Characterization of the *cis*-Regulatory Region of the Drosophila Homeotic Gene *Sex combs reduced*

Joseph G. Gindhart, Jr.,[1] A. Nicole King[2] and Thomas C. Kaufman

Howard Hughes Medical Institute and Department of Biology, Indiana University, Bloomington, Indiana 47405

Manuscript received July 13, 1994
Accepted for publication November 3, 1994

ABSTRACT

The Drosophila homeotic gene *Sex combs reduced* (*Scr*) controls the segmental identity of the labial and prothoracic segments in the embryo and adult. It encodes a sequence-specific transcription factor that controls, in concert with other gene products, differentiative pathways of tissues in which *Scr* is expressed. During embryogenesis, Scr accumulation is observed in a discrete spatiotemporal pattern that includes the labial and prothoracic ectoderm, the subesophageal ganglion of the ventral nerve cord and the visceral mesoderm of the anterior and posterior midgut. Previous analyses have demonstrated that breakpoint mutations located in a 75-kb interval, including the *Scr* transcription unit and 50 kb of upstream DNA, cause *Scr* misexpression during development, presumably because these mutations remove *Scr cis*-regulatory sequences from the proximity of the *Scr* promoter. To gain a better understanding of the regulatory interactions necessary for the control of *Scr* transcription during embryogenesis, we have begun a molecular analysis of the *Scr* regulatory interval. DNA fragments from this 75-kb region were subcloned into *P*-element vectors containing either an *Scr-lacZ* or *hsp70-lacZ* fusion gene, and patterns of reporter gene expression were assayed in transgenic embryos. Several fragments appear to contain *Scr* regulatory sequences, as they direct reporter gene expression in patterns similar to those normally observed for *Scr*, whereas other DNA fragments direct *Scr* reporter gene expression in developmentally interesting but non-*Scr*-like patterns during embryogenesis. *Scr* expression in some tissues appears to be controlled by multiple regulatory elements that are separated, in some cases, by more than 20 kb of intervening DNA. Interestingly, regulatory sequences that direct reporter gene expression in an *Scr*-like pattern in the anterior and posterior midgut are imbedded in the regulatory region of the segmentation gene *fushi tarazu* (*ftz*), which is normally located between 10 and 20 kb 5′ of the *Scr* transcription start site. This analysis provides an entry point for the study of how *Scr* transcription is regulated at the molecular level.

THE homeotic genes of *Drosophila melanogaster* are necessary for the establishment and maintenance of segmental identity during embryonic and larval development. These genes act as developmental switches that trigger a cascade of events leading to the characteristic morphology of each segment (OUWENEEL 1976; GARCIA-BELLIDO 1977; LEWIS 1978; WAKIMOTO and KAUFMAN 1981). Genetic analyses have shown that Antennapedia complex (ANT-C) homeotic genes *labial* (*lab*), *proboscipedia* (*pb*), *Deformed* (*Dfd*), *Sex combs reduced* (*Scr*) and *Antennapedia* (*Antp*) control the segmental identity of the head and thorax, whereas the BX-C genes *Ultrabithorax* (*Ubx*), *Abdominal-A* (*abd-A*) and *Abdominal-B* (*Abd-B*) control the identity of posterior thoracic and abdominal segments (reviewed in DUNCAN 1987; KAUFMAN *et al.* 1990). The establishment of segment-specific patterns of homeotic gene expression during embryogenesis is controlled primarily by the gap and segmentation class genes (reviewed in AKAM 1987; SCOTT and CARROLL 1987; INGHAM 1988). Once established, patterns of homeotic gene expression are maintained by various mechanisms, including autogenous control (BIENZ and TREMML 1988; KUZIORA and MCGINNIS 1988; MANN and HOGNESS 1990; CHOUINARD and KAUFMAN 1991), cross-regulatory interactions between homeotic gene products (HAFEN *et al.* 1984; STRUHL and WHITE 1985) and the activity of the *Polycomb* (*Pc*) and *trithorax* (*trx*) groups of *trans*-regulators (reviewed in PARO 1990; KENNISON 1993). To better understand the mechanism by which segmental identity is established and maintained during Drosophila development, we have begun an analysis of how expression of the homeotic gene *Sex combs reduced* (*Scr*) is initiated by gap and pair-rule gene products and how later interactions refine and maintain the established *Scr* expression pattern.

The *Scr* gene is required for the normal development of the labial and prothoracic segments in the embryo and adult (reviewed in KAUFMAN *et al.* 1990). The expression pattern of *Scr* is consistent with its develop-

Corresponding author: Thomas C. Kaufman, Howard Hughes Medical Institute, Indiana University, Jordan Hall, 3rd St. and Faculty, Bloomington, IN 47405.

[1] *Present address:* Howard Hughes Medical Institute, Division of Cellular and Molecular Medicine, UCSD School of Medicine, 9500 Gilman Dr., La Jolla, CA 92093–0683.

[2] *Present address:* Department of Cellular and Developmental Biology, Harvard University, 16 Divinity Ave., Cambridge, MA 02138.

mental function. *Scr* protein is first detected in PS2 during gastrulation (LEMOTTE *et al.* 1989). In germband extended embryos, *Scr* protein accumulates in the PS2 and lateral PS3 ectoderm (RILEY *et al.* 1987; CARROLL *et al.* 1988; MAHAFFEY *et al.* 1989). Later in development, *Scr* protein continues to accumulate in the labial and prothoracic ectoderm, with additional accumulation in the subesophageal ganglion of the central nervous system (CNS) (MAHAFFEY and KAUFMAN 1987; RILEY *et al.* 1987) and the anterior and posterior parts of the visceral mesoderm (LEMOTTE *et al.* 1989; REUTER and SCOTT 1990). Putative *trans*-acting regulators of *Scr* have been identified by studying the pattern of *Scr* expression in embryos mutant for individual loci. Using this strategy, gap genes such as *hunchback* (*hb*), *giant* (*gt*) and *Krüppel* (Kr); the segmentation genes *fushi tarazu* (*ftz*), *even-skipped* (*eve*) and *odd-paired* (*opa*) and the homeotic genes *Antennapedia* (*Antp*) and *teashirt* (*tsh*) have been implicated as possible regulators of the establishment of *Scr* expression at the transcriptional level (INGHAM and MARTINEZ-ARIAS 1986; RILEY *et al.* 1987; TREMML and BIENZ 1989b; PANGANIBAN *et al.* 1990; FASANO *et al.* 1991). Thus, the pattern of *Scr* expression established during embryogenesis is thought to be controlled by *trans*-acting factors that exert their effects by binding to enhancer sequences in the *Scr* regulatory region.

The molecular map of the *Scr* region of the ANT-C is shown in Figure 1. The *Scr* transcript is composed of three exons, separated by introns 6 and 14 kb in length (LEMOTTE *et al.* 1989). The segmentation gene *ftz* is located ~15 kb 5′ of the *Scr* transcription start site, and the 3′ end of the homeotic gene *Antp* is located ~50 kb upstream of *Scr* (SCOTT *et al.* 1983; KUROIWA *et al.* 1985). A set of DNA breakpoints associated with mutations that affect *Scr* function or regulation have been mapped to the 75-kb interval between the 3′ end of *Scr* and the 3′ end of *Antp* (vertical arrows) (SCOTT *et al.* 1983; KUROIWA *et al.* 1985; PATTATUCCI and KAUFMAN 1991; PATTATUCCI *et al.* 1991). The 5′ limit of the *Scr* regulatory region is not precisely known; however, the most proximal *Antp* breakpoint mutation, $Antp^{14}$, fully complements *Scr* and maps to +120 on the ANT-C molecular map, ~70 kb upstream of the *Scr* promoter (KAUFMAN *et al.* 1990). These data suggest that the regulatory region controlling *Scr* expression is quite large. *Scr* breakpoint mutations can be placed into two groups based upon their ability to complement null *Scr* alleles. The first group consists of breakpoints that fail to complement *Scr*, resulting in embryonic lethality and strong cuticular transformations of the labial segment toward a maxillary identity and the prothoracic segment toward a mesothoracic identity (WAKIMOTO and KAUFMAN 1981; PATTATUCCI *et al.* 1991). They map to a 45-kb interval that includes the *Scr* transcription unit and 20 kb of upstream DNA (Figure 1). The second class of mutations includes hypomorphic alleles that partially complement *Scr* null mutations, resulting in larval and pupal lethality but some adult survival. Hypomorphic alleles map to the interval between *ftz* and *Antp*, located 20–50 kb 5′ of the *Scr* transcription start site (Figure 1). Larvae hemizygous for Scr hypomorphic mutations exhibit few or no cuticular transformations (PATTATUCCI *et al.* 1991). Therefore, these hypomorphic alleles direct the expression of *Scr* at high enough levels to rescue embryonic lethality but lack regulatory activity necessary for normal postembryonic development.

This paper describes our attempt to further localize sequences that control the expression of *Scr* during embryogenesis. The genetically defined *Scr* regulatory region was subdivided into a series of overlapping fragments (Figure 1) and subcloned into *P*-element vectors containing either an *hsp70-lacZ* or *Scr-lacZ* fusion gene. Germline transformants containing these constructs were generated, and the pattern of β-galactosidase (β-gal) accumulation in transformant embryos was observed. Using this strategy, regulatory sequences that direct reporter gene expression in a subset of the *Scr* expression domain were identified. The *Scr* regulatory region also appears to contain regulatory sequences that block ectopic *Scr-lacZ* fusion gene expression. In addition, several DNA fragments from the *Scr* regulatory interval direct reporter gene expression in novel spatial patterns. These results and the insights they provide with respect to *Scr* regulation are discussed.

MATERIALS AND METHODS

Plasmid construction: Genomic DNA fragments from the *Scr* locus were isolated from phage clones characterized during a chromosome walk through the Antennapedia complex (SCOTT *et al.* 1983). The size and location of these fragments are shown in Figure 1. These fragments were subcloned into the shuttle vectors pHSS7 (SEIFERT *et al.* 1986) or pHSS7phil (RANDAZZO 1991), two Kan^r vectors that contain several unique restriction sites flanked by *Not*I restriction sites. *Scr* regulatory fragments were excised from pHSS7 and pHSS7phil by *Not*I digestion, then the *Not*I fragments were inserted into a unique *Not*I site 5′ of the reporter genes in BSR (Big *Scr* Reporter gene), BSRN (Big *Scr* Reporter gene with Nuclear localization sequence), SSRN (Small *Scr* Reporter gene with Nuclear localization sequence), HZR (HZ-white Reporter gene) and/or HZ50PL (HIROMI and GEHRING 1987).

