Zehring et al., 1984; Goldberg et al., 1983). However, these effects are usually not so strong as to totally preclude the function of the inserted genes. In the case of the 10.4 kb fragment from the *Bx-hdp-a* region, each of five intact insertions failed to express any identifiable activity.

A less likely possibility is that the 4 kb transcript is only coincidentally affected by *hdp-a* and *Bx* mutations and that the true *hdp-a*$^+$ gene is affected in some subtle way or is located outside the region that we examined.

We do not believe that our results can be ascribed to an alteration in the 10.4 kb fragment which was acquired during *in vitro* manipulations or subsequent

growth in bacteria, since two transformed lines which we made with a similar plasmid (a fragment from the region between -11.2 and -1.6 in pPA2) also fail to complement *hdp-a*^{D30r}.

Whatever the explanation, we feel it would be enlightening to examine RNA produced by these transformant lines. If the 4 kb transcript is present in these lines then it seems likely that some other transcript encodes the *hdp-a*$^{+}$ function.

ACKNOWLEDGMENTS

We thank Carol Mayeda for helpful instructions on performing embryo injections. We are grateful to Elliot Meyerowitz for allowing us to use his injection facilities. Madeline Crosby, Mark Garfinkel and Beverley Bond Matthews provided fly stocks and plasmid DNAs used in this work. This work was supported by a National Research Service Award (GM-0766) and an NIH grant (GM-10991).

LITERATURE CITED

Beachy, P.A., S.L. Helfand and D.S. Hogness. 1985. Segmental distribution of bithorax complex proteins during *Drosophila* development. Nature 313:545-551.

Benyajati, C., A.R. Place and W. Sofer. 1983. Formaldehyde mutagenesis in *Drosophila*: molecular analysis of ADH-negative mutants. Mut. Res. 111:1-10.

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 34:59-75.

Green, M.M. 1953a. The *Beadex* locus in *Drosophila melanogaster*: On the nature of the mutants Bx^{r} and Bx^{1}. Genetics 38:91-105.

Green, M.M. 1953b. The *Beadex* locus in *Drosophila melanogaster*: Genetic analysis of the mutant Bx^{r49k}. Zeitschrift fur indukt. Abstammungs- und Vererbungslehre. 85:435-449.

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.

Lifshytz, E. and M.M. Green. 1979. Genetic identification of dominant overproducing mutations: The *Beadex* gene. Mol. Gen. Genet. 171:153-159.

Spradling, A.C. and G.M. Rubin. 1982. Transposition of cloned P elements into 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 on *Drosophila* larvae. EMBO 4:167-171.

Vigue, C. and W. Sofer. 1976. Chemical selection of mutants that affect ADH activity in *Drosophila*. III. Effects of ethanol. Biochem. Genet. 14,127-135.

Zehring, W.A., D.A. Wheeler, P. Reddy, R.J. Konopka, C.P. Kyriacou, M. Rosbash and J.C. Hall. 1984. P element transformation with *period* locus DNA restores rhythmicity to mutant, arrhythmic *Drosophila melanogaster*. Cell 39:369-376.

APPENDIX 1

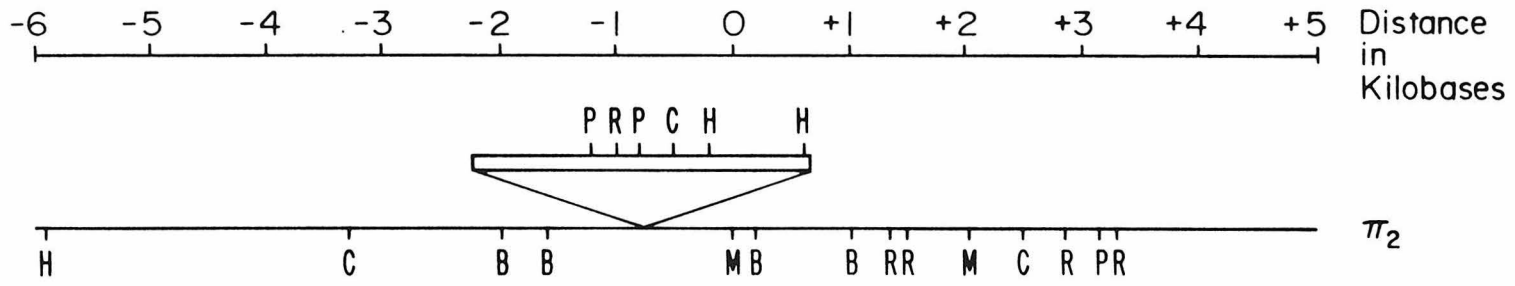
Characterization of More Hybrid Dysgenesis Induced Deletion Mutations at the *Beadex* Locus

William W. Mattox and Norman Davidson

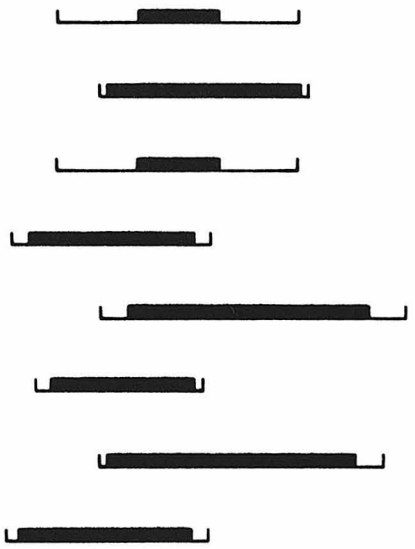
In addition to the two hybrid dysgenesis-induced deletion mutations that we analyzed in Chapter 1, we have characterized the DNA structure of six other dysgenesis-induced *Beadex* (*Bx*) mutations (Bx^{MH6}, Bx^{MH25}, Bx^{MH26}, Bx^{MH32}, Bx^{MH48} and $Bx^{3.02.1}$) by genomic blot hybridizations. Each of these mutations was induced on the π_2 X-chromosome (Berg, et al., 1980), which contains a P element at -0.8. As was the case for Bx^{9} and Bx^{15}, each of these six mutations is associated with the loss of this P element. This is evidenced by the loss of restriction sites that are contained within the P element sequences and the observation that the 17C segment of mutant chromosomes does not hybridize with a P element probe (C. Preston and W. Engels, personal communication). In five of the six mutant DNAs we found that restriction sites flanking the P element insertion site were also lost. In each of these five cases the added sizes of the lost fragments exceeded the size of a single novel fragment which replaced them. By examining the fragments produced when the DNA from each of these five *Bx* strains were digested with Bam HI, Xba I, Sst I, Pst I, or Ava I (see Figure 1), we are able to conclude that each mutant strain has lost a segment of DNA which flanks the P element insertion site and includes the P element as well. The deleted regions are shown in the figure. In each case at least some of the sequences located between -1.6 and -0.8 are lost. This is consistent with our conclusion in Chapter 1 that sequences essential to the normal function of the *Bx* locus lie in this region.

We were surprised to find that the mutants Bx^{MH6} and Bx^{MH26} had identical restriction maps even though they are the result of independent mutational events. In all of the blots that we have performed with DNA from these two mutants, we have not been able to discern a difference in the size of any of their restriction fragments. This may indicate that there are preferred sites for the termination of these imprecise excision events. However, it is also

Figure 1. *Bx* deletion mutants. The restriction map is that of π_2. The inserted P element is shown at -0.8. The deleted region in each of the mutant strains is indicated below the map. The entire P element is deleted in each of the mutants examined. The dark bar indicates regions that are certainly deleted. The lines indicate regions that may or may not be deleted. The sizes of the deletions, not including the lost P element sequences, are indicated at the right.



B = Bam HI
C = Sac I
H = Hind III
M = Sma I
R = Eco RI
P = Pvu I



Bx^{MH6}	Δ 1.2 kb
Bx^{MH25}	Δ 1.5 kb
Bx^{MH26}	Δ 1.2 kb
Bx^{MH32}	Δ 1.5 kb
Bx^{MH48}	Δ 2.2kb
Bx^{9}	Δ 1.3kb
Bx^{15}	Δ 2.0kb
hdp-a^{D30r}	Δ 1.5kb

possible that the resolution of our analysis has not been great enough to discern a small difference which exists between these two deletions.

