The EMBO Journal vol.7 no.8 pp.2559-2567, 1988

Functional redundancy in the tissue-specific enhancer of the *Drosophila Sgs-4* gene

Thomas A.Jongens, Timothy Fowler, Antony W.Shermoen and Steven K.Beckendorf

Department of Molecular Biology, University of California, Berkeley, CA 94720, USA

Communicated by F.Kafatos

During the last day of larval development, the *Sgs-4* glue gene of *Drosophila melanogaster* is expressed at high levels in a single tissue, the larval salivary glands. As shown by transformation experiments and by DNA sequence analysis of *Sgs-4* underproducing strains, an essential regulatory region for *Sgs-4* expression lies between 149 and 568 bp upstream from the transcribed part of the gene. This region shows the positional independence of a transcriptional enhancer and directs at least three regulatory activities: tissue specificity, developmental timing and high-level expression. Here we use a transient transformation assay to identify three elements within this enhancer that are involved in tissue specificity. For at least this regulatory activity the enhancer is internally redundant. Any pairwise combination of the three elements is sufficient to direct salivary gland expression, although none of the three can act alone.

Key words: salivary glands/glue gene/transient transformation

Introduction

During the development of multicellular eukaryotes, gene expression occurs in specific and often elaborate spatial and temporal patterns. Because such patterns depend on cellular interactions and because they change as development proceeds, their regulation is best studied in the intact organism. In *Drosophila*, DNA-mediated transformation methods have been developed to allow analysis of this tissue- and stage-specific gene expression. Genes that have been modified *in vitro* can be integrated into the germline via P element vectors and their activity monitored in subsequent generations (Rubin and Spradling, 1982). Alternatively, transient expression of injected DNA can be monitored in the somatic tissues of the injected animals themselves (Steller and Pirotta, 1984; Martin *et al.*, 1986). Using these methods we have begun to dissect the regulation of the *Drosophila* glue gene, *Sgs-4* (McNabb and Beckendorf, 1986; Shermoen *et al.*, 1987).

Sgs-4 is one of seven glue genes that, in response to the steroid hormone ecdysone, are abundantly expressed in the salivary glands of third instar larvae (Korge, 1975; Beckendorf and Kafatos, 1976). The products of these genes are small glycoproteins that are secreted into the lumen of the salivary gland a few hours before the end of third instar and are then expelled, as the prepupa forms, to attach the pupal case to the underlying substrate (Fraenkel and Brookes, 1953). The glue genes are co-ordinately expressed and are distributed among five sites on the X, second and third chromosomes. *Sgs-4*, the only one on the X chromosome, is located within an intermolt puff at 3C11-12 (Korge, 1975) and produces an approximately 1-kb, unspliced RNA (Muskavitch and Hogness, 1980).

Several lines of evidence have identified an essential regulatory region upstream from *Sgs-4*. When the gene is active, three DNase I-hypersensitive sites are found in salivary gland nuclei at 480, 405 and 330 bp upstream from the *Sgs-4* cap site (Shermoen and Beckendorf, 1982). Deletion of parts of this hypersensitive region in natural *Sgs-4* underproducers or in transformed lines, eliminates or dramatically reduces expression (Muskavitch and Hogness, 1982: McGinnis *et al.*, 1983; McNabb and Beckendorf, 1986). Further transformation experiments showed that a 420-bp fragment which included the hypersensitive region could act as an enhancer element. When placed adjacent to the *Drosophila Adh* gene, this fragment increased *Adh* expression in tissues such as the anterior midgut and fat body which normally express *Adh*, and it directed novel *Adh* expression in salivary glands during late third instar. All three regulatory activities—enhancement, tissue specificity and developmental timing—showed the positional flexibility of enhancer elements (Shermoen *et al.*, 1987).