The *Scr-lacZ* fusion plasmid constructs are based upon the *P*-element transformation vector CaSPeRB-gal (THUMMEL *et al.* 1988). This vector contains a *white* minigene for transformant screening and the bacterial *lacZ* gene for generating protein fusions. First, CaSPeRB-gal was cut with *Bam*HI and *Xba*I to remove its polylinker, which was replaced with a polylinker that created unique cloning sites for *Not*I, *Eco*RI and *Bam*HI. This new plasmid, called CaSPeRB-galM, was digested with *Bam*HI and *Eco*RI, then ligated to an 8.5-kb genomic *Eco*RI/*Bam*HI fragment containing 2.3 kb of sequences upstream of the *Scr* promoter, *Scr* exon I, intron I and ~100 bp of exon II (LEMOTTE *et al.* 1989). This construct, called BSR, has an *Scr-lacZ* protein fusion containing the amino-terminal three amino acids of *Scr* fused to the amino terminus of β-gal at the corresponding *Bam*HI site. The genomic fragments

3.0 *Eco*RI, 5.5 *Hin*dIII, 3.5 *Sal*I/*Kpn*I, 3.7 *Hin*dIII and 7.0 *Eco*RI were subcloned as *Not*I fragments into BSR. The vector SSR is identical to BSR, except it does not contain intron I. This vector was constructed by ligating CaSPeRB-galM cut with *Bam*HI and *Eco*RI to a 3.0-kb *Bam*HI/*Eco*RI fragment containing 2.3 kb of DNA 5′ of *Scr* exon I and an exonI–exonII fusion created by ligating genomic and cDNA sequences (supplied by DAVID MILLER). SSR was then digested with *Bam*HI, and a 39-bp double-stranded oligonucleotide encoding the nuclear localization sequence of SV40 large T antigen (KALDERON *et al.* 1984; RIDDIHOUGH and ISH-HOROWICZ 1991) was inserted into the *Bam*HI site to create SSRN. The sequence of NLS1 is GATCCCACCCCCCCCAAGAAGAAGCGCAAGGTGGAGGAC, whereas the sequence GATCGTCCTCCACCTTGCGCTTCTTCTTGGGGGGGGTGG represents NLS2. The nuclear localization sequence encodes amino acids 3–13 (TPPKKKRKVED) of the *Scr-lacZ* fusion protein and enables it to be recognized by the nuclear protein import pathway. All genomic subclones shown in Figure 1 except 3.8 *Hin*dIII, 5.2 *Hin*dIII, 2.6 *Hin*dIII and 3.0 *Eco*RI were subcloned as *Not*I fragments SSRN. The vector BSRN was constructed by cutting SSRN with *Bam*HI and *Eco*RI and then ligating the cut vector to an 8.5-kb *Bam*HI-*Eco*RI fragment from BSR that contains *Scr* genomic sequences. The resulting vector is identical to BSR, except BSRN contains a nuclear localization sequence. The genomic DNA fragments 8.2 *Xba*I and 10.0 *Xba*I were subcloned as *Not*I fragments into BSRN. Maps of BSRN and SSRN are shown in Figure 1.

The transformation vector HZR (formerly HZ-white) (GINDHART 1993) was constructed by replacing the *rosy* gene in HZ50PL (HIROMI and GEHRING 1987) with a *white* minigene (PIRROTTA 1988). To do this, a 4.1-kb *Spe*I fragment containing the *white* minigene (PIRROTTA 1988, subclone provided by DAVID CRIBBS) was inserted into a unique *Xba*I site in HZ50PL. The *rosy* gene was removed from this intermediate vector by digestion with *Hin*dIII and religation of the appropriate fragments. HZR contains an *hsp70-lacZ* reporter gene that has *hsp70* promoter sequences to −50 with respect to the transcription start site (HIROMI and GEHRING 1987). The genomic fragments 5.2 *Hin*dIII, 2.6 *Hin*dIII, 3.0 *Eco*RI, 2.4 *Hin*dIII, 5.5 *Hin*dIII and 3.0 *Xba*I were subcloned as *Not*I fragments into HZ50PL, and the remaining fragments were subcloned as *Not*I fragments into HZR.

***P*-element transformation:** Germline transformation of P{w^+} constructs was performed essentially as described in ROBERTSON *et al.* (1988). A solution of 0.5–1.0 mg ml^{-1} of each construct was injected into 0- to 45-min AED embryos resulting from the cross *w*;P{*ry*$^+$,Δ2,3} females × *w*; *TM3/TM6B* males. The P{ry$^+$} constructs were co-injected with the *P*-element helper plasmid p*P*25.7wc (KARESS and RUBIN 1984) into *ve st ry* embryos. The resulting G_0 adults were crossed to w^1 or *ve st ry*, depending upon the construct injected. Transformants were isolated on the basis of the rescue of *white* or *rosy*. Multiple transformant lines were isolated for each construct.

Immunohistochemistry: Transgenic and control embryos were fixed and stained using the procedure described in MAHAFFEY and KAUFMAN (1987). After staining, embryos were dehydrated in ethanol, cleared in methylsalicylate and then viewed and photographed on a Zeiss Axiophot microscope with Nomarski optics. Polyclonal rabbit SCR antisera were provided by MARIE MAZZULLA and used at a 1:150 dilution. Monoclonal mouse anti-SCR is described in GLICKSMAN and BROWER (1988). 6H4 (anti-SCR) cell supernatant was concentrated using a Centricon 30 column (Amicon) and then used at a 1:10 dilution. Monoclonal anti-β-gal was obtained from Boehringer Mannheim and used at a 1:2000 dilution. Goat anti-rabbit (1:150) and goat anti-mouse (1:200) secondary antibodies conjugated to horseradish peroxidase were purchased from Bio-Rad. Polyclonal rabbit anti-β-gal was provided by DAVID MILLER and used at a 1:150 dilution. Diaminobenzidene was the chromagen substrate used in all experiments. Embryos were staged according to CAMPOS-ORTEGA and HARTENSTEIN (1985).

RESULTS

Reporter gene expression: Four reporter gene constructs were made to facilitate the search for *Scr* regulatory elements. The plasmid maps of these constructs are shown in Figure 1. SSRN is an *Scr-lacZ* fusion gene vector that contains 2.3 kb of upstream DNA, exon I and part of exon II fused to *lacZ*. Plasmid constructs BSR and BSRN contain all *Scr* sequences that are in SSRN as well as intron I sequences. HZR and HZ50PL (HIROMI and GEHRING 1987) are *hsp70-lacZ* fusion gene vectors. Genomic DNA fragments from the *Scr* locus were subcloned into these vectors, transformants were generated and the pattern of β-gal accumulation in transformant embryos was assayed (Table 1). The rationale of this approach was that if a DNA fragment contains enhancer elements that direct all or a subset of the endogenous *Scr* expression pattern, then these same fragments, when placed upstream of a reporter gene, should direct its expression in a similar pattern. Observation of multiple independent transformant lines reduced the possibility that observed β-gal patterns resulted from the influence of DNA sequences flanking the site of chromosome insertion rather than regulatory sequences within the enhancer tester construct. Each reporter gene expression pattern described below, unless otherwise noted, is the "consensus" pattern observed in the majority of transformant lines containing a particular construct.

The use of both *hsp70-lacZ* and *Scr-lacZ* reporter genes to assay the regulatory activity of genomic DNA fragments from the *Scr* locus is primarily the result of the strategy we used to detect *Scr* enhancers. At first, genomic DNA fragments from *Scr* were subcloned into HZR or HZ50PL to identify *Scr* positive regulatory elements. Whereas some fragments directed *hsp70-lacZ* reporter gene product accumulation in *Scr*-like patterns, most fragments contained regulatory elements that directed non-*Scr*-like expression patterns (Table 1). These results led us to ask the following question: Why is *Scr* not expressed in these cryptic patterns? One possibility is that the *Scr* promoter and the *hsp70* promoter are functionally heterogeneous, in that some enhancers can activate transcription from the *Scr* promoter but not the *hsp70* promoter, whereas other enhancers activate the *hsp70* promoter but not *Scr*. An example of promoter-enhancer specificity is described by WEFALD *et al.* (1990), who showed that a human myoglobin enhancer can activate transcription from its own promoter but not the SV40 promoter. To determine if enhancers in the *Scr* regulatory region are promoter specific, the