The sixth *Bx* mutant that we analyzed, $Bx^{3.02.1}$, was associated with a complex rearrangement which in some cases produced as many as three novel bands on our genomic blots. We did not analyze this mutant in sufficient detail to determine the exact nature of the rearrangement, but it seems to be confined to the region between -3.4 and +2.2.

We wish to note here that we also included $hdp\text{-}a^{D30r}$ in this set of blot hybridizations. Careful measurements of the fragments produced by DNA from this strain in digests using Ava I and Pst I indicate that this deletion is close to 1.5 kb in length rather than 1.4 kb as reported in Chapter 1.

LITERATURE CITED

Berg, R., W.R. Engels, and R.A. Kreber. 1980. Site-specific X-chromosome rearrangements from hybrid dysgenesis in *Drosophila melanogaster*. Science 210:427-429.

APPENDIX 2

Drosophila melanogaster Has Only One Myosin Alkali Light-Chain Gene which Encodes a Protein with Considerable Amino Acid Sequence Homology to Chicken Myosin Alkali Light-Chain

SCOTT FALKENTHAL, VANN P. PARKER, WILLIAM W. MATTOX AND NORMAN DAVIDSON

Molecular and Cellular Biology

4:956-965 (1984)

MOLECULAR AND CELLULAR BIOLOGY, May 1984, p. 956–965
0270-7306/84/050956-10$02.00/0

Vol. 4, No. 5

Drosophila melanogaster Has Only One Myosin Alkali Light-Chain Gene Which Encodes a Protein with Considerable Amino Acid Sequence Homology to Chicken Myosin Alkali Light Chains

SCOTT FALKENTHAL,[1†*] VANN P. PARKER,[2‡] WILLIAM W. MATTOX,[2] AND NORMAN DAVIDSON[1]

Department of Chemistry[1] and Division of Biology,[2] California Institute of Technology, Pasadena, California 91125

Received 9 January 1984/Accepted 27 February 1984

A chimeric lambda DNA molecule containing the myosin alkali light-chain gene of *Drosophila melanogaster* was isolated. The encoded amino acid sequence was determined from the nucleic acid sequence of a cDNA homologous to the genomic clone. The identity of the encoded protein was established by two criteria: (i) sequence homology with the chicken alkali light-chain proteins and (ii) comparison of the two-dimensional gel electrophoretic pattern of the peptides synthesized by in vitro translation of hybrid-selected RNA to that of myosin alkali light-chain peptides extracted from *Drosophila* myofibrils. There is only one myosin alkali light-chain gene in *D. melanogaster*; its chromosomal location is region 98B. This gene is abundantly expressed during the development of larval as well as adult muscles. The *Drosophila* protein appears to contain one putative divalent cation-binding domain (an EF hand) as compared with the three EF hands present in chicken alkali light chains.

Myosin light chains are proteins which occur abundantly and in a defined stoichiometry in myofibrils. They are members of an evolutionarily related group of calcium-binding proteins known as the troponin C superfamily, which includes calmodulin, troponin C, and the myosin alkali and 5,5'-dithiobis(2-nitrobenzoic acid) (DTNB) light chains. The primary amino acid sequence has been determined for at least one vertebrate example of each of these polypeptides (2). The principal sequence homology between these proteins resides in the putative Ca^{2+}-binding domains, which are known as EF hands (14). The roles of all of these proteins, except for the myosin alkali light chain, in muscle function have been determined (10, 13, 29; R. A. Murphy, M. O. Askoy, P. F. Dillon, W. T. Gerthoffer, and K. E. Kanim, Fed. Proc. **42**:51–57, 1983).

The skeletal muscle myosin alkali light chains are so named because of the high pH required to dissociate them from the myosin heavy chain (39). For vertebrate muscles, they are sometimes called MLC-1 and MLC-3. The two skeletal muscle alkali light chains of mammals and chickens, which have molecular weights of about 21,000 (MLC-1) and 17,000 (MLC-3), are virtually identical in sequence over their C-terminal 141 residues, but diverge in sequence at the amino terminus. MLC-1, depending upon the tissue from which it is isolated, has an additional alanine-proline- or alanine-lysine-rich sequence of 40 amino acids at its amino terminus. There is evidence that in rats the two proteins are encoded by a single gene (L. Garfinkel, R. Gubits, B. Nadal-Ginard, and N. Davidson, manuscript in preparation). At one time, these peptides were thought to be essential for the actin-activated adenosine triphosphatase activity of myosin (16, 32, 38), but recent in vitro studies have suggested that such is not the case (36). Thus, the function of the alkali light chains is unknown.

Drosophila melanogaster is an advantageous organism for further studies of structure and function of gene products such as the myosin light chains because one may use both molecular and genetic approaches. Here, we describe our initial steps in isolating and identifying the myosin alkali light-chain gene of *D. melanogaster* and in determining from the nucleic acid sequence the primary structure of the protein. We also report on the reiteration frequency and developmental expression of this gene.

* Corresponding author.
† Present address: Department of Genetics, 927 Biological Sciences Building, The Ohio State University, Columbus, OH 43210.
‡ Present address: Amgen, Newberry Park, CA 91320.

MATERIALS AND METHODS

Isolation of RNA. Total cellular RNA was prepared by homogenizing developmental-stage whole animals in 4 M guanidine thiocyanate–1 M 2-mercaptoethanol–0.05 M sodium acetate–0.001 M EDTA (pH 6.0) and banding in cesium chloride (11, 35). Polyadenylate-containing [poly(A)$^+$] RNA was selected by oligodeoxythymidylate-cellulose (type T3; Collaborative Research) chromatography as described by Anderson and Lengyel (1).

Isolation and purification of DNA. Charon phage DNA was isolated as described by Yen and Davidson (41) with the modifications described by Snyder et al. (33). Plasmid DNA was isolated as described by Fyrberg et al. (11). High-molecular-weight pupal DNA was prepared from 40- to 60-h pupae as follows. Two grams of pupae frozen in liquid nitrogen were ground to a fine powder in a mortar at −70°C. The frozen powder was homogenized by 10 to 15 strokes with a B pestle Dounce homogenizer in 30 ml of homogenization buffer (50 mM Tris, pH 7.2, 25 mM KCl, 5 mM $MgCl_2$, 350 mM sucrose, 0.15 mM spermine, 0.15 mM spermidine) and then filtered through Nitex cloth to remove cuticular debris. The nuclei were collected by centrifugation in a Sorvall HB-4 rotor at 4°C for 15 min at 3,000 rpm. The pelleted nuclei were washed with 30 ml of homogenization buffer and centrifuged to reduce the mitochondrial contamination of the nuclear pellet. The washed and pelleted nuclei were then suspended in 1.0 ml of nuclear suspension buffer (60 mM NaCl, 10 mM Tris, pH 7.2, 10 mM EDTA, 0.15 mM spermine, 0.15 mM spermidine). The nuclei were then lysed by the addition of 2.0 ml of lysis buffer (200 mM Tris, pH 8.5, 30 mM EDTA, 2% (wt/vol) Sarkosyl). Proteinase K was added to a concentration of 50 μg/ml, and the solution was gently mixed. After 2 to 4 h at 42°C, CsCl and ethidium

bromide were added to concentrations of 0.925 g/ml and 500 μg/ml, respectively. The solution was spun in a table-top clinical centrifuge at top speed for 20 min to float the denatured protein. The cleared solution was centrifuged for 18 h at 53,000 rpm (20°C) in a Beckman VTi65 rotor. The banded genomic DNA was visualized by UV irradiation, and the band was collected by side puncture. The ethidium bromide was removed by butanol extraction. The sample was diluted to 0.5 ml and dialyzed twice against 5,000 volumes of 10 mM Tris (pH 8.0)–1 mM EDTA. The salt concentration was adjusted to 0.15 M with sodium acetate, and the DNA was precipitated by the addition of 2.0 volumes of absolute ethanol. After 1 h at −20°C, the DNA was collected by centrifugation, rinsed with 70% ethanol, and air dried at 4°C for 24 h. The DNA was resuspended by overlaying the pellet with 10 mM Tris (pH 8.0)–1 mM EDTA and allowing it to hydrate for 48 to 72 h at 4°C.