To locate the functional regions within this enhancer, we have prepared a number of constructs which contain truncated or rearranged copies of the previously tested fragment. To decrease the time required for their analysis, we have employed a transient transformation assay developed by Martin *et al.* (1986). They showed that constructs containing the *Drosophila Adh* gene could be qualitatively

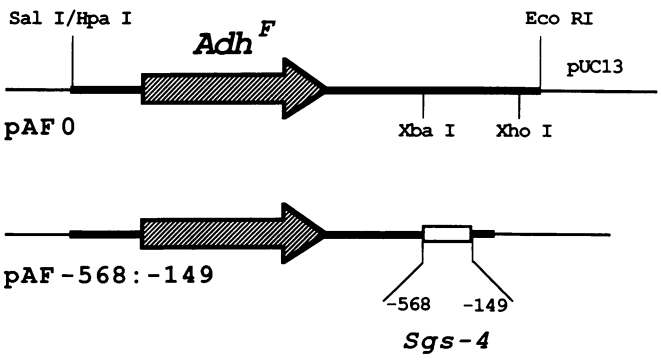

Fig. 1. Organization of the pAF 0 and pAF −568:−149 plasmids. pAF 0 includes the *HpaI/EcoRI* fragment of Adh^F, a fragment that includes the proximal but not the distal promoter for *Adh* (Posakony *et al.*, 1985). This fragment is inserted into the polylinker of pUC13. The pAF −568:−149 plasmid is similar but the *XbaI/XhoI* fragment 3′ to *Adh* has been replaced with the −568 to −149 *DdeI* fragment of *Sgs-4*. This fragment was previously shown to contain *Sgs-4* enhancer functions (Shermoen *et al.*, 1987).