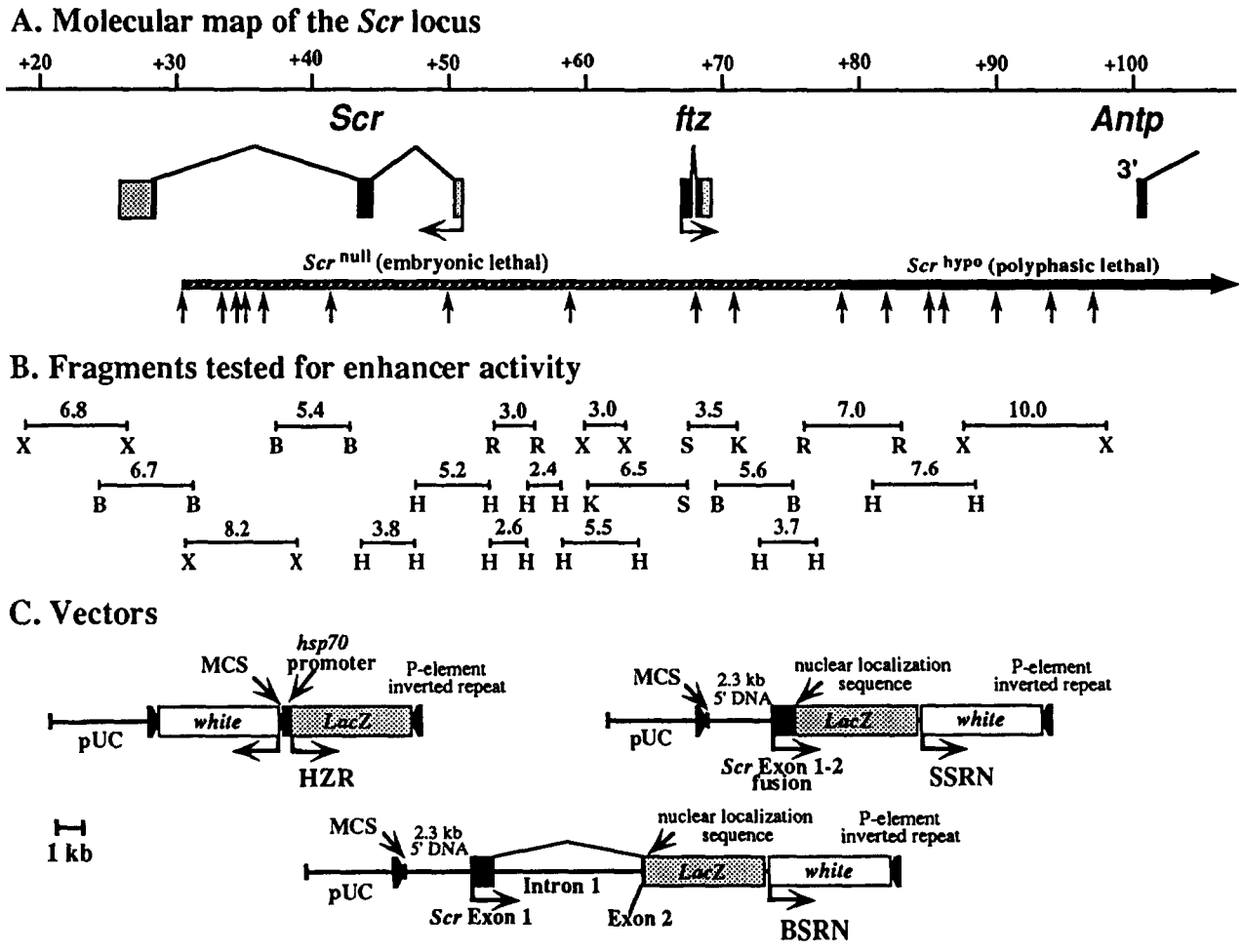

FIGURE 1.—Map of the *Scr* locus, its regulatory region and *P*-element transformation vectors used in this analysis. (A) Molecular map of the *Scr-Antp* interval of the ANT-C. The coordinates at the top of the figure are in kilobases and are numbered according to the chromosome walk of SCOTT *et al.* (1983). Exons are denoted by rectangles, filled portions indicate coding regions and noncoding regions are hatched. The *Scr* transcription unit is shown at left. It is composed of three exons and is transcribed from right to left. The *ftz* transcription unit is located 15 kb 5′ of the *Scr* transcript and is transcribed from left to right. The 3′ end of the *Antp* transcript is shown at right. Vertical arrows below the exon-intron maps are the molecular breakpoints of *Scr* null (Scr^{null}) and *Scr* hypomorphic (Scr^{hypo}) alleles according to PATTATUCCI *et al.* (1991). The uncertainty associated with the location of each arrow is, in most cases, <5 kb in either direction (PATTATUCCI *et al.* 1991). (B) Genomic DNA fragments tested for enhancer activity. The number above each line indicates the size of each fragment in kilobases, and the letters below each fragment indicate the restriction sites that define individual fragments. X, *Xba*I; B, *Bam*HI; H, *Hin*dIII; R, *Eco*RI; K, *Kpn*I; S, *Sal*I. (C) Pictoral representations of the *P*-element transformation vectors used in this analysis, with the exception of HZ50PL, which is described in HIROMI and GEHRING (1987). The restriction fragments shown in B were subcloned into the multiple cloning site (MCS) of one or more of these vectors to test for enhancer activity. For more information, see MATERIALS AND METHODS.

set of genomic DNA fragments previously subcloned into HZR or HZ50PL were subcloned into SSRN, which contains the *Scr* promoter and 2.3 kb of upstream *Scr* DNA. A comparison of the β-gal accumulation patterns directed by each fragment in HZR or HZ50PL and SSRN (Table 1) shows that some tissue-specific enhancers activate transcription of both promoters, whereas other enhancers activate transcription of only the *hsp70-lacZ* promoter or the *Scr-lacZ* promoter. These results demonstrate that *Scr* promoter sequences interact with an overlapping, but not identical, set of enhancers as the *hsp70* promoter. Regulatory elements described as *Scr* enhancers in the following sections activate transcription of both the *Scr-lacZ* and *hsp70-lacZ* promoters or, in some cases, only the *Scr-lacZ* promoter. No regulatory elements were identified that activate *hsp70-lacZ* expression in an *Scr*-like pattern without directing expression of the *Scr-lacZ* fusion in a similar pattern (Table 1).

To determine the expression pattern of the basal reporter constructs, embryos containing either BSR, BSRN or SSRN were stained with antibodies recognizing β-gal. The patterns of BSRN and SSRN expression are shown in Figure 2. SSRN is expressed weakly in the ectoderm of the maxillary and labial segments and in the dorsal ridge. β-gal accumulation is first detected

TABLE 1

Summary of enhancer data

DNA fragment	*hsp70-lacZ* fusion (HZR and HZ50PL)	*Scr-lacZ* fusion (BSR and BSRN)[a]	*Scr-lacZ* fusion (SSRN)
No insert	Basal expression	Maxillary segment-T3 (ectoderm), dorsal ridge, anterior midgut	Weak maxillary-labial expression (ectoderm), *en*-like stripes (ectoderm), dorsal ridge
6.8 kb *Xba*I	Hindgut, anal plate, ventral nerve cord (T2-A8)	Not tested	Hindgut, anal plate
6.7 kb *Bam*HI	Antennal-maxillary complex, head musculature	Not tested	Same as reporter gene
8.2 kb *Xba*I	Dorsal and ventral pharyngeal muscles, subset of body wall muscle	Same as reporter gene	Same as reporter gene
5.4 kb *Bam*HI	Gut mesoderm anterior of normal *Scr* gut expression, CNS	Not tested	Gut mesoderm anterior of normal *Scr* gut expression, CNS
3.8 kb *Hin*dIII	No expression	Not tested[b]	Not tested[b]
5.2 kb *Hin*dIII	No expression	Not tested[b]	Not tested[b]
3.0 kb *Eco*RI	Somatic mesoderm (T1)	Same as reporter gene	Not tested[c]
2.6 kb *Hin*dIII	Somatic mesoderm (T1)	Not tested[d]	Not tested[d]
2.4 kb *Hin*dIII	Antennal-maxillary complex, pharnyx	Not tested	Same as reporter gene
5.5 kb *Hin*dIII	Anterior and posterior midgut (mesoderm), *ftz* stripes (ectoderm), subset of PNS	Anterior and posterior midgut (mesoderm), weak *ftz* stripes	Anterior and posterior midgut (mesoderm), weak *ftz* stripes
3.0 kb *Xba*I	*ftz* stripes (ectoderm)	Not tested	Not tested
0.8 kb *Xba*I/*Hin*dIII	Anterior and posterior midgut (mesoderm)	Not tested	Not tested
6.5 kb *Kpn*I/*Sal*I	Anterior and posterior midgut (mesoderm), *ftz* stripes (ectoderm)	Not tested	Anterior and posterior midgut (mesoderm), weak *ftz* stripes
3.5 kb *Kpn*I/*Sal*I	CNS	Same as reporter gene	Same as reporter gene
5.6 kb *Hin*dIII	Procephalon-T1 (ectoderm), proventriculus, posterior midgut, Filzkorper	Not tested	Procephalon-T1 (ectoderm), proventriculus, posterior midgut, Filzkorper
3.7 kb *Hin*dIII	Procephalon-T1 (ectoderm), proventriculus, posterior midgut, Filzkorper	Procephalon-T3 (ectoderm), proventriculus, posterior midgut, Filzkorper	Procephalon-T1 (ectoderm), proventriculus, posterior midgut, Filzkorper
7.0 kb *Eco*RI	Somatic mesoderm	Maxillary segment-T1 (ectoderm) (BSR expression repressed in T2-T3), salivary gland placode	Labial lobe, salivary gland placode, clypeolabrum
7.6 kb *Hin*dIII	Somatic mesoderm	Not tested	PS1-PS2 ventral ectoderm, salivary gland placode, clypeolabrum
10.0 kb *Xba*I	No expression	Labial-T1 segment (ectoderm) (BSRN expression repressed in maxillary segment and T2-T3)	T1 ectoderm, posterior T2 ectoderm, weak posterior T3 ectoderm (SSRN expression repressed in head)

Accumulation patterns described as "segment a-segment b" denote that all segments between segment a and segment b accumulate β-gal. For example, the accumulation pattern "maxillary segment-T3" signifies that the maxillary, labial, prothoracic, mesothoracic and metathoracic segments have reporter gene product accumulation. Unless otherwise stated, the "no insert" pattern of BSR, BSRN and SSRN is present in transformants containing these constructs.

[a] *Scr* genomic DNA fragments were tested for enhancer activity in SSRN; in some cases, DNA fragments were also subcloned into BSR or BSRN.

[b] The 3.8 *Hin*dIII and 5.2 *Hin*dIII fragments are included in the 8.5-kb *Bam*HI/*Eco*RI fragment in the vector BSRN (Figure 1) and therefore were not tested in SSRN.

[c] The 3.0 *Eco*RI fragment was tested for enhancer activity in BSR, so it was not tested in SSRN.

[d] The 2.6 *Hin*dIII overlaps the 3.0 *Eco*RI fragment by 2.3 kb, and it had the same expression pattern as 3.0 *Eco*RI in HZR, so it was not tested further.

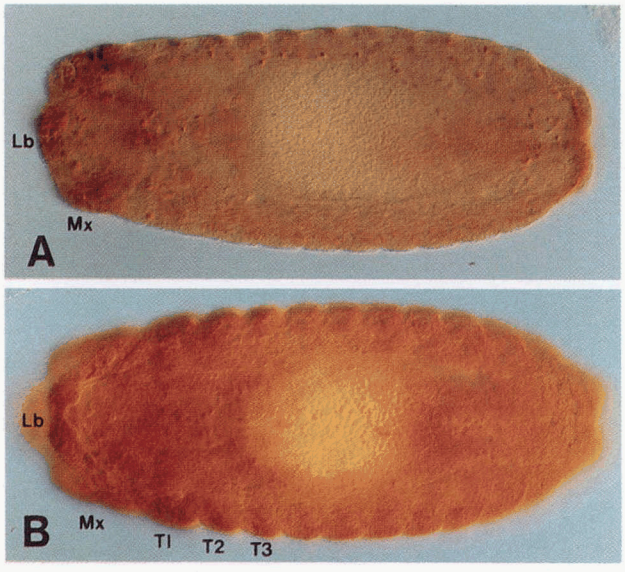

FIGURE 2.—Expression pattern of *Scr-lacZ* fusion vectors SSRN and BSRN. Embryos were stained with an anti-β-gal antibody. Anterior is to the left, and horizontal views of stage 14 embryos are shown. (A) SSRN. β-gal accumulation is observed in the labial segment (Lb) and maxillary (Mx) segments. Additional expression is observed as a narrow band at the posterior edge of each thoracic and abdominal segment. (B) BSRN. Fusion gene expression is observed in a broad domain from the maxillary segment (Mx) to the metathoracic segment (T3).

in the maxillary and labial segments during stage 13. Additional expression is seen in narrow bands, approximately one cell in width, in the posterior compartment of each segment. In this respect, the reporter gene in SSRN appears to be expressed in a similar register to the segment polarity gene *engrailed* (*en*), although the possibility that *en* and the reporter gene are coexpressed was not directly tested. Other investigators have also noted a basal *en*-like accumulation pattern associated with *lacZ* reporter genes in Drosophila (BIENZ *et al.* 1988; BELLEN *et al.* 1989; SIMON *et al.* 1990), suggesting that the *en*-like β-gal accumulation of SSRN is due to cryptic enhancers within *lacZ* and not to *Scr* sequences. Although the SSRN reporter gene appears to contain positive regulatory elements that direct its expression in the labial segment (where *Scr* is normally expressed), it also contains regulatory sequences that direct reporter gene expression outside the *Scr* expression domain.

The *Scr-lacZ* fusion protein in BSRN transformants accumulates at low levels in the maxillary and labial ectoderm, the dorsal ridge and at higher levels in the ectoderm of the three thoracic segments (Figure 2). Additional weak accumulation is seen in the visceral mesoderm of the anterior midgut in cells that normally express *Scr*, as well as cells both anterior and posterior of the normal *Scr* expression domain. In nearly half of the transformant lines tested, there are detectable levels of β-gal in the ectoderm of the abdominal segments. The *Scr* reporter genes in BSR and BSRN are expressed in the same pattern. These results demonstrate that *Scr* sequences in BSR and BSRN reporter genes contain positive regulatory elements sufficient for expression in tissues where *Scr* is normally expressed, such as the ectoderm of the labial and prothoracic segments and the visceral mesoderm of the anterior midgut. In addition, these reporter genes contain positive regulatory elements that activate transcription in the ectoderm both anterior and posterior of the normal *Scr* expression domain. The expression of BSRN differs from SSRN in that BSRN has expression in the ectoderm of the thoracic segments. Because BSRN contains *Scr* intron I, whereas SSRN does not, the regulatory elements responsible for BSRN expression in the thoracic ectoderm appear to be located in *Scr* intron I.