Screening libraries. A *Drosophila* genomic DNA library prepared by J. Lauer was used. It consisted of randomly sheared Canton S embryonic DNA of 12 to 20 kilobases (kb) inserted into Charon 4 via synthetic *Eco*RI linkers (17). The library was screened by the high-density plaque hybridization technique of Benton and Davis (4).

Subcloning fragments of genomic clones. DNA fragments of chimeric lambda clone λdmpT102 were obtained by digestion with *Hin*dIII and *Eco*RI. These fragments were ligated into plasmid pBR322 as described previously (41).

DNA labeling and hybridizations. ^{32}P-labeled pupal cDNA probes were prepared from pupal poly(A)$^+$ RNA by using oligo(dT) primers (P-L Biochemicals) by the method of Mullins et al. (24). Preparation of nick-translated probes and hybridization of ^{32}P-labeled DNA probes to filter-bound DNA were performed essentially as described by Mullins et al. (24). Nick-translated probes were hybridized at a concentration of 10^5 dpm/ml (10^8 dpm/μg), whereas the labeled cDNA probes were hybridized at a concentration of 1×10^6 to 2×10^6 dpm/ml (10^7 dpm/μg).

Modified Okayama and Berg technique for cDNA library preparation. The goal of this approach was to make a cDNA library with efficient utilization of poly(A)$^+$ RNA and with long inserts which contained the entire 3′ untranslated region, the protein-coding region, and most of the 5′ untranslated region, following the general spirit of the method of Okayama and Berg (28). Their procedure can be readily adapted and indeed simplified by using a modern cloning vector with a suitable polylinker. The first version of our procedure is depicted in Fig. 1. In this proposed procedure, intermediate 4 is treated with dGTP and terminal transferase to give intermediate 5 and to provide a G tail for oligodeoxycytidylate priming of second-strand synthesis from intermediate 6. In this procedure, we used Mn^{2+} ion as a catalyst based on the recommendation of Deng and Wu (9). If successful, this procedure would have regenerated the *Eco*RI site upon final ligation. In fact, none of the resulting clones examined at the sequence level had a GC oligonucleotide segment or the *Eco*RI site. However, this procedure has worked well in the hands of other investigators (F. K. Lin, personal communication). We presume that our lack of success in this step was due to a failure of the terminal transferase tailing reaction. We believe that an alternative and equally effective procedure would be to carry out the replacement synthesis directly on intermediate 4. This procedure, which is analogous to that used by Charles Rice (personal communication), relies on RNase H action to generate the primers for second-strand synthesis. Several of our clones for other genes have been shown to include up to ca. 10 nucleotides from the 5′ end of the mRNA. Thus a 5′-terminal primer is hardly necessary. This revised procedure would result in converting the *Bam*HI site into a blunt end. The final ligation would thus preserve the *Eco*RI site.

It should be noted that with the pUC8 vector there is an upstream *lac* promoter and translation start system. Thus, the resulting cDNA clones are useful for expression of the encoded proteins in *E. coli*.

pUC8 DNA (10 μg) was digested with 50 U of *Pst*I in 40 μl containing 100 mM Tris-hydrochloride (pH 7.5), 50 mM KCl, 10 mM $MgCl_2$, and 100 μg of bovine serum albumin (Pentex) per ml. After 60 min the solution was heated to 65°C for 15 min and then quickly cooled. Poly $(T)_{50-80}$ tails were added to the *Pst*I-digested DNA with terminal transferase. The reaction solution (200 μl) contained the following: 10 μg of *Pst*I-cut pUC8 DNA, 0.25 mM unlabeled TTP, 50 μCi of [^{3}H]TTP (77.8 Ci/mmol), 0.2 M cacodylate, 0.05 M Tris (pH 7.0), 1 mM $CoCl_2$. This was incubated for 10 min at 37°C, after which dithiothreitol and terminal transferase were added to concentrations of 0.1 mM and 110 U/ml, respectively. After 20 min the nucleic acid was precipitated by the addition of 2 volumes of absolute ethanol. The precipitated nucleic acid was collected by centrifugation, washed once with 90% ethanol, and dried under vacuum.

At this point, the vector has polythymidylate tails at each end of the molecule. To act as a good primer there must be only one tail per molecule. One tail is eliminated by digesting the molecule with *Bam*HI, whose recognition site is 12 nucleotides 5′ to the T-tailed *Pst*I site. The small resulting oligonucleotide was separated from the tailed vector by gel filtration on Sephadex G-150.

The T-tailed DNA was resuspended in 200 μl of 100 mM Tris (pH 7.5)–50 mM KCl–10 mM $MgCl_2$–100 μg of BSA per ml. *Bam*HI (10 U) was added, and the solution was incubated at 37°C. After 1 h, EDTA was added to a concentration of 50 mM, and the solution was layered onto a 5.0-ml Sephadex G-150 column (0.5 by 28 cm) equilibrated with 50 mM NaCl–10 mM Tris (pH 8.0)–10 mM EDTA. Fractions containing the excluded peak were pooled, and the nucleic acid was recovered by ethanol precipitation and centrifugation.

First-strand cDNA synthesis was performed in a volume of 100 μl containing 2 μg of poly(T)-tailed pUC8 DNA, 10 μg of pupal poly(A)$^+$ RNA, 50 mM Tris (pH 8.3), 50 mM KCl, 8 mM $MgCl_2$, 4 mM sodium PP_i, 2 mM DTT, 2 mM TTP, 2 mM dCTP, 2 mM dATP, 2 mM dGTP, and 5 μCi of [^{32}P]dCTP (410 Ci/mmol). Reverse transcriptase (54 U) was added, and the solution was incubated at 46°C for 45 min. After phenol extraction, the aqueous phase was layered onto a 5.0-ml Sephadex G-150 column (0.5 by 28 cm) equilibrated in 50 mM NaCl–10 mM Tris (pH 8.0)–10 mM EDTA. The excluded peak fractions were pooled, and the nucleic acid was recovered by ethanol precipitation and centrifugation. The second-strand replacement synthesis was done with RNase H and DNA polymerase 1 as described by Okayama and Berg (28). The resulting double-stranded DNA molecules were circularized by blunt end ligation as follows. The DNA was ethanol precipitated after second-strand synthesis, collected by centrifugation, and resuspended in 0.5 ml of a solution containing 50 mM Tris (pH 7.8), 10 mM $MgCl_2$, 20 mM DTT, 1 mM ATP, and 50 μg of BSA per ml. Incubation with 2,000 U of T4 DNA ligase proceeded for 12 h at 12°C. Bacterial cells (*Escherichia coli* MC1061) were transformed, and 5% of the transformation mix was plated onto L plates containing 50 μg of ampicillin per ml, whereas the remainder of the library was amplified for 3 h in liquid culture and then frozen in 50% glycerol at −40°C. Our library contained

200,000 independent recombinants from 2 μg of starting vector DNA.

Purification of *Drosophila* myosin. *Drosophila* indirect flight muscle myofibrils were purified from 24- to 48-h adults essentially as described by Mogami et al. (23). Myosin was prepared by high-salt extraction of isolated myofibrils as described by Whalen et al. (40). EDTA was present at a concentration of 10 mM during the isolation of myofibrils and extraction of actomyosin, which probably resulted in the dissociation of DTNB light-chain peptides from the myosin (V. P. Parker, S. Falkenthal, and N. Davidson, manuscript in preparation).

Electrophoresis of RNA in formaldehyde gels. Electrophoresis of RNA in formaldehyde gels and transfer to nitrocellulose paper was performed as described by Rozek and Davidson (31). Prehybridization, hybridization with 10% dextran sulfate, and the washing of filters after hybridization were done as described by Mullins et al. (25).

Positive selection and translation of RNA. RNA was selected by a procedure similar to that described by Ricciardi et al. (30). Bacteriophage DNA (10 μg) was denatured by heating and applied to nitrocellulose filters (0.3 by 0.3 cm) in 10× SSC (1× SSC is 0.15 M NaCl plus 0.015 sodium citrate). Up to 15 filters containing different cloned *Drosophila* DNA sequences were prehybridized in 200 μl of 70% formamide–0.4 M NaCl–0.1 M PIPES [piperazine-*N*,*N*′-bis(2-ethanesulfonic acid)] (pH 6.5) for 2 h at 50°C. Hybridizations were done for 4 h at 50°C in 100 μl of the same buffer containing 100 to 150 μg of pupal poly(A)$^+$ RNA (70 to 75 h postpupariation). After hybridization, the filters were washed en masse 10 times in 5 ml of 1× SSC–0.5% sodium dodecyl sulfate at 65°C, then three times in 5 ml of 0.01 M Tris (pH 7.8)–1 mM EDTA at room temperature. The filters were then placed in individual vials, and the hybridized RNA was eluted in boiling distilled water and recovered by ethanol precipitation. Translation was in a commercial (Bethesda

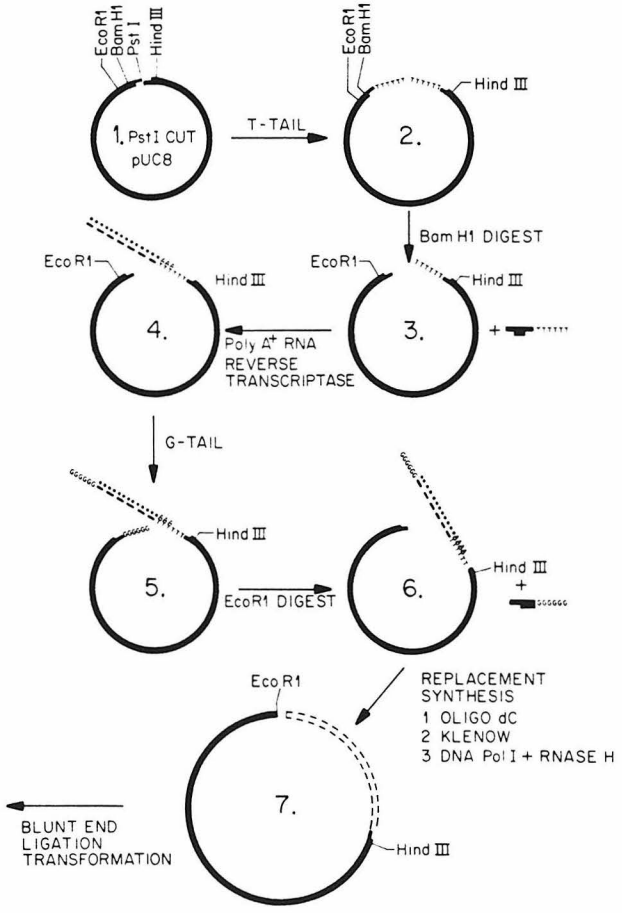


FIG. 1. Diagrammatic representation of the cDNA cloning strategy.

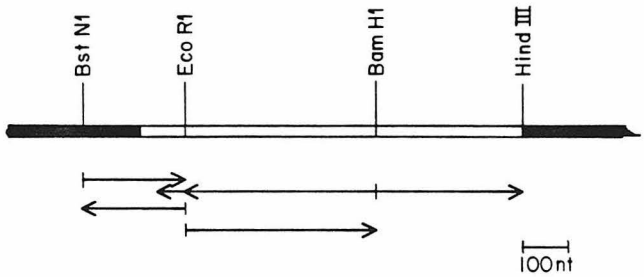


FIG. 2. Sequencing strategy of cDNA clone pcMLC-1.10. The DNA sequencing strategy is such that a vertical line (|) denotes ^{32}P labeling at the 3′ terminus. The arrows indicate the direction and extent of sequencing from the labeled terminus.

Research Laboratories) micrococcal nuclease-treated rabbit reticulocyte translation system. For each 30-μl translation, 50 μCi of 800-Ci/mmol [^{35}S]methionine (New England Nuclear) was used. Translation was terminated after 60 min by the addition of 50 ng each of RNase A and DNase I and subsequent incubation at 4°C for 30 min.

One-third of each translation assay was analyzed on a two-dimensional polyacrylamide gel as described by O'Farrell (27). The lysis buffer was modified to contain 0.1% sodium dodecyl sulfate (6). Molecular weights were determined by electrophoresing ^{14}C-labeled protein standards (Bethesda Research Laboratories) in an adjacent slot. Radioactively labeled proteins were detected by fluorography by the method of Laskey and Mills (15). Dried gels were exposed to preflashed Kodak XAR X-ray film.

DNA sequencing. The DNA sequence of a cDNA clone was determined by the method of Maxam and Gilbert (20) with modifications as described by Snyder et al. (34). The sequencing strategy used is shown in Fig. 2. Regions which were sequenced for one strand only were confirmed by a comparison with the DNA sequence of the chromosomal gene (S. Falkenthal, V. P. Parker, and N. Davidson, manuscript in preparation).

Amino acid homology comparisons. One homology comparison was carried out with a standard dot matrix program that asks for matches of five out of eight contiguous amino acids, but accepts conservative amino acid replacements (Val, Ileu, and Leu; Arg and Lys; Gln and Asn; Glu and Asp) as a match. In practice, this provides a rather stringent comparison of amino acid sequence homology (T. Hunkapiller, personal communication). The second two-dimensional matrix analysis, denoted the "best-fit" program, is useful for comparisons of sequences which are more distantly related. A match of the central amino acid is given a score of one, and matches at increasing distances from the center are given a reduced added score [M. Hunkapiller, S. Kent, M. Carruthers, W. Dreyer, J. Firca, C. Griffin, S. Horvath, T. Hunkapiller, P. Tempst, and L. Hood, Nature (London), in press].

RESULTS

Isolation of genes abundantly expressed in the indirect flight muscle. Our approach for isolating the myosin light-chain genes was based on the assumption that the concentration of their mRNAs would be high during the developmental stage when maximal synthesis of adult musculature occurs. Protein labeling studies showed that this occurred 70 to 75 h after puparium formation (data not shown). In vitro translation of RNA isolated from dissected thoraces at this developmental stage revealed that greater than 50% of the incorporation of [^{35}S]methionine was into myofibrillar proteins (data not shown). Accordingly, cDNA synthesized from pupal thoracic poly(A)$^+$ RNA was used to screen a Canton S random shear library of *Drosophila* genomic DNA in the vector Charon 4 (17). The 73 positive phages so selected were all plaque purified. These isolates were counterscreened with [^{32}P]cDNA homologous to early embryo RNA, a developmental time in which muscle-specific genes are not expressed, and with actin and myosin heavy-chain probes. The isolates which screened negatively with the above hybridization probes were rescreened with [^{32}P]cDNA synthesized from RNA extracted from the dorsal-lateral indirect flight muscle of late pupae (70 to 75 h). By restriction mapping, the resulting positive clones fell into 24 groups of nonoverlapping DNA inserts.

Initial characterization of these clones involved cytological localization of the DNA inserts by in situ hybridization to salivary gland chromosomes (Table 1). It may be seen that these inserts, representing genes abundantly expressed in the indirect flight muscle, do not show a general pattern of clustering on the *Drosophila* chromosomes. Tight clustering, as has been demonstrated for the cuticle genes on the second

TABLE 1. In situ localization of λdmpT recombinant clones

Clone	Band
31[a]	100B
49	28C
50[a]	97A
57[a]	99E
61[a]	17A
63	36B
73	88F
75	30B
85[a]	64C
101	102EF
102	98B
104	30EF
106	Repeated
115	72DE
116	53F
120	64F
121[a]	66F
123	94E

[a] Clone blots and in vitro translation of hybrid-selected RNA indicated that more than one gene is contained within the *Drosophila* insert (data not shown).