A labial segment enhancer is located 25 kb 5′ of the *Scr* promoter: In wild-type embryos, SCR protein accumulates at high levels in the ventral and lateral ectoderm of the labial segment. SCR is first detected in the labial lobe as it forms during germ band extension, and this accumulation persists throughout embryogenesis (MAHAFFEY and KAUFMAN 1987; LEMOTTE *et al.* 1989; RILEY *et al.* 1987) (Figure 3). A 7.0-kb *Eco*RI fragment normally located 25–32 kb upstream of the *Scr* transcription start site is capable of directing the expression of an *Scr* reporter gene in the labial segment. When the 7.0 *Eco*RI fragment is subcloned 2.3 kb 5′ of the *Scr-lacZ* transcription start site in SSRN, it directs fusion gene expression in the labial segment in an *Scr*-like spatiotemporal pattern (Figure 3). The labial segment accumulation of β-gal in SSRN+7.0RI transformant embryos differs from that observed in SSRN transformants in that regulatory sequences of the 7.0 *Eco*RI fragment activate detectable reporter gene expression in the labial segment much earlier in development than the reporter gene alone. In addition, the level of β-gal accumulation in the labial segment appears to be much higher in SSRN+7.0RI transformants than in SSRN transformants. β-gal accumulation in SSRN+7.0RI embryos is first detected during germ band extension in the lateral labial ectoderm but not in the ventral PS2 ectoderm. At the end of germ band extension (stage 11), the reporter gene is expressed at high levels in the labial lobe, which is located in the lateral ectoderm (Figure 3, A–D). This expression appears to be segmental (Figure 3D, arrows). Additional expression in SSRN+7.0RI transformants is detected in the dorsal ridge (Figure 3B, arrow), in weak engrailed-like stripes at the posterior of each segment (Figure 3B) and in the maxillary segment. Because SSRN is expressed in these tissues, it is difficult to assess the presence of corresponding enhancers in the 7.0 *Eco*RI fragment. These results suggest that the 7.0 *Eco*RI fragment may contain regulatory sequences that direct *Scr* expression in the labial segment, a subset of its normal expression domain.

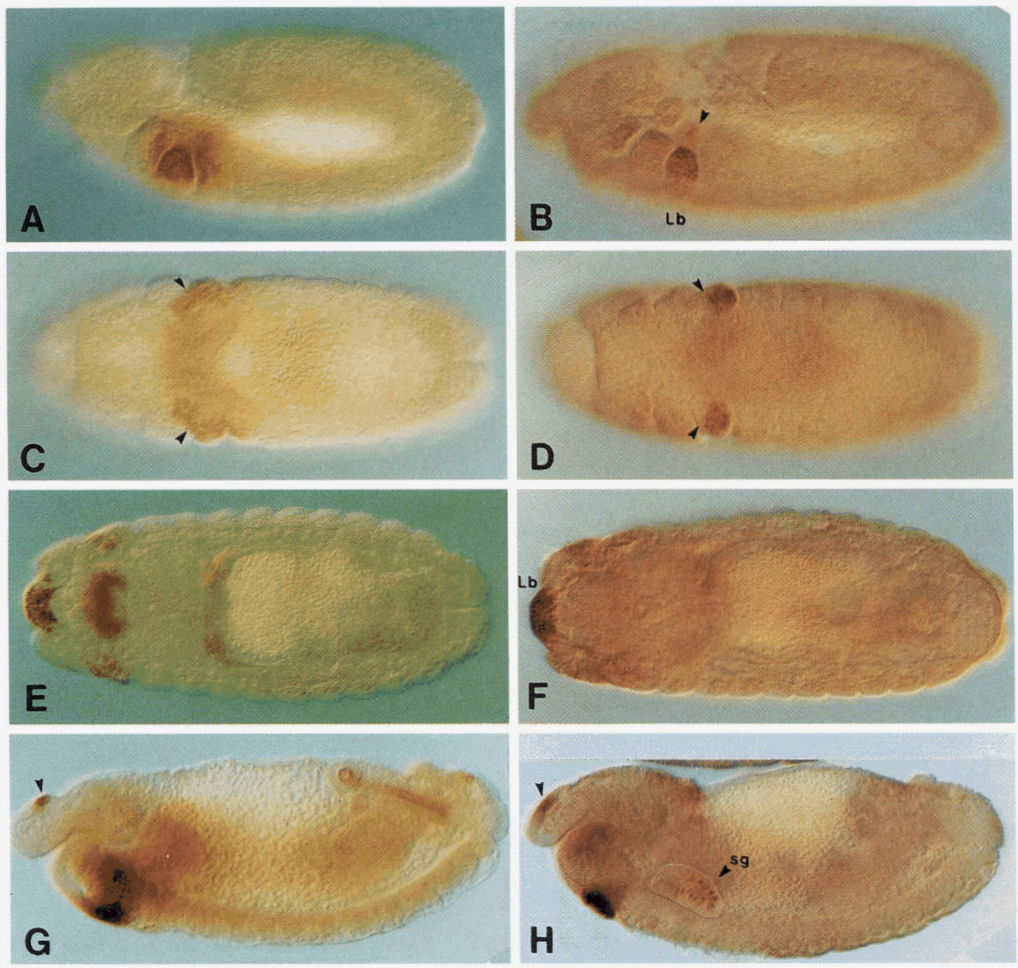

FIGURE 3.—A 7.0-kb *Eco*RI fragment directs *Scr-lacZ* reporter gene expression in the labial segment and other tissues. Embryos were stained with anti-Scr polyclonal antisera (A, C, E, G) or anti-β-gal antibody (B, D, F, H). Anterior is to the left. Embryos in A–D are stage 12. Compare the labial lobe (Lb) expression of *Scr* in a wild-type embryo (A) and β-gal in an SSRN+7.0RI transformant embryo (B) (lateral views). (B) SSRN+7.0RI transformant embryos also have expression in the dorsal ridge (arrow) and in weak *en*-like stripes. (C) *Scr* is expressed in a parasegmental stripe in the ventral part of a wild-type embryo (anterior to arrows). (D) In SSRN+7.0RI transformant embryos, β-gal is not expressed in a ventral parasegmental stripe (arrows are similarly positioned in C and D). Labial lobe expression in SSRN+7.0RI transformants (F) appears to completely overlap the *Scr* domain of wild-type embryos (E) at stage 15 (horizontal view). Embryos in G and H are at stage 14 (sagittal views). (G) Ectopic *Scr* expression is present in a subset of cells in the clypeolabrum (arrow) in *Polycomb* mutant embryos. (H) In SSRN+7.0RI transformant embryos, β-gal accumulation is detected in an overlapping set of cells in the clypeolabrum (arrow). Additional expression is detected in the salivary gland (sg, from a different focal plane of the same embryo).

The 7.0 *Eco*RI fragment also contains regulatory elements that direct *Scr* reporter gene expression not typical of the normal *Scr* expression domain. For example, in SSRN+7.0RI transformants, β-gal accumulates in the salivary glands, which are derived from the ventral PS2 ectoderm (CAMPOS-ORTEGA and HARTENSTEIN 1985). This accumulation is detected in stage 13 embryos, but, in some transformant lines, accumulation is first detected in the salivary gland placode at stage 12. In stage 14 embryos, β-gal is present in the distal part of the salivary gland. SCR protein normally accumulates in the precursor cells of the salivary gland but not in the gland itself (PANZER *et al.* 1992). The accumulation of β-gal in the salivary gland may be due to reporter gene expression outside of the normal *Scr* domain or to the perdurance of β-gal made in salivary gland precursor cells. Ectopic *Scr* reporter gene expression is also observed in a small subset of cells in the clypeolabrum (Figure 3H, arrow). Interestingly, this SSRN+7.0RI expression is quite similar to ectopic SCR accumulation observed in embryos mutant for *Polycomb*, a gene necessary for the maintenance of homeotic gene expression patterns (Figure 3G, arrow). These results demonstrate that the 7.0 *Eco*RI fragment, in addition to containing labial segment enhancers, contains sequences that are capable of directing *Scr-lacZ* expression outside of the normal *Scr* expression domain.

Finally, the 7.0 *Eco*RI fragment also contains sequences that block expression of the *Scr-lacZ* fusion gene present in the vector BSR. Figure 4 compares the *lacZ* expression

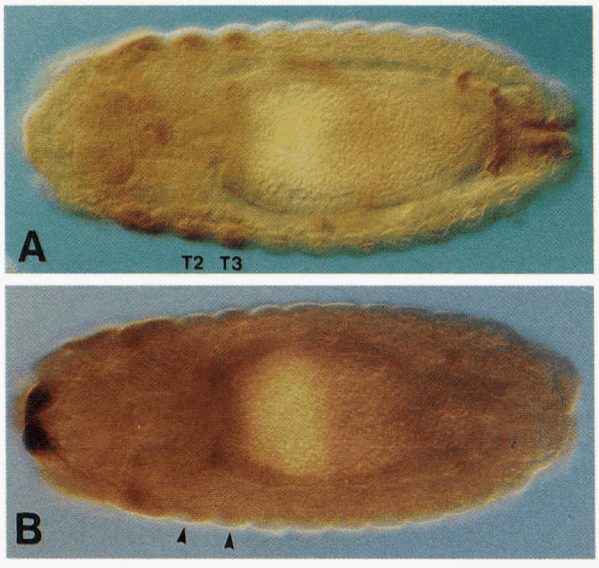

FIGURE 4.–Sequences in the 7.0 *Eco*RI fragment repress BSR reporter gene expression posterior of the prothoracic segment. Embryos were stained with an antibody that recognizes β-gal. Anterior is to the left. (A) *Scr-lacZ* fusion gene expression in BSR transformants. Note that the level of β-gal accumulation in the mesothoracic (T2) and metathoracic (T3) segments are approximately equivalent to levels seen in the prothoracic segment. In BSR+7.0RI transformants (B), the level of β-gal accumulation in the mesothoracic and metathoracic segments (arrows) is reduced relative to the prothoracic segment.

pattern in transformants containing either BSR or BSR+7.0RI. BSR contains *Scr* intron I sequences that direct reporter gene expression in the thoracic ectoderm (Figure 4A). Although there is detectable accumulation of β-gal in the mesothoracic and metathoracic ectoderm of embryos containing BSR+7.0RI, the relative level of expression in these segments is reduced relative to the level of prothoracic accumulation (Figure 4B, arrows). This effect is probably not due to enhanced reporter gene accumulation in the prothoracic segment, because the 7.0 *Eco*RI fragment does not appear to contain any prothoracic enhancers, as assayed by the expression of SSRN+7.0RI (Figure 3). These results suggests that the 7.0 *Eco*RI fragment contains regulatory sequences that partially block transcriptional activation of the *Scr* promoter by regulatory elements located in *Scr* intron I.