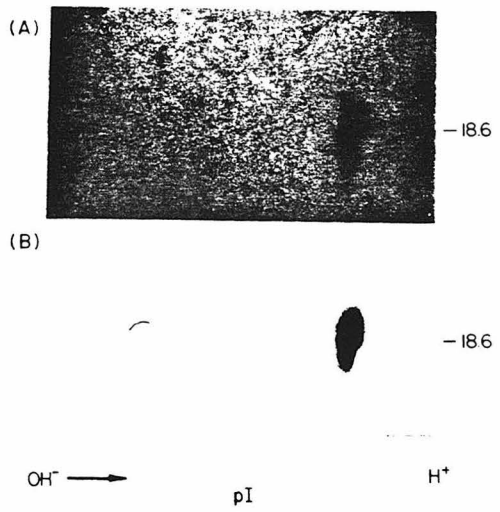


FIG. 3. Translation of myosin light-chain mRNA in a rabbit reticulocyte lysate. Myosin light-chain mRNA was selected by hybridization to filter bound λdmpT102 DNA as described in the text. This RNA was translated in an mRNA-dependent rabbit reticulocyte lysate containing [^{35}S]methionine, and the translation products were coelectrophoresed with 5 μg of *Drosophila* myosin purified from adult myofibrils on a 16% polyacrylamide gel. (A) Coomassie brilliant blue-stained gel. (B) Autoradiographic exposure (36 h) of the gel in A.

chromosome (33, 34) and the glue protein genes on the third chromosome (21), is not excluded by this analysis. Clone blots with labeled pupal cDNA as a hybridization probe and in vitro translation of hybrid-selected RNA revealed that 11 of the inserts contained only one gene which was expressed during pupal myogenesis, whereas the other inserts contained two or possibly three separate but closely linked transcription units (data not shown). Therefore, there is some tight clustering of genes abundantly expressed in the indirect flight muscle. Only one clone, λdmpT73, which probably contains a tropomyosin gene (3) hybridized in situ to the same region, 88F, where an actin gene expressed in the indirect flight muscle maps (12). Mutations which result in dominant flightless behavior have been found at regions 36B, where the myosin heavy-chain gene maps, and 88F (5, 23). None of our other clones mapped in these two regions.

Identification of the clone which contained the myosin light-chain sequence. The DNA insert which most probably encodes a myosin light-chain gene was identified by hybrid selection of RNA and in vitro translation (see above). The resulting ^{35}S-labeled polypeptides were compared to purified *Drosophila* myosin light-chain protein by two-dimensional gel electrophoretic analysis. The in vitro translation products of two of these clones (λdmpT75 and λdmpT102) had molecular weights of 17,000 to 20,000, the molecular weight range expected for the myosin light-chain protein. Whereas the polypeptide encoded by λdmpT75 had a more basic isoelectric point (data not shown), that encoded by λdmpT102 had the identical electrophoretic mobility of the extracted myosin light-chain protein (Fig. 3). The molecular weight of the latter protein(s) ranged from 18,000 to 19,000. Note that there is heterogeneity in both the molecular weight and the isoelectric point of the myosin light-chain isolated from adult muscle as well as for the in vitro-synthesized translation products.

Isolation and sequencing of a myosin light-chain cDNA clone. To achieve a more positive identification of the protein encoded by the insert of λdmpT102, a cDNA clone for which it codes was isolated and sequenced. Initially, the coding region of λdmpT102 was localized onto 3.2 kb of DNA by clone blots with [^{32}P]cDNA from total pupal poly(A)$^+$ RNA as a hybridization probe. This region contains two contiguous restriction fragments, a 1.8-kb *Hin*dIII-*Eco*RI fragment and a 1.35-kb *Eco*RI fragment. These two fragments were subcloned (see Fig. 4 for relevant restriction maps) and used to screen a cDNA library.

The cDNA library was prepared from late pupal poly(A)$^+$ RNA by a simplified procedure modeled on that of Okayama and Berg (28) to increase the probability of obtaining long inserts. Of 10,000 recombinants screened (5% of the total library), 10 clones containing sequences homologous to the insert of p102.6 were obtained. The clone, pcMLC-1.10, which contained the longest insert (880 nucleotides), was

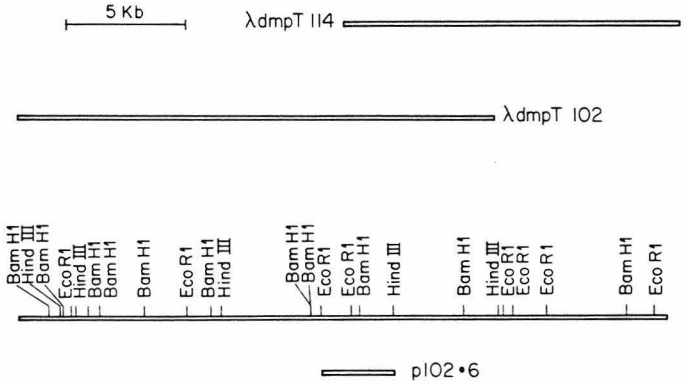


FIG. 4. Restriction endonuclease map of λdmpT102 and λdmpT114. A composite map of the two bacteriophage lambda clone inserts is shown. The subcloned coding region is shown below the map.

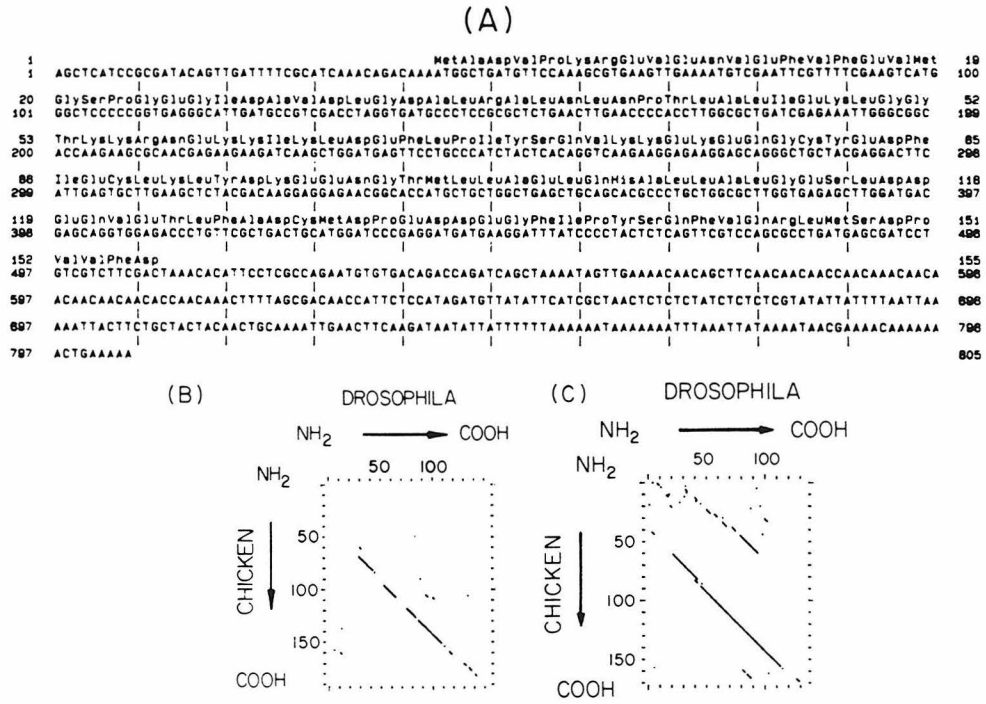


FIG. 5. Predicted protein coding sequence of the insert of pcMLC-1.10. (A) The amino acid sequence of the putative myosin light chain is depicted over the encoding DNA sequence of the 880-nucleotide insert of pcMLC-1.10. The DNA sequence was determined as described in the text. (B) Homology search between the predicted amino acid sequence of the *Drosophila* protein and the amino acid sequence of chicken skeletal muscle myosin light chain-1 protein. The comparison is such that every eight amino acids of the *Drosophila* protein are compared with the chicken protein (a positive score is marked if five out of eight amino acids match). A diagonal line indicates homology. (C) A best-fit comparison between the predicted amino acid sequence of the *Drosophila* protein and the chicken skeletal muscle myosin light chain-1 protein calculated by the method of Hunkapiller et al. (in press). The comparison is such that 10 amino acids of the Drosophila protein are compared with the chicken protein, allowing conservative amino acid replacements. The confidence limit for this fit was set at 99%.

sequenced (Fig. 5A). An AUG codon at nucleotide 44 begins a reading frame 456 nucleotides in length, which terminates with a TAA at nucleotide 509. Downstream from this termination codon are two additional in-frame termination signals (TAA and TGA at nucleotides 552 and 561, respectively).