An additional repressor element is located in a 10.0-kb *Xba*I fragment normally located 40–50 kb upstream of the *Scr* transcription start site (Figure 1). When subcloned into BSRN, sequences in this fragment are able to block transcriptional activation of the *Scr-lacZ* fusion gene posterior to the prothoracic segment and anterior to the labial segment in transformants containing BSRN+10.0Xb (Table 1) (GINDHART and KAUFMAN 1995). Therefore, at least two repressor elements, located in the 7.0 *Eco*RI and the 10.0 *Xba*I fragments, attenuate the transcriptional activation of an *Scr-lacZ* fusion gene outside of the normal *Scr* expression domain.

Labial and prothoracic segment enhancers are located just distal of *ftz*: Additional *Scr* regulatory sequences were identified in a 3.7-kb *Hin*dIII fragment normally located 3 kb 3′ of the *ftz* transcription unit or 22 kb 5′ of the *Scr* promoter (Figure 1). This fragment is just proximal to the 7.0 *Eco*RI fragment and overlaps the 7.0 *Eco*RI fragment by ~600 base pairs. In SSRN+3.7H, β-gal accumulation is observed in the dorsolateral prothoracic ectoderm at stage 12 (Figure 5B, arrow). Reporter gene expression is seen in only a subset of the prothoracic cells that express *Scr* at this stage (compare to Figure 5A). During germ band retraction, reporter gene product accumulation in the prothoracic segment expands ventrally (Figure 5D) but does not yet fully overlap *Scr* expression in the prothorax (Figure 5C). By stage 14 of embryogenesis, reporter gene accumulation is observed throughout the prothoracic segment (Figure 5F) and appears to be coincident with *Scr* accumulation in this segment (Figure 5E). These results demonstrate that sequences in the 3.7 *Hin*dIII fragment can transcriptionally activate the *Scr* promoter in the prothoracic segment.

The 3.7 *Hin*dIII fragment also contains a labial enhancer, as well as enhancers that direct reporter gene expression outside of the *Scr* domain. By stage 13, SSRN+3.7H transformants accumulate β-gal in the labial and maxillary segments, the procephalon, the proventriculus and the posterior midgut (Figure 5D), although this accumulation is not observed earlier in development (Figure 5B). Like the prothoracic segment, 3.7 *Hin*dIII-directed reporter gene expression is first detected in only a subset of the labial segment (Figure 5D). However, by stage 14 the number of cells in the labial segment expressing the reporter gene increases (Figure 5F), causing β-gal accumulation in the labial segment to more fully resemble that of *Scr* (Figure 5E). Although the 3.7 *Hin*dIII fragment overlaps the 7.0 *Eco*RI fragment by 600 bp, the labial enhancer sequences present in the 3.7 *Hin*dIII fragment are distinct from the labial enhancers located in the 7.0 *Eco*RI fragment because a 5.6-kb *Bam*HI fragment that overlaps the 3.7 *Hin*dIII fragment, but not the 7.0 *Eco*RI fragment, behaves similarly to the 3.7 *Hin*dIII fragment (Table 1).

Because the basal reporter gene SSRN is expressed in the labial and maxillary segments (Figure 2A), it is difficult to determine, by using the SSRN+3.7H construct, if the 3.7 *Hin*dIII fragment has labial and maxillary segment enhancer activity. Therefore, the enhancer activity of 3.7 *Hin*dIII was tested in HZR, which itself has no labial or maxillary segment enhancers. HZR+3.7H transformants show a reporter gene expression pattern similar to SSRN+3.7H in the labial and maxillary segments, as well as in other tissues (Table 1). This result demonstrates that the 3.7 *Hin*dIII fragment contains labial and maxillary segment enhancers.

A 10.0-kb *Xba*I fragment normally located 40–50 kb

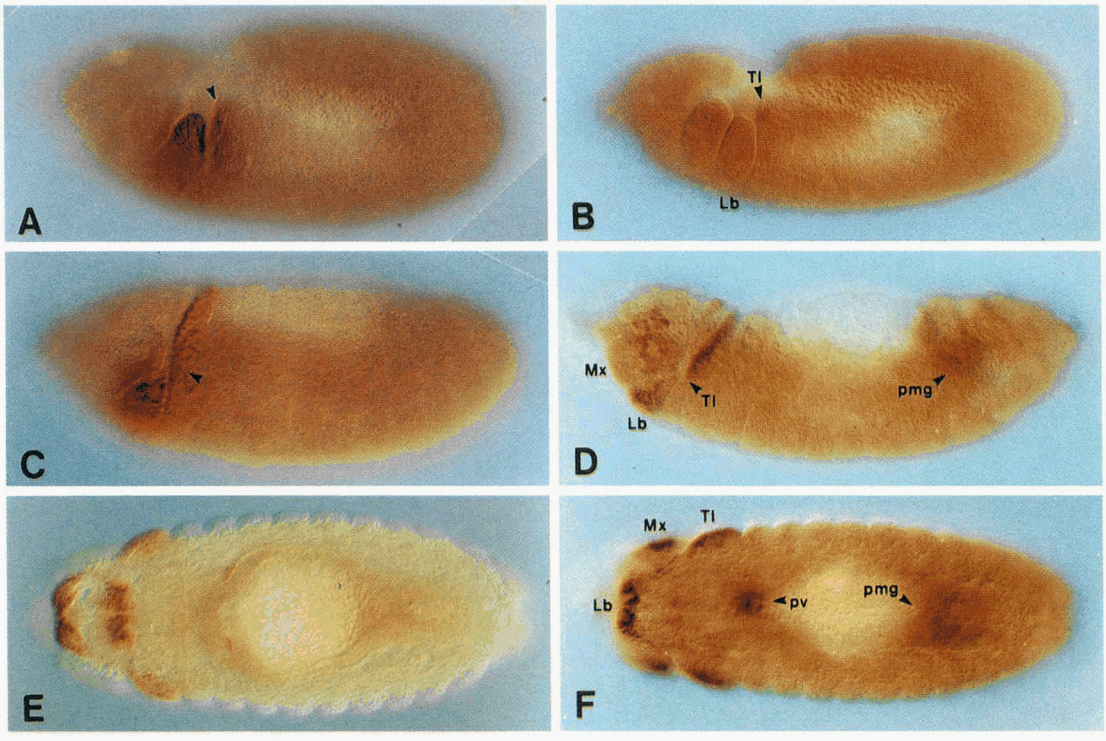

FIGURE 5.—A 3.7-kb *Hin*dIII fragment directs *Scr-lacZ* fusion gene expression in the labial and prothoracic ectoderm. Embryos were stained with a monoclonal antibody (A and C) or polyclonal antisera (E) that recognize Scr protein or an antibody that recognizes β-gal (B, D, F). A–D are sagittal views, whereas E and F are horizontal views. Anterior is to the left. (A and B) At stage 12 of embryogenesis, SSRN+3.7H transformants (B) exhibit reporter gene expression in the dorsolateral part of the prothoracic segment (T1) in a subset of the prothoracic cells that accumulate *Scr* at this stage (A, arrow). No reporter gene expression is detected in the labial segment (Lb) at this stage (B). (C and D) expression pattern of *Scr* in a wild-type embryo (C) and β-gal in a SSRN+3.7H transformant (D) at stage 13 of embryogenesis. Reporter gene expression in the prothoracic segment (T1) has extended more ventrally than at stage 12 (B) but does not yet completely overlap *Scr* accumulation in this segment (arrow in C). Patchy β-gal accumulation is also dectected in the labial (Lb) and maxillary (Mx) segments and in the procephalon. Weaker expression is also detected in the posterior midgut at this stage. (E and F) Scr and β-gal accumulation, respectively, at stage 14 of embryogenesis. SSRN+3.7H reporter gene expression (F) in the prothoracic segment (T1) appears to completely overlap *Scr* expression at this point in development. Reporter gene expression in the labial segment (Lb) is still somewhat patchy but includes a larger proportion of the labial segment than at stage 13 (D). β-gal is also present in the proventriculus (pv) and posterior midgut (pmg), as well as in the maxillary segment (Mx) (out of focal plane).

upstream of the *Scr* transcription start site also contains an enhancer that directs *Scr-lacZ* expression in the prothoracic segment, although its spatiotemporal profile of enhancer activity suggests that it is not actually an *Scr* enhancer (Figure 1; Table 1). β-gal accumulation is first detected late in germ band retraction in the posterior part of T1 and encompasses most, if not all, of the T1 ectoderm by late in embryogenesis (data not shown). However, SCR protein forms an anterior-posterior gradient of accumulation, with highest levels in the anterior part of T1 (Figure 5A), whereas the gradient of SSRN+10.0 *Xba*I in T1 is posterior-anterior (data not shown).

An 0.8-kb fragment contains anterior and posterior mesodermal enhancers: SCR accumulation in the visceral mesoderm of the anterior midgut is first visible during germ band retraction in a band approximately four to five cells in width near the anterior tip of the midgut (TREMML and BIENZ 1989a; REUTER and SCOTT 1990). *Scr* expression in the anterior midgut persists until the end of embryogenesis (LEMOTTE *et al.* 1989; REUTER and SCOTT 1990). Additional Scr protein accumulation is detected during stage 15 of embryogenesis in a cluster of visceral mesoderm cells in the most posterior region of the midgut (LEMOTTE *et al.* 1989). Whereas Scr protein is needed in the anterior midgut to form the gastric caeca (REUTER and SCOTT 1990), the developmental function, if any, of *Scr* in the posterior midgut is unclear.

A DNA fragment normally located in the upstream regulatory region of the segmentation gene *ftz* appears to contain regulatory sequences that direct reporter gene expression in an *Scr*-like pattern in the visceral mesoderm of both the anterior and posterior midgut. This 750-bp 0.8 *Xba*I / *Hin*dIII fragment is a subclone of the 5.5-kb *Hin*dIII fragment located in the interval between *Scr* and *ftz* (Figure 1). The 0.8 *Xba*I / *Hin*dIII fragment is contained in the "dispensable fragment" defined by HIROMI and GEHRING (1987). It is located ~12 kb 5′ of the *Scr* transcription start site. In trans-