The amino acid sequence of this reading frame was deduced and compared with the amino acid sequence of bovine brain calmodulin (37), rabbit troponin C (7), chicken skeletal muscle myosin light chain-1 (MLC-1) (19), and chicken skeletal muscle myosin light chain-2 (MLC-2) (18). These evolutionarily related proteins all possess the divalent metal binding (Ca^{2+}, Mg^{2+}) structure referred to as the EF hand (14).

By the stringent dot matrix amino acid homology comparison described above, only the chicken skeletal muscle myosin alkali light chains (MLC-1 and MLC-3) gave a significant positive score with the *Drosophila* protein (Fig. 5B). This fact and the in vitro translation study described above are the primary bases on which we identified the gene as that for the *D. melanogaster* alkali light chain.

In all regions where the dot matrix method showed homology between the *Drosophila* protein and the chicken MLC-1 protein, the latter is identical in sequence with the chicken MLC-3 protein (see Fig. 8 and below). The *Drosophila* protein has approximately the same molecular length as does the chicken MLC-3 protein, without the amino-terminal tail of the MLC-1 protein. A more detailed comparison of these sequences is presented later. We note here simply that overall the region of homology extends from amino acids 32 to 144 of the *Drosophila* protein with amino acids 72 to 167 and 31 to 143 of chicken MLC-1 and MLC-3, respectively (see Fig. 8). On this basis, we identified this gene as the *Drosophila* myosin alkali light chain (MLC-ALK).

The *Drosophila* MLC-ALK and chicken MLC-1 sequences were compared by the best-fit program which is designed to identify more distant sequence relationships (Fig. 5C). This comparison confirmed the major segment of homology noted above. In addition, it revealed that the sequence from amino acids 90 to 105 of *D. melanogaster* is related to the amino terminus proximal sequence from amino acids 56 to 71 of chicken MLC-1.

Developmental expression of the MLC-ALK gene. The MLC-ALK gene was selected by using a probe isolated from the pupal stage of development. The question arose as to whether the MLC-ALK gene was transcribed at other times in development, particularly at other times of muscle biosynthesis. To answer this question, total poly(A)$^+$ RNA was

isolated from animals at various developmental stages, electrophoresed on denaturing formaldehyde agarose gels, transferred to nitrocellulose paper, and hybridized with nick-translated MLC-ALK-coding DNA (p102.6). The results (Fig. 6) demonstrated that the MLC-ALK gene is transcribed during late embryogenesis (the time of formation of the larval musculature), during the larval instars (a time of rapid tissue, and therefore muscle, growth), and during the late stages of pupariation (the time of rapid synthesis of the indirect flight muscle and other adult muscles). There was no accumulation of MLC-ALK transcripts during early pupal development, the time of histolysis of larval musculature (8). The most abundant RNA showed a broad distribution with a mean molecular length of around 0.95 kb. There is a less intense band at 1.3 kb, which was most evident in Fig. 6 in the late pupal indirect flight muscle RNA lane, because of the greater fraction of muscle-specific RNA in this preparation. A 3.2-kb band which we believe is an unspliced nuclear transcript (unpublished data) is faintly visible in several lanes.

Reiteration frequency of the MLC-ALK gene. There are indications of multiple polypeptides in the in vitro translation products of RNA selected by the MLC-ALK gene we isolated (Fig. 3). Additionally, cross-hybridizing RNAs are expressed in several stages of development (Fig. 6). Similar results were observed for the actin genes of *D. melanogaster*, which comprise a small multigene family (11), and for the myosin heavy-chain gene which is single copy (5, 31). If the MLC-ALK gene is reiterated, all of the copies must reside at the chromosomal region 98B (Table 1), because this is the only site labeled by in situ hybridization.

We carried out genome blot studies to resolve this question. The gel blots were hybridized to the MLC-ALK probes at a moderate stringency. Quantitative comparisons of intensity were made by genome reconstruction experiments. All these data (Fig. 7) show that at the moderate stringency of the hybridization experiments there is only one MLC-ALK gene per haploid genome.

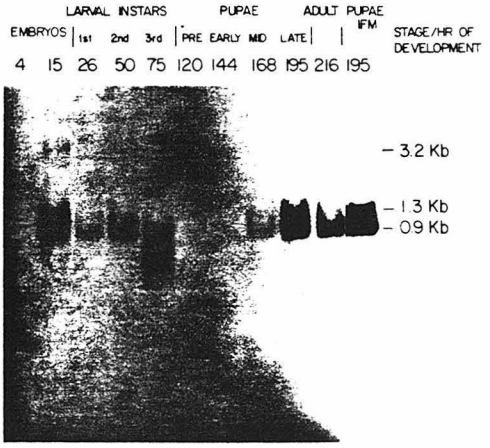


FIG. 6. Developmental expression of myosin alkali light chain mRNAs. Poly(A)$^+$ RNA (1 μg) from different developmental stages of synchronized populations of *D. melanogaster* was subjected to electrophoresis on 1.5% agarose gels containing 2.2 M formaldehyde. The RNA was transferred to nitrocellulose and hybridized with a nick-translated probe synthesized from p102.6 DNA. Time (in hours) was calculated from the time of egg deposition. Pupae were resynchronized by floatation on water at 5 h post-pupariation (22). *E. coli* 16S and 23S rRNAs as well as HeLa 28S and 18S rRNAs were used as length standards.

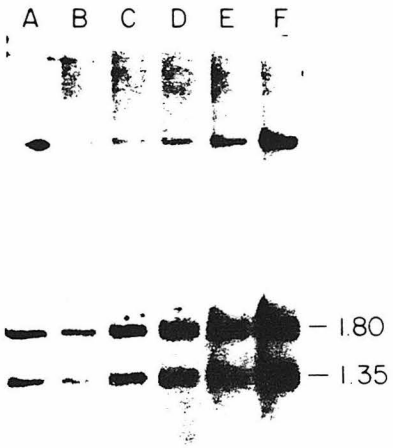


FIG. 7. Genomic representation of the coding region of λdmpT102. (A) Genomic *Drosophila* Canton S DNA (4 μg) was digested with *Eco*RI and *Hin*dIII, and the resulting fragments were separated on 1% agarose gels. In adjacent lanes of the same gels, amounts of *Eco*RI-*Hin*dIII-restricted λdmpT102 DNA equivalent to (B) 0.5, (C) 1.0, (D) 3.0, (E) 5.0, and (F) 25.0 copies per genome (a haploid genome size of 1.6×10^8 base pairs was assumed) plus 4 μg of *Eco*RI-*Hin*dIII-restricted *E. coli* DNA were subjected to electrophoresis. After electrophoresis, the fragments were transferred to nitrocellulose and hybridized with ^{32}P-labeled nick-translated p102.6 DNA.