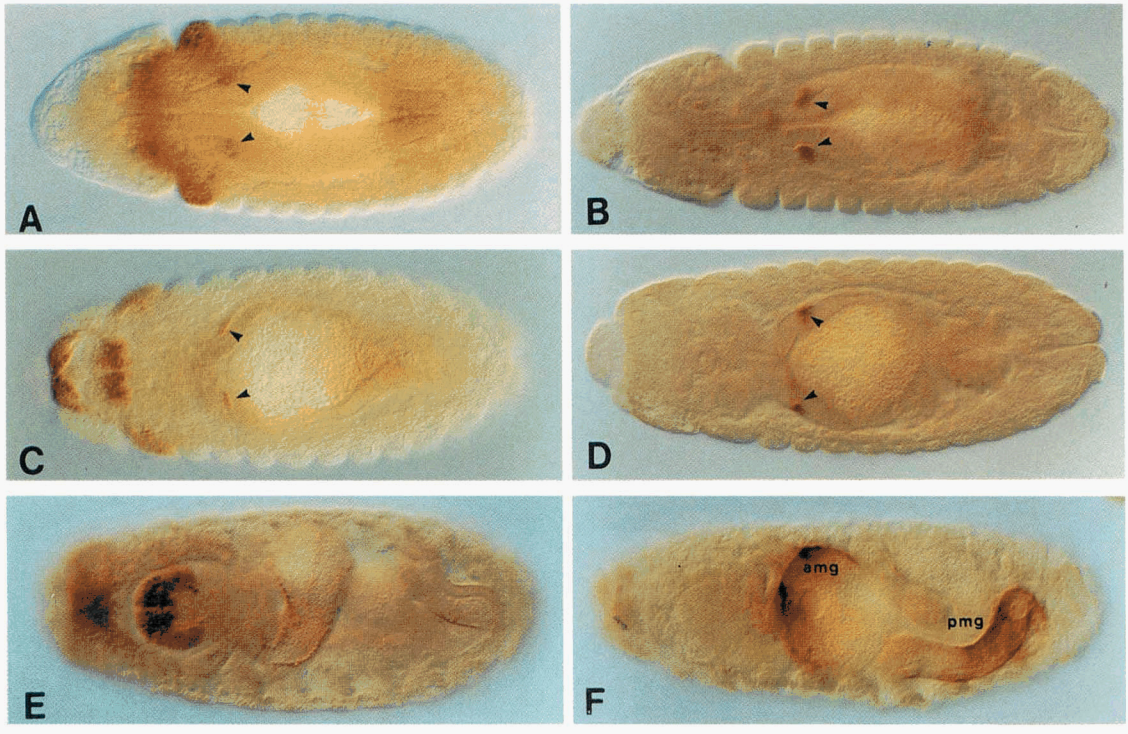

FIGURE 6.—Regulatory sequences in the *ftz* regulatory region direct reporter gene expression in an *Scr*-like pattern in the visceral mesoderm. Embryos were stained with a monoclonal antibody (A and E) or polyclonal antisera (C) that recognizes Scr protein or an antibody that recognizes β-gal (B, D, F). All embryos are horizontal sections and anterior is to the left. (A and B) At stage 13 of embryogenesis, Scr accumulation is clearly visible in the visceral mesoderm of the anterior midgut (A, arrows). In HZR+0.8X/H transformants (B), β-gal accumulation is observed in a similar pattern in the anterior midgut (arrows). (C and D) β-gal accumulation in stage 14 transformants (D) parallels the *Scr* accumulation pattern (C) in the anterior midgut (arrows). At stage 16 of development, HZR+0.8X/H transformants (F) exhibit reporter gene expression in the visceral mesoderm of the anterior (amg) and posterior midgut (pmg). Similar patterns of Scr accumulation (E) are observed in the visceral mesoderm.

formants containing the 0.8 *Xba*I/*Hin*dIII fragment subcloned into HZR, *hsp70-lacZ* fusion gene expression is detected in a pattern indistinguishable from the *Scr* expression pattern in the visceral mesoderm (Figure 6). At stage 13, HZR+0.8X/H transformants exhibit β-gal accumulation in the visceral mesoderm of the anterior midgut (Figure 6B, arrows). Figure 6F shows that the 0.8 *Xba*I/*Hin*dIII fragment also contains regulatory sequences that direct *hsp70-lacZ* expression in the visceral mesoderm of the posterior midgut. Subtle differences in the pattern of β-gal and *Scr* expression in stage 14 (Figure 6D) and stage 16 (Figure 6F) are possibly due to the fact that *hsp70-lacZ* fusion protein accumulates in the cytoplasm, whereas *Scr* protein is localized to the nucleus. The regulatory activity of the 0.8 *Xba*I/*Hin*dIII fragment in SSRN has not been tested, but the 5.5 *Hin*dIII fragment and the 6.5 *Kpn*I/*Sal*I fragments, both of which contain the 0.8 *Xba*I/*Hin*dIII fragment, also direct *Scr-lacZ* expression in the anterior and posterior midgut (Table 1). These results suggest that the 0.8 *Xba*I/*Hin*dIII fragment, which is normally located in the *ftz* regulatory region, contains regulatory sequences that direct *Scr* expression in the visceral mesoderm.

Regulatory sequences activate *Scr-lacZ* transcription outside of the *Scr* expression domain: Whereas some genomic DNA fragments contain sequences that positively regulate the expression of an *Scr-lacZ* fusion gene in an *Scr*-like pattern, other DNA fragments (Table 1) direct β-gal accumulation in tissues where *Scr* is not normally expressed. These effects do not appear to be caused by the influence of sequences near the site of chromosome insertion, as multiple independent transformant lines exhibit similar patterns of reporter gene expression. An example of this type of cryptic regulatory element is a 6.8-kb *Xba*I fragment normally located at the 3′ end of *Scr* (Figure 1). In SSRN+6.8X transformants, β-gal accumulates in the hindgut and anal pads (Table 1). A 5.4-kb *Bam*HI fragment normally located in *Scr* intron II (Figure 1) contains regulatory sequences that activate the transcription of an *Scr-lacZ* fusion gene in the visceral mesoderm at the base of the proventriculus in cells just anterior to the cells that normally accumulate Scr protein.

Another developmentally interesting *Scr-lacZ* expression pattern is obtained when a 7.6-kb *Hin*dIII fragment normally located between *ftz* and *Antp* (Figure 1) is subcloned into SSRN. Transformants containing this

construct exhibit β-gal accumulation in the ventral part of PS1 and PS2, with higher expression levels PS1 than PS2. *Scr* is normally expressed in PS2 but not PS1. The β-gal level decreases during germ band retraction and is no longer detectable at the end of germ band retraction; therefore, it is difficult to determine the fate of these cells in the developing embryo. It is possible that some of these cells are precursors of the subesophageal ganglion, a part of the CNS in which *Scr* is expressed (reviewed in KAUFMAN *et al.* 1990), although further analysis is required to determine this unequivocally.

DISCUSSION

The goal of this analysis was to identify *cis*-regulatory sequences important for transcriptional regulation of the homeotic gene *Sex combs reduced* (*Scr*) during embryogenesis. A straightforward approach consisting of mutant rescue using genomic *Scr* DNA fragments was deterred by the large size of the *Scr* locus. Therefore, a less direct strategy was used in which genomic DNA fragments from the genetically defined *Scr* locus were tested for enhancer activity in *P*-element vectors containing either the *Scr* or *hsp70* promoter fused to the bacterial gene *lacZ*. By observing the pattern of β-gal accumulation in transformant embryos, positive regulatory elements controlling *Scr-lacZ* transcription in the labial and prothoracic ectoderm, as well as the visceral mesoderm of the anterior and posterior midgut, were identified. In addition, putative negative regulatory elements that repress *Scr-lacZ* transcription posterior of the prothoracic segment were identified. Surprisingly, some DNA fragments from the *Scr* regulatory region activate *Scr-lacZ* transcription in non-*Scr*-like patterns in the embryo.

Multiple enhancers control *Scr* expression during embryogenesis: The molecular mapping of *Scr* mutations suggested that *cis*-acting regulatory sequences controlling *Scr* transcription were dispersed throughout a 75-kb interval of ANT-C DNA (Figure 1). The *ftz* gene divides the *Scr* regulatory region into a proximal half that contains the *Scr* transcription unit and 15 kb of upstream DNA and a distal half that contains ~35 kb of regulatory DNA between the *ftz* and *Antp* transcription units. Breakpoints that remove the distal regulatory region are phenotypic nulls that exhibit weak *Scr* expression in an approximately normal pattern (PATTATUCCI and KAUFMAN 1991; M. GORMAN, personal communication). Mutations that remove smaller portions of the distal regulatory region form an allelic series in which the amount of DNA removed is correlated with the severity of the *Scr* mutant phenotype (PATTATUCCI *et al.* 1991). The experiments presented in this paper demonstrate that the loss-of-function phenotype of mutations that remove the distal regulatory region is caused by the removal of *cis*-regulatory sequences present in this region (Figure 7). Enhancers that direct reporter gene expression in an *Scr*-like pattern in the labial segment are located ~22 and 25 kb upstream of the *Scr* promoter, and prothoracic segment enhancer sequences are located ~22 kb upstream of the *Scr* promoter. Interestingly, this group of *Scr* enhancers is located only 7–12 kb away from the *ftz* promoter, yet they do not appear to contribute to the *ftz* transcription pattern and are dispensible for *ftz* function (HIROMI *et al.* 1985).

The weak but detectable *Scr* accumulation pattern of mutants in which the distal regulatory region is removed predicts that *Scr* regulatory sequences are also located proximal of *ftz*. The expression patterns of SSRN and BSRN, which contain sequences normally located near the *Scr* promoter, suggest that labial segment enhancers are located in the 2.3 kb 5′ of the *Scr* transcription start site and that thoracic ectoderm and anterior midgut mesoderm enhancers are located in *Scr* intron I. These results demonstrate that the spatiotemporal regulation of *Scr* expression is controlled by multiple redundant regulatory elements, located both relatively close to the *Scr* promoter and tens of kilobases away. In addition, these results suggest that weak *Scr* expression observed in breakpoint mutations that remove the distal regulatory region may be controlled by regulatory sequences present in SSRN and BSRN.

The distal regulatory region also contains sequences that repress reporter gene expression outside of the *Scr* expression domain. Two DNA fragments, a 7.0-kb *Eco*RI fragment normally located 25 kb 5′ of the *Scr* promoter and a 10.0-kb *Xba*I fragment normally 38 kb 5′ of the *Scr* promoter, block transcriptional activation of the *Scr-lacZ* fusions in BSR and BSRN, respectively. Sequences within the 7.0 *Eco*RI fragment repress *Scr-lacZ* fusion protein accumulation in the mesothoracic and metathoracic segments, whereas the 10.0 *Xba*I fragment prevents transcriptional activation by BSRN reporter gene sequences anterior of the labial segment, as well as posterior of the prothoracic segment (GINDHART and KAUFMAN 1995). This suggests that the function of regulatory sequences contained within these fragments is to block *Scr* transcription outside of its normal expression domain. However, the lack of ectopic *Scr* expression in embryos containing mutations that remove the distal regulatory region suggests that unidentified negative regulatory elements are present in the proximal regulatory region and that these negative regulatory elements may, as demonstrated for the positive regulatory elements, be functionally redundant.