DISCUSSION

We cloned a *D. melanogaster* gene coding for a myosin alkali light-chain protein. The gene is expressed to provide a relatively high concentration of RNA during those specific times in development when extensive myofibrillar assembly occurs, and it is not expressed before larval myogenesis or during early pupariation when larval muscles are undergoing histolysis (8). We showed by in vitro translation of hybrid selected RNA that the gene encodes a polypeptide which comigrates in two-dimensional gels with myosin alkali light-chain protein extracted from adult indirect flight muscle. The strongest evidence that the gene encodes the MLC-ALK protein was obtained from the nucleotide sequence of a cDNA clone derived from the mRNA transcript of the gene.

```
                                                          0         10         20
DRO L-1                                                                MADVPKREVENVEFVFEVMG
CSM L-1     PKK*DVKKPAAAAAPAPAPAPAPAPAPAKPKEPAIDLKSIKIEFSKEQQDDFK-AFLLFDR
CSM L-3                                                                SFS-DQIDDFK-AFLLFDR
CCM L-1     PKKPEPKKAPEPKKEEPKPAPKPAEPEPK*KEVEFNPASIKVEFT-DQIEEFK-AFSLFDR
CGM 17K                                                                CDFSEEQTA-FK-AFQLFDR

            30         40         50         60         70         80
SPGE**GIDAVDLGDALRALNLNPTLALIEK*LGGTKKR**NEKKIKLDEFLPIYSNVKKEKEQGCTEDFIE
TGDA**K-TLSQV--IV---GQ---N-E-N-I--NPS-EEM-A---TFE----MLQAAANN-D--TF---V-
TGDA**K-TLSQV--IV---GQ---N-E-N-I--NPS-EEM-A---TFE----MLQAAANN-D--TF---V-
T-KSEMK-TYAQC--V----GQ---Q-EVM-V--RP-QEEM-S-M-DFET---MLQHIS-T-DT-TY---V-
TGDG**K-LYSQC--VM---GQ---N-EVM-V--NP-SDEM-L-TLNFEQ---MMQTIA-N-D---F--YV-
                                                                        n
 90        100        110        120        130        140        150
CLKLTDKEENGTMLLAELNHALLALGESLDDEQVETLFADCMDPEDDEGFIPTSNFVNRLMSDPVVFD
G-RVF---G---VMG---R-V-AT---KMTE-E--E-MKGQE-SNGCINYEAFVKHIMSV
G-RVF---G---VMG---R-V-AT---KMTE-E--E-MKGQE-SNGCINYEAFVKHIMSV
G-RVF---G---VMG---R-V-AT---R-TE-E-DK-M-GQE-ANGCINYEAFVKHIMAN
G-RVF---G---VMG--IR-V-VT---KMTE-E--Q-V-GHE-SNGCINYEELVRM-LSG
nn  n            n  nn n
     x y z  -x
           -y   -z
```

FIG. 8. Comparison of the amino acid sequence of *Drosophila* myosin light chain with the sequence of chicken myosin light chains (7, 19). Abbreviations: DRO L-1, *Drosophila* myosin alkali light chain; CSM L-1, chicken skeletal muscle myosin light chain-1; CSM L-3, chicken skeletal muscle myosin light chain-3; CCM L-1, chicken cardiac muscle myosin light chain-1; CGM 17K, chicken gizzard muscle 17,000-dalton myosin light chain. *Drosophila* amino acids identical to those of the chicken myosin light chains are indicated by a hyphen. The absence of a particular amino acid in the sequence is denoted by an asterisk; n indicates those residues which form the core of the E and F α helices. X, Y, Z, −X, −Y, and −Z are residues which might be involved in the binding of divalent metal ions in EF hand domain III.

The corresponding amino acid sequence showed a high degree of sequence homology with that of the chicken alkali light chains and much less homology with any other of the evolutionarily related calcium ion-binding proteins.

The consensus EF hand structural domain of the calcium-binding proteins consists of an α helix of 9 amino acids (the E helix), a loop of 12 amino acids (6 of which have side chains containing oxygen atoms which are capable of octahedral coordination to a Ca^{2+} ion), and an α helix of eight amino acids (the F helix) (14). This structure is denoted n(2), n(5), n(6), n(9), X(10), Y(12), Z(14), −Y(16), −X(18), −Z(21), n(22), n(25), n(26), n(28), n(29). The numbers denote the relative positions of residues, where n represents the hydrophobic or inner core of the α helix and X, Y, Z, −X, −Y, and −Z represent the calcium ligands. The amino acids at these positions are variable from protein to protein. To bind Ca^{2+}, there must be at least four aspartic and glutamic acid residues in the calcium loop, but others may be serine, threonine, asparagine, and glutamine. In some calcium-binding proteins, one of the six positions is occupied by glycine which is capable of coordinating Ca^{2+} through the oxygen of a hydrogen-bonded water molecule (14). The vertebrate myosin alkali light chains do not, in fact, bind Ca^{2+} in vitro (39). However, because of the very similar amino acid composition of their EF hand domains with those of other calcium-binding proteins, they are included in this group.

A comparison of the amino acid sequence of the *Drosophila* MLC-ALK gene with those of the highly homologous chicken skeletal, cardiac, and gizzard alkali light chains is shown in Fig. 8. The region of greatest homology between the chicken proteins and the *Drosophila* protein extends from amino acids 31 to 124 of the latter. There is an EF hand (denoted EF hand III) of the chicken proteins in a region of very high sequence homology with the *Drosophila* protein, amino acids 85 to 112. The chicken MLC-ALK proteins have four domains which are identified as resembling the EF hand domains of other calcium-binding proteins. However, due to an insertion of two amino acids in the Ca^{2+} loop of EF hand domain II, it is believed that this structure has been disrupted. Weeds et al. (39) note that only the EF hand in domain III has four acidic residues among its six ligating groups.

Sequence comparisons of aligned amino acids for chicken MLC-3 and the *Drosophila* protein are shown in Table 2. The highest degree of homology is in domain III, with a moderate degree of homology in domain II and very low homology in domains I and IV. Analysis of domain II did not show a good EF hand structure. Therefore, we believe that the only potential calcium-binding site of the *Drosophila* protein is that found in the region homologous to domain III of the chicken proteins.

The best-fit matrix analysis showed a definite but more distant relationship between domain III of the *Drosophila* protein and domain I of the chicken protein (Fig. 5C), calmodulin, and troponin C (data not shown). This is reasonable because the four Ca^{2+}-binding domains of calcium-binding proteins are thought to have arisen from an ancestral Ca^{2+} binding domain that underwent two gene duplication events followed by sequence divergence (14). Similar comparisons show sequence homology between this domain of the *Drosophila* protein with the EF hand III domains of bovine brain calmodulin and of rabbit skeletal muscle troponin C (data not shown). This EF hand domain contains the high-affinity Ca^{2+}–Mg^{2+}-binding site.

We conclude that the *Drosophila* alkali light chain contains one EF hand domain which is related in sequence to those of the family of calcium-binding proteins.

The *Drosophila* MLC-ALK gene is present at only one copy per haploid genome and is transcribed during larval and pupal myogenesis. In this respect, it is analogous to the myosin heavy-chain gene (5, 31). However, the latter shows developmental differences in transcript lengths and splicing patterns (31) which lead to developmental differences in amino acid sequences (C. Rozek, personal communication).

TABLE 2. Homology by domain of the chicken and *Drosophila* proteins

Domain (amino acids)	% Homology		% Protein homology allowing conservative amino acid changes	% DNA homology discounting third-base changes	% Nonhomology due only to third-base changes	Protein homology/DNA homology
	Protein[a]	DNA[b]				
I (1–32)	15	34	21	39	8	0.44
II (33–76)	46	52	50	59	15	0.88
III (77–115)	58	57	70	68	33	1.02
IV (116–157)	15	27	18	30	4	0.55

[a] Homology between the chicken skeletal muscle myosin light chain-3 and *Drosophila* myosin alkali light chain (Fig. 4).
[b] Homology between the *Drosophila* myosin light-chain sequence and that of chicken skeletal muscle myosin light chain-3 cDNA clone (26).

There are indications of heterogeneity in molecular weight and isoelectric point for the *Drosophila* MLC-ALK in vitro translation product and for this protein extracted from adult indirect flight muscle. This heterogeneity could be due to post-translational modification, such as phosphorylation. Alternatively, or in addition, there could be subtle differences in transcripts, thereby encoding proteins which differ slightly in amino acid composition, which are not revealed by the RNA gel blots. We have evidence that some of this heterogeneity is due to differential splicing generating transcripts encoding two proteins which differ at their carboxy-terminal ends. Studies of these and related questions dealing with the fine structure of the gene will be reported in a later communication (Falkenthal et al., in preparation).

ACKNOWLEDGMENTS

We thank E. Fyrberg and C. Rozek for their generous gifts of actin and myosin heavy-chain-coding DNAs and T. Hunkapiller for help in the execution of the protein homology comparisons. We also thank H. Mitchell and N. Peterson for guidance and direction in pupal dissection and sharing their results of pupal indirect flight muscle development.

S.F. was supported by a Muscular Dystrophy Postdoctoral Fellowship and NRSA Fellowship 1 F32 GM08895-01. The research was supported by Public Health Service grants from the National Institutes of Health.