The localization of *Scr* regulatory elements throughout its regulatory interval is reminiscent of the distribution of enhancers of the homeotic gene *Ubx* (BIENZ *et al.* 1988; SIMON *et al.* 1990; IRVINE *et al.* 1991; MÜLLER and BIENZ 1991). By placing subsets of the *Ubx* regulatory region upstream of *Ubx-lacZ* reporter genes, it has been shown that enhancers controlling both the establishment of *Ubx* expression by gap and segmentation

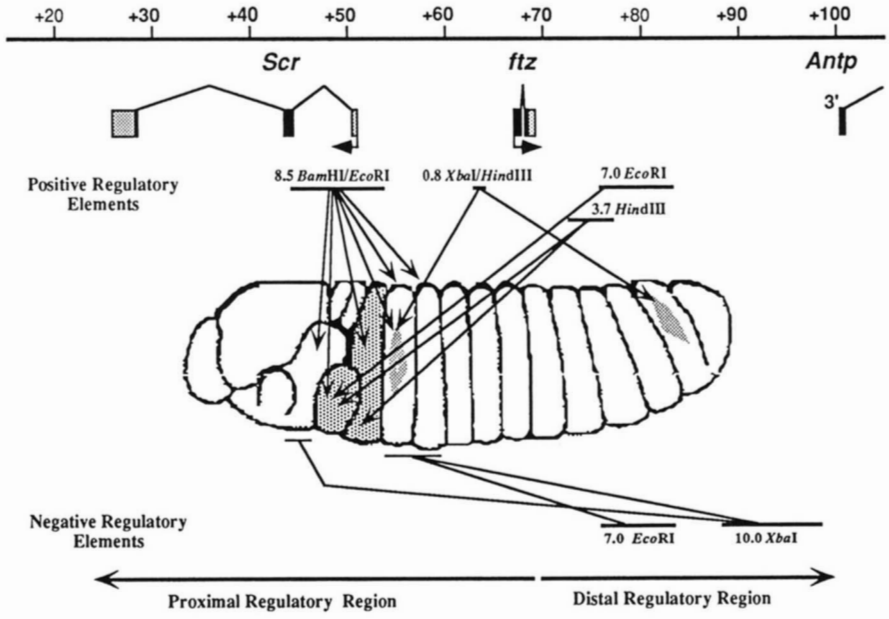

FIGURE 7.—Organization of positive and negative *Scr* regulatory sequences in the *Scr-Antp* interval of the ANT-C. The coordinate map and exon-intron structure of transcripts in the *Scr* regulatory region are shown at top. A representation of a germ band retraction stage embryo is shown in the center of the figure. Tissues that accumulate *Scr* at this stage are stippled and include the labial and prothoracic segments, as well as the visceral mesoderm of the anterior and posterior midgut. *Scr* expression in the subesophageal ganglion is not indicated, as regulatory sequences controlling this aspect of *Scr* expression have not been identified. The location of DNA fragments containing regulatory sequences that direct reporter gene expression in a subset of the *Scr* expression domain are shown above the embryo. DNA fragments containing regulatory sequences that block transcriptional activation by sequences in the 8.5 *Bam*HI/*Eco*RI fragment are shown below the embryo. Arrows indicate the tissues in which reporter genes containing these fragments accumulate β-gal. Lines ending with a bar indicate the segments in which negative regulatory sequences block transcriptional activation. At the bottom of the figure, the arbitrarily defined limits of the proximal and distal *Scr* regulatory regions are shown.

gene products, as well as regulatory elements that maintain these established patterns in a *Polycomb* Group-dependent manner, are distributed over 80 kb of genomic DNA, both 5′ and 3′ of the *Ubx* promoter. Similarly, we have shown that *Scr* enhancer elements are distributed over ≥60 kb of regulatory DNA (Figure 7) (GINDHART and KAUFMAN 1995). Previous analyses have implicated several gene products as being important for the establishment of *Scr* transcription in its normal pattern. These include the gap genes *hunchback* (*hb*), *giant* (*gt*) and *Krüppel* (*Kr*); the segmentation genes *fushi tarazu* (*ftz*), *even-skipped* (*eve*) and *odd-paired* (*opa*); the homeotic genes *Antennapedia* (*Antp*) and *teashirt* (*tsh*) and the TGF-β homologue *decapentaplegic* (INGHAM and MARTINEZ-ARIAS 1986; RILEY *et al.* 1987; TREMML and BIENZ 1989b; PANGANIBAN *et al.* 1990; FASANO *et al.* 1991). It is likely that the *Scr* regulatory elements identified in this analysis contain binding sites for *trans*-acting factors that establish the characteristic pattern of *Scr* expression.

***Scr* enhancers are present in the *ftz* upstream regulatory region:** An intriguing result of this analysis is the identification of *Scr* regulatory sequences in the *ftz* regulatory region. A 750-bp fragment contained within the "dispensable fragment" defined by HIROMI and GEHRING (1987) directs reporter gene expression in the visceral mesoderm of the anterior and posterior midgut in an *Scr*-like pattern. This result demonstrates that the *ftz* and *Scr* regulatory regions, instead of merely being located near each other, are actually intermingled. This result leads to a number of questions. How did an *Scr* enhancer come to reside in the *ftz* regulatory region? Is the presence of *Scr* enhancers in the *ftz* regulatory region evolutionarily conserved? MAIER *et al.* (1990) showed that although the size and spacing of *ftz* regulatory sequences of the *ftz* homologue in *D. hydei*, a species that diverged from *D. melanogaster* 60 million years ago (BEVERLY and WILSON 1984), is conserved, the DNA fragment containing *Scr* midgut enhancers is not located in the *D. hydei ftz* regulatory region. However, the location of this DNA fragment does appear to be conserved in *D. pseudoobscura*, a species more closely related to *D. melanogaster* (GINDHART 1993; MAIER *et al.* 1993). These results lead to a hypothesis in which a DNA fragment containing the *Scr* midgut enhancers entered the *ftz* regulatory region through a nearby transposition event and that this event occurred after the divergence of *D. melanogaster* and *D. hydei*. Interestingly, MAIER *et al.* (1990) also demonstrated that the *ftz* gene and its regulatory sequences are inverted in *D. hydei* with respect to *Scr* and *Antp*. In *D. hydei*, the labial and prothoracic enhancers located in the 3.7

*Hin*dIII fragment are also inverted, causing them to be located <10 kb 5′ of the *D. hydei Scr* promoter. However, *Scr* regulatory sequences distal of the 3.7 *Hin*dIII fragment do not appear to be inverted in *D. hydei* (MAIER *et al.* 1990). Although the division of the *Scr* regulatory region by *ftz* appears to be conserved in Drosophilids, these results show that the linear arrangement of *Scr* enhancer sequences has been altered since the divergence of *D. hydei* and *D. melanogaster*.

Cryptic enhancers reside in the *Scr* regulatory region: Another surprising result of this analysis was the identification of sequences that direct the expression of an *Scr-lacZ* fusion gene in non-*Scr*-like patterns. These cryptic enhancers can be divided into two broad classes: enhancers that activate *Scr-lacZ* transcription in patterns resembling *Scr* expression in *trans*-acting factor mutant backgrounds and those that have no apparent relationship to *Scr* expression. An example of the former class is the 7.0 *Eco*RI fragment, which directs *Scr-lacZ* expression in the clypeolabrum, a pattern that resembles ectopic *Scr* accumulation observed in embryos lacking *Polycomb* function. This result suggests that the clypeolabrum enhancer in the 7.0 *Eco*RI fragment may be negatively regulated, either directly or indirectly, by *Polycomb*. An example of the latter class is the 6.8 *Xba*I fragment, which contains enhancer sequences that transcriptionally activate an *Scr-lacZ* fusion gene in the hindgut and anal pads. *Scr* expression is limited, for the most part, to parasegment 2- and 3-derived structures, whereas the hindgut and anal pads originate from cells located posterior of parasegment 15 (CAMPOS-ORTEGA and HARTENSTEIN 1985).

A possible reason that some of these fragments have cryptic enhancer activity is that in the transgenic constructs they are located closer to the promoter compared with their position in the normal chromosome. Some cryptic enhancers, such as the hindgut and anal pads enhancer in the 6.8 *Xba*I fragment, are normally tens of kilobases away from the *Scr* promoter. When placed 2.3 kb 5′ of the *Scr* promoter in SSRN, this enhancer activates *Scr-lacZ* transcription in these tissues, whereas, in the resident *Scr* locus, 25 kb of intervening DNA may prevent efficient enhancer-promoter interactions. If a DNA sequence cannot interact with a promoter due to topological constraints and thus plays no role in normal gene expression, this sequence would be freed from selective constraints and allowed to diverge. If this is the case, one would predict that a corresponding fragment from a related species might not contain a hindgut and anal pad enhancer sequence but may contain no or different tissue-specific enhancers.

Another possible explanation for the effect of cryptic enhancers on *Scr-lacZ* fusion gene expression is that negative regulatory elements that block transcriptional activation of the *Scr* promoter by cryptic enhancers are absent from SSRN and BSRN. An example of this type of cryptic enhancer may be the clypeolabrum enhancer in the 7.0 *Eco*RI fragment. Because its tissue-specific enhancer activity resembles ectopic *Scr* expression observed in *Polycomb* mutant embryos, it is possible that regulatory sequences located outside of the 7.0 *Eco*RI fragment normally block, in a *Polycomb*-dependent manner, the ability of the clypeolabrum enhancer to activate *Scr* transcription.

One of the unanticipated outcomes of this analysis is the discovery that the *Scr* and *hsp70* promoters interact with overlapping yet distinct sets of enhancers. Some enhancers direct the expression of *Scr-lacZ* and *hsp70-lacZ* reporter genes in similar tissue-specific patterns, whereas other enhancers interact with only one promoter. Two possible models can be proposed to explain this result. First, heterogeneity of the binding sites of basal transcription factors may result in differences in transcription preinitiation complex assembly at each promoter, resulting in differences in the ability of tissue-specific transcription factors to stimulate the rate of transcription (PTASHNE 1988; STRUHL 1991). For example, SIMON *et al.* (1988) demonstrated that altering the TATA box of the human *hsp70* promoter resulted in the abolishing of its transcriptional activation by adenovirus E1A protein but not by heat shock inducible transcription factors. In this model, differences in the core promoter sequences of *Scr* and *hsp70* are responsible for promoter-enhancer specificity. The second model proposes that sequences in the 2.3 kb of DNA upstream of the *Scr-lacZ* fusion gene promoter either block or augment the ability of tissue-specific enhancers to interact with the *Scr-lacZ* promoter. An example of this model is context-dependent activity of the tissue-specific transcription factor encoded by *dorsal*. DOR protein can act as a transcriptional activator or repressor, depending upon sequences upstream of the promoters DOR regulates (JIANG *et al.* 1992). This model of promoter-enhancer specificity proposes that enhancers in the 2.3 kb upstream of the *Scr-lacz* promoter can block cryptic enhancer activity and/or interact cooperatively with *Scr* enhancers normally located tens of kilobases from the *Scr* promoter. These models can also be used to explain how the *ftz* gene, which is located in the middle of the *Scr* regulatory region (Figure 7), insulates itself from the influence of *Scr* enhancers, as well as cryptic enhancer sequences.

We thank MARIE MAZZULLA for providing polyclonal SCR antisera and generating a subset of the *P*-element constructs used in this analysis. We thank DAVID MILLER and ZAMIN YANG for providing plasmids and DANNY BROWER for providing the 6H4 cell line. We thank the Bloomington Stock Center for providing mutant strains. We are grateful to MAUREEN GORMAN, CARL BAUER, THOMAS BLUMENTHAL, MARC MUSKAVITCH and WILLIAM SAXTON for critical reading of the manuscript. We also thank SCOTT CHOUINARD, MAUREEN GORMAN, ANN KAPOUN, DAVID MILLER and ANGELA PATTATUCCI for invigorating discussions and useful suggestions. Finally, we thank DEE VEROSTKO for her administrative assistantship and good humor. This work was supported by National Institutes of Health (NIH) grant GM-24299 to T.C.K., NIH Predoctoral Fellowship GM-07757 to J.G.G. and the

Howard Hughes Medical Institute. T.C.K. is Investigator of the Howard Hughes Medical Institute

LITERATURE CITED

AKAM, M., 1987 The molecular basis for metameric pattern in the *Drosophila* embryo. Development **101:** 1–22.

BELLEN, H. J., C. J. O'KANE, C. WILSON, U. GROSSNIKLAUS, R. KURTH-PEARSON *et al.*, 1989 P-element-mediated enhancer detection: a versatile method to study development in *Drosophila*. Genes Dev. **3:** 1288–1300.

BEVERLY, S. M., and A. C. WILSON, 1984 Molecular evolution in *Drosophila* and the higher Diptera. II. A time scale for fly evolution. J. Mol. Evol. **21:** 1–13.

BIENZ, M., and G. TREMML, 1988 Domains of *Ultrabithorax* expression in *Drosophila* visceral mesoderm from autoregulation and exclusion. Nature **333:** 576–578.

BIENZ, M., G. SAARI, G. TREMML, J. MÜLLER, B. ZÜST *et al.*, 1988 Differential regulation of *Ultrabithorax* in two germ levels of *Drosophila*. Cell **53:** 567–576.

CAMPOS-ORTEGA, J. A., and V. HARTENSTEIN, 1985 *Embryonic Development of Drosophila melanogaster*. Springer-Verlag, Berlin.

CARROLL, S. B., S. DINARDO, P. H. O'FARRELL, R. A. H. WHITE and M. P. SCOTT, 1988 Temporal and spatial relationships between segmentation and homeotic gene expression in *Drosophila* embryos: distributions of the *fushi tarazu*, *engrailed*, *Sex combs reduced*, *Antennapedia*, and *Ultrabithorax* proteins. Genes Dev. **2:** 350–360.

CHOUINARD, S., and T. C. KAUFMAN, 1991 Control of expression of the homeotic *labial* (*lab*) locus of *Drosophila melanogaster*: evidence for both positive and negative autogenous regulation. Development **113:** 1267–1280.

DUNCAN, I., 1987 The *bithorax* complex. Annu. Rev. Genet. **21:** 285–319.

FASANO, L., L. RÖDER, N. CORÉ, E. ALEXANDRE, C. VOLA *et al.*, 1991 The *teashirt* gene is required for the development of *Drosophila* embryonic trunk segments and encodes a protein with widely spaced zinc finger motifs. Cell **64:** 63–79.

GARCIA-BELLIDO, A., 1977 Homeotic and atavistic mutations in insects. Am. Zool. **17:** 613–629.

GINDHART, J. G., 1993 Developmental regulation of the Drosophila homeotic gene *Sex combs reduced*. Ph.D. Thesis, Indiana University, Bloomington.

GINDHART, J. G., and T. C. KAUFMAN, 1995 Identification of *Polycomb* and *trithorax* group responsive elements in the regulatory region of the Drosophila homeotic gene *Sex combs reduced*. Genetics **139:** 797–814.

GLICKSMAN, M. A., and D. L. BROWER, 1988 Expression of the *Sex combs reduced* protein in *Drosophila* larvae. Dev. Biol. **127:** 113–118.

HAFEN, E., M. LEVINE and W. J. GEHRING, 1984 Regulation of *Antennapedia* transcript distribution by the *bithorax* complex in *Drosophila*. Nature **307:** 287–289.

HIROMI, Y., and W. J. GEHRING, 1987 Regulation and function of the *Drosophila* segmentation gene *fushi tarazu*. Cell **50:** 963–74.

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

INGHAM, P., 1988 The molecular genetics of embryonic pattern formation in *Drosophila*. Nature **335:** 25–34.

INGHAM, P. W., and A. MARTINEZ-ARIAS, 1986 The correct activation of *Antennapedia* and *bithorax* complex genes requires the *fushi tarazu* gene. Nature **324:** 592–597.

IRVINE, K. D., S. L. HELFAND and D. S. HOGNESS, 1991 The large upstream control region of the *Drosophila* homeotic gene *Ultrabithorax* (*Ubx*). Development **111:** 407–424.

JIANG, J., C. A. RUSHLOW, Q. ZHOU, S. SMALL and M. LEVINE, 1992 Individual *dorsal* morphogen binding sites mediate activation and repression in the *Drosophila* embryo. EMBO J. **11:** 3147–3154.

KALDERON, D., B. L. ROBERTS, W. D. RICHARDSON and A. E. SMITH, 1984 A short amino acid sequence able to specify nuclear location. Cell **39:** 499–509.

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

KAUFMAN, T. C., M. A. SEEGER and G. OLSON, 1990 Molecular and genetic organization of the *Antennapedia* gene complex of *Drosophila melanogaster*. Adv. Genet. **27:** 309–362.

KENNISON, J. A., 1993 Transcriptional activation of *Drosophila* homeotic genes from distant regulatory elements. Trends Genet. **9:** 75–78.

KUROIWA, A., U. KLOTER, P. BAUMGARTNER and W. J. GEHRING, 1985 Cloning of the homeotic *Sex combs reduced* gene in *Drosophila* and in situ localization of its transcripts. EMBO J. **4:** 3757–3764.

KUZIORA, M. A., and W. MCGINNIS, 1988 Autoregulation of a *Drosophila* homeotic selector gene. Cell **55:** 477–85.

LEMOTTE, P. K., A. KUROIWA, L. I. FESSLER and W. J. GEHRING, 1989 The homeotic gene *Sex combs reduced* of *Drosophila*: gene structure and embryonic expression. EMBO J. **8:** 219–227.

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

MAHAFFEY, J. W., and T. C. KAUFMAN, 1987 The distribution of the *Sex combs reduced* gene product during development of *Drosophila melanogaster*. Genetics **117:** 51–60.

MAHAFFEY, J. W., R. J. DIEDERICH and T. C. KAUFMAN, 1989 Novel patterns of homeotic protein accumulation in the head of the *Drosophila* embryo. Development **105:** 167–174.

MAIER, D., A. PREISS and J. R. POWELL, 1990 Regulation of the segmentation gene *fushi tarazu* has been functionally conserved in *Drosophila*. EMBO J. **9:** 3957–3966.

MAIER, D., D. SPERLICH and J. R. POWELL, 1993 Conservation and change of the developmentally crucial *fushi tarazu* gene in *Drosophila*. J. Mol. Evol. **36:** 315–326.

MANN, R. S., and D. S. HOGNESS, 1990 Functional dissection of *Ultrabithorax* proteins in *D. melanogaster*. Cell **60:** 597–610.

MÜLLER, J., and M. BIENZ, 1991 Long range repression conferring boundaries of *Ultrabithorax* expression in the *Drosophila* embryo. EMBO J. **10:** 3147–3155.

OUWENEEL, W. J., 1976 Developmental genetics of homeosis. Adv. Genet. **18:** 179–248.

PANGANIBAN, G. E. F., R. REUTER, M. P. SCOTT and F. M. HOFFMAN, 1990 A *Drosophila* growth factor homolog, *decapentaplegic*, regulates homeotic gene expression within and across germ layers during midgut morphogenesis. Development **110:** 1041–1050.

PANZER, S., D. WEIGEL and S. K. BECKENDORF, 1992 Organogenesis in *Drosophila melanogaster* embryonic salivary gland determination is controlled by homeotic and dorsoventral patterning genes. Development **114:** 49–57.

PARO, R., 1990 Imprinting a determined state into the chromatin of *Drosophila*. Trends Genet. **6:** 416–421.

PATTATUCCI, A. M., and T. C. KAUFMAN, 1991 The homeotic gene *Sex combs reduced* of *Drosophila melanogaster* is differentially regulated in the embryonic and imaginal stages of development. Genetics **129:** 443–461.

PATTATUCCI, A. M., D. C. OTTESON and T. C. KAUFMAN, 1991 A functional and structural analysis of the *Sex combs reduced* locus of *Drosophila melanogaster*. Genetics **129:** 423–441.

PIRROTTA, V., 1988 Vectors for P-mediated transformation in *Drosophila*. Biotechnology **10:** 437–456.

PTASHNE, M.,1988 How eukaryotic transcriptional activators work. Nature **335:** 683–689.

RANDAZZO, F., 1991 Evolution of homeotic gene complexes and the rescue and regulation of proboscipedia: a homeotic gene of the Antennapedia complex in *Drosophila melanogaster*. Ph.D. Thesis, Indiana University, Bloomington.

REUTER, R., and M. P. SCOTT, 1990 Expression and function of the homeotic genes *Antennapedia* and *Sex combs reduced* in the embryonic midgut of *Drosophila*. Development **109:** 289–303.

RIDDIHOUGH, G., and D. ISH-HOROWICZ, 1991 Individual stripe regulatory elements in the *Drosophila hairy* promoter respond to maternal, gap, and pair-rule genes. Genes Dev. **5:** 840–854.

RILEY, P. D., S. B. CARROLL and M. P. SCOTT, 1987 The expression and regulation of *Sex combs reduced* protein in *Drosophila* embryos. Genes Dev. **1:** 716–730.

ROBERTSON, H. M., C. R. PRESTON, R. W. PHILLIS, D. M. JOHNSON-SCHLITZ, W. K. BENZ *et al.*, 1988 A stable genomic source of *P*-element transposase in *Drosophila melanogaster*. Genetics **118:** 461–470.

SCOTT, M. P., and S. B. CARROLL, 1987 The segmentation and homeotic gene network in early *Drosophila* development. Cell **51:** 689–698.

SCOTT, M. P., WEINER, A. J., HAZELRIGG, T. I., POLISKY, B. A., PIRROTTA

et al., 1983 The molecular organization of the *Antennapedia* locus of *Drosophila*. Cell **35:** 763–776.

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

SIMON, J., M. PEIFER, W. BENDER and M. O'CONNOR, 1990 Regulatory elements of the *bithorax* complex that control expression along the anterior-posterior axis. EMBO J. **9:** 3945–3956.

SIMON, M. C., T. M. FISCH, B. J. BENECKE, J. R. NEVINS and N. HEINTZ, 1988 Definition of mutiple, functionally distinct TATA elements, one of which is a target in the *hsp70* promoter for E1A regulation. Cell **52:** 723–729.

STRUHL, K., 1991 Mechanisms for diversity in gene expression patterns. Neuron **7:** 177–181.

STRUHL, G., and R. A. H. WHITE, 1985 Regulation of the *Ultrabithorax* gene of *Drosophila* by other *bithorax* complex genes. Cell **43:** 507–519.

THUMMEL, C. S., A. M. BOULET and H. D. LIPSHITZ, 1988 Vectors for P-element mediated transformation and tissue culture transfection. Gene **74:** 445–456.

TREMML, G., and M. BIENZ, 1989a Homeotic gene expression in the visceral mesoderm of *Drosophila* embryos. EMBO J. **8:** 2677–2685.

TREMML, G., and M. BIENZ, 1989b An essential role of *even-skipped* for homeotic gene expression in the *Drosophila* visceral mesoderm. EMBO J. **8:** 2687–2693.

WAKIMOTO, B. T., and T. C. KAUFMAN, 1981 Analysis of larval segmentation in lethal genotypes associated with the *Antennapedia* gene complex in *Drosophila melanogaster*. Dev. Biol. **81:** 51–64.

WEFALD, F. C., B. H. DEVLIN and R. S. WILLIAMS, 1990 Functional heterogeneity of mammalian TATA-box sequences revealed by interaction with a cell-specific enhancer. Nature **344:** 260–262.

Communicating editor: R. E. DENELL