LITERATURE CITED

1. **Anderson, K. V., and J. A. Lengyel.** 1979. Rates of synthesis of major classes of RNA in Drosophila embryos. Dev. Biol. **70**:217–231.
2. **Barker, W. C., L. K. Ketcham, and M. O. Dayhoff.** 1978. The troponin C superfamily, p. 273–284. *In* M. O. Dayhoff (ed.), Atlas of protein sequence and structure, vol. 5. National Biomedical Research Foundation, Silver Spring, Md.
3. **Bautch, V. L., R. V. Storti, D. Mischke, and M. L. Pardue.** 1982. Organization and expression of Drosophila tropomyosin genes. J. Mol. Biol. **162**:231–250.
4. **Benton, W. D., and R. W. Davis.** 1978. Screening λgt recombinant clones by hybridization to single plaques in situ. Science **196**:180–182.
5. **Bernstein, S. I., K. Mogami, J. J. Donady, and C. P. Emerson.** 1983. Drosophila muscle myosin heavy chain is encoded by a single gene located in a cluster of muscle mutations. Nature (London) **302**:393–397.
6. **Brandhorst, B. P.** 1976. Two-dimensional gel patterns of protein synthesis before and after fertilization of sea urchin eggs. Dev. Biol. **52**:310–317.
7. **Collins, J. H., J. D. Potter, M. J. Horn, G. Wilshire, and N. Jackman.** 1973. The amino acid sequence of rabbit skeletal muscle troponin C: gene replication and homology with calcium binding proteins from carp and hake muscle. FEBS Lett. **36**:268–272.
8. **Crossley, A. C.** 1978. The morphology and development of the Drosophila muscular system, p. 499–560. *In* M. Ashburner and T. R. F. Wright (ed.), The genetics and biology of Drosophila, vol. 2b. Academic Press, Inc., New York.
9. **Deng, G., and R. Wu.** 1981. An improved procedure for utilizing terminal transferase to add homopolymers to 3′ termini of DNA. Nucleic Acids Res. **9**:4173–4188.
10. **Ebashi, S.** 1963. Third component participating in the superprecipitation of "natural actomyosin." Nature (London) **200**:1010–1012.
11. **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.
12. **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**:115–123.
13. **Klee, C. B., T. H. Crouch, and P. G. Richman.** 1980. Calmodulin. Annu. Rev. Biochem. **49**:489–515.
14. **Kretsinger, R. H.** 1980. Structure and evolution of calcium modulated proteins. Crit. Rev. Biochem. **8**:119–174.
15. **Laskey, R. A., and A. D. Mills.** 1975. Quantitative fluorography of acrylamide gels. Eur. J. Biochem. **56**:335–341.
16. **Lowey, S., and D. Risby.** 1971. Light chain from fast and slow muscle myosins. Nature (London) **234**:81–85.
17. **Maniatis, T., R. C. Hardison, E. Lacy, Jr., J. Lauer, C. O'Connell, D. Quon, G. K. Sim, and A. Efstratiadis.** 1978. The isolation of structural genes from libraries of eukaryotic DNA. Cell **15**:687–702.
18. **Matsuda, G., T. Maita, Y. Kato, J. Chen, and T. Umegane.** 1981. Amino acid sequences of the cardiac L-2A, L-2B, and gizzard 17,000 M light chains of chicken muscle myosin. FEBS Lett. **135**:232–236.
19. **Matsuda, G., T. Maita, and T. Umegane.** 1981. The primary structure of L-1 light chain of chicken fast skeletal muscle myosin and its genetic implication. FEBS Lett. **126**:111–113.
20. **Maxam, A. M., and W. Gilbert.** 1980. Sequencing end-labeled DNA with base-specific chemical cleavages. Methods Enzymol. **65**:499–560.
21. **Meyerowitz, E. M., and D. S. Hogness.** 1982. Molecular organization of a Drosophila puff site that responds to ecdysone. Cell **28**:165–176.
22. **Mitchell, M. K., and A. Mitchell.** 1964. Mass culture and age selection in Drosophila. Dros. Inf. Ser. **39**:135–137.
23. **Mogami, K., S. C. Fujita, and Y. Hotta.** 1982. Identification of Drosophila indirect flight muscle myofibrillar proteins by means of two-dimensional electrophoresis. J. Biochem. **91**:643–650.
24. **Mullins, J. I., J. W. Casey, M. O. Nicolson, K. B. Burck, and N. Davidson.** 1981. Sequence arrangement and biological activity of cloned feline leukemia virus proviruses from a virus-productive human cell line. J. Virol. **38**:688–703.
25. **Mullins, J. I., J. Casey, M. Nicolson, and N. Davidson.** 1980. Sequence organization of feline leukemia virus DNA in infected cells. Nucleic Acids Res. **8**:3287–3305.
26. **Nabeshima, Y., Y. Fuju-Kuriyama, M. Muramatsu, and K. Ogata.** 1982. Molecular cloning and nucleotide sequences of the complementary DNAs to chicken skeletal muscle myosin two alkali light chain mRNAs. Nucleic Acids Res. **10**:6099–6110.
27. **O'Farrell, P. H.** 1975. High resolution two-dimensional electrophoresis of proteins. J. Biol. Chem. **250**:4007–4021.
28. **Okayama, H., and P. Berg.** 1982. High efficiency cloning of full length cDNAs. Mol. Cell. Biol. **2**:161–170.
29. **Pemrick, S. M.** 1980. The phosphorylated L2 light chain of

skeletal myosin is a modifier of the actomyosin ATPase. J. Biol. Chem. **255**:8836–8841.
30. **Ricciardi, R. P., J. S. Miller, and B. E. Roberts.** 1979. Purification and mapping of specific mRNAs by hybridization-selection and cell-free translation. Proc. Natl. Acad. Sci. U.S.A. **76**:4927–4931.
31. **Rozek, C. E., and N. Davidson.** 1983. Drosophila has one myosin heavy chain gene with three developmentally regulated transcripts. Cell **32**:23–34.
32. **Sarkar, S., F. A. Streter, and J. Gergely.** 1971. Light chains of myosins from white, red, and cardiac muscles. Proc. Natl. Acad. Sci. U.S.A. **68**:946–950.
33. **Snyder, M., J. Hirsh, and N. Davidson.** 1981. The cuticle genes of Drosophila: a developmentally regulated gene cluster. Cell **25**:165–177.
34. **Snyder, M., M. Hunkapiller, D. Yuen, D. Silvert, J. Fristrom, and N. Davidson.** 1982. Cuticle protein genes of Drosophila: structure, organization, and evolution of four clustered genes. Cell **29**:1027–1040.
35. **Ullrich, A., J. Shine, J. Chirguin, R. Pictet, E. Tischer, W. J. Rutter, and H. M. Goodman.** 1977. Rat insulin genes: construction of plasmids containing the coding sequences. Science **196**:1313–1319.
36. **Wagner, P. D., and E. Giniger.** 1981. Hydrolysis of ATP and reversible binding to F-actin by myosin heavy chains free of all light chains. Nature (London) **292**:560–562.
37. **Watterson, D. M., F. Sharief, and T. Vanaman.** 1980. The complete amino acid sequence of the calcium dependent modulator protein (calmodulin) of bovine brain. J. Biol. Chem. **255**:962–975.
38. **Weeds, A. G., and S. Lowey.** 1971. Substructure of the myosin molecule II. The light chains of myosin. J. Mol. Biol. **61**:701–725.
39. **Weeds, A., P. Wagner, R. Jakes, and J. Kendrick-Jones.** 1977. Structure and function of myosin light chains, p. 222. *In* R. H. Wasserman et al. (ed.), Calcium binding proteins and calcium function. Elsevier/North-Holland Publishing Co., Amsterdam.
40. **Whalen, R. G., G. S. Butler-Browne, and F. Gros.** 1978. Identification of a novel form of myosin light chain present in embryonic muscle tissue and cultured muscle cells. J. Mol. Biol. **126**:415–431.
41. **Yen, P. H., and N. Davidson.** 1980. The gross anatomy of a tRNA gene cluster at region 42A of the Drosophila melanogaster chromosome. Cell **22**:137–148.