©IRL Press Limited, Oxford, England

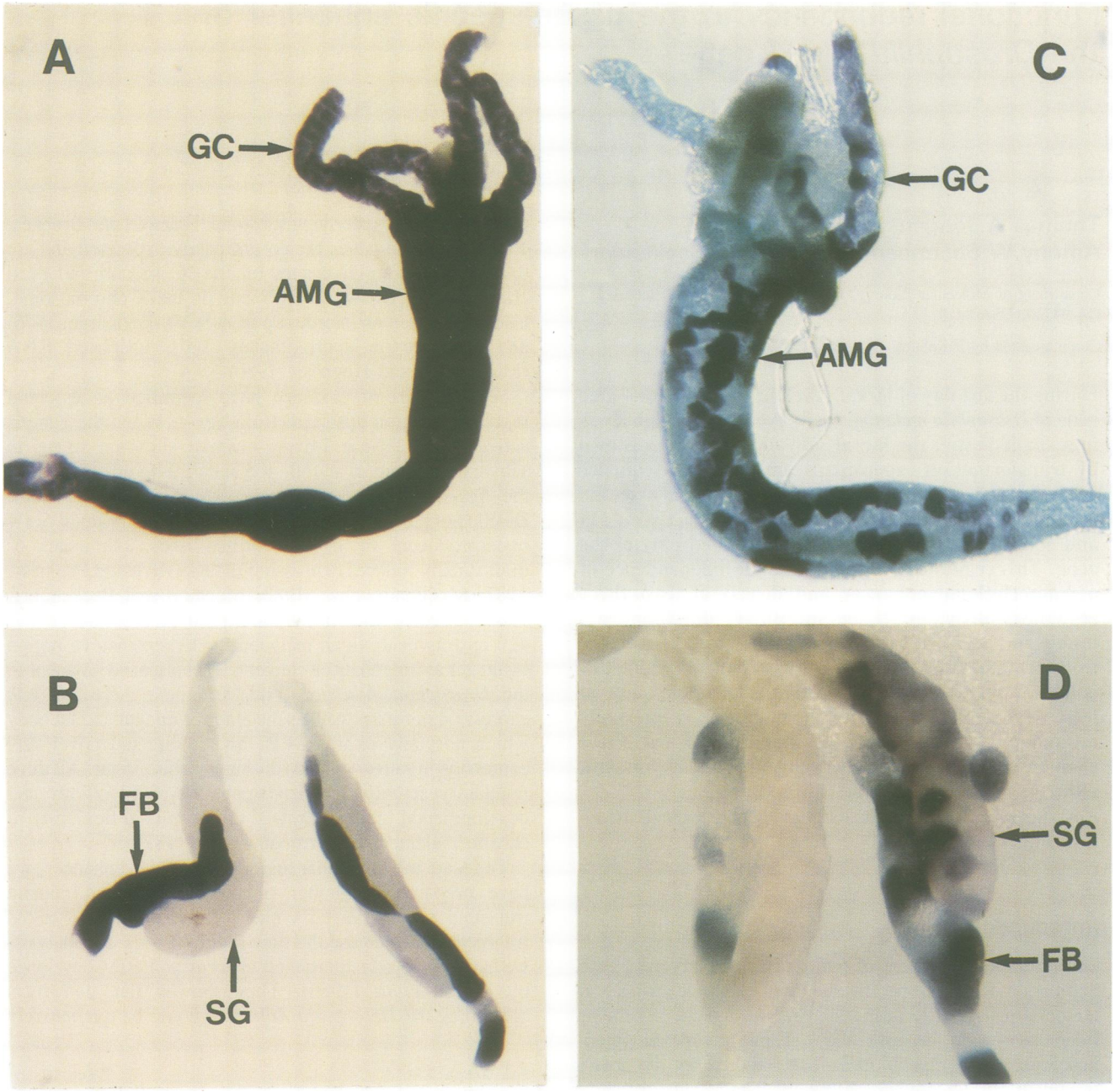

Fig. 2. Pattern of *Adh* activity in injected and control larvae. (**A**) Staining pattern of the anterior mid gut (AMG) and gastric caecae (GC), and (**B**) salivary glands (SG) and fat body (FB) from wild-type Oregon R-B larvae. There was no staining in tissues from the ACR strain which carries the Adh^{fn6} null allele (results not shown). ACR was used as the recipient in the transient assay. (**C,D**) Examples of stained anterior midgut and salivary glands from ACR larvae which were injected as embryos with pAF −568:−149.

tested for expression following injection into preblastoderm embryos of an *Adh* null strain. When cells formed at blastoderm, the injected DNA was incorporated into some of them. After development to an appropriate stage, the injected animals were dissected and stained for *Adh* activity. Since the injected animals, rather than their progeny, were analyzed, this transient assay dramatically reduced the time required to evaluate any particular construct.

We have now used this assay to define sequences required for the tissue specificity imparted by the *Sgs-4* enhancer. The results identify three separable elements, any two of which can combine to specify salivary gland expression.

Results

Tests of the transient assay

The parental plasmid for the transient analysis was pAF −568:−149 (Figure 1) which contains all of the previously identified *Sgs-4* enhancer region (Shermoen *et al.*, 1987). The *Sgs-4* sequence in this plasmid and in all of its derivatives was placed 3′ to the *Adh* gene, about 1.7 kb downstream from the proximal *Adh* promoter. The 3′ location of the *Sgs-4* sequence was chosen so that only long-range, enhancer-like effects would be measured, rather than any cryptic promoter activity of the *Sgs-4* sequences. A similar

Table I. Expression of modified *Sgs-4* enhancer constructs in larval tissues

Construct	Tissues						Number
	SG	AMG	GC	Gut	FB	MT	
pAF 0	0	65	41	5	5	0	19
pAF −568:−149	17	71	71	17	12	8	24
pAF −515:−149	27	36	44	0	12	0	59
pAF −498:−149	21	71	58	0	21	0	24
pAF −464:−149	0	74	67	11	16	0	70
pAF −568:−274	52	33	24	5	14	0	21
pAF −568:−316	21	71	50	7	14	7	14
pAF −568:−371	16	68	52	5	20	0	19
pAF −568:−397	30	81	67	2	37	0	54
pAF −568:−440	8	61	50	0	15	0	26
pAF −568:−473	0	81	58	3	12	0	74
pAF −498:−440	0	83	66	4	13	0	103
pAF −498:−392	0	70	50	3	6	0	34
pAF −498:−274	18	79	59	0	18	0	39
pAF −515:−440	0	84	79	0	18	0	61
pAF −568:−149 (ΔII)	9	85	81	4	17	0	53
pAF −568:−274 (ΔII)	10	78	52	3	16	2	63
pAF −437:−149/−568:−498	0	88	85	0	10	0	41
Average	22[a]	70	58	5	15	0	

Each construct was tested for its ability to direct *Adh* expression in Adh^{fn6} larvae. Following histochemical staining of dissected third instar larvae, tissue was scored as positive if between one cell and all the cells stained. For any construct that was expressed in salivary glands, at least one of the larvae had >30% of its salivary glands stained. The results are expressed as per cent of the number of larvae examined in which the specified tissue was positive. The structures of the constructs are shown in Figure 4. The last column indicates the number of larvae examined for each construct. SG, salivary glands; AMG, anterior mid gut; GC, gastric caecae; Gut, remainder of mid and hind gut; FB, fat body; MT, malpighian tubules. Infrequent staining of the malpighian tubules results from the anterior injection site rather than an inability of the *Adh* sequences to direct expression in this tissue (see Figure 3).

[a]Average for those constructs that expressed *Adh* in their salivary glands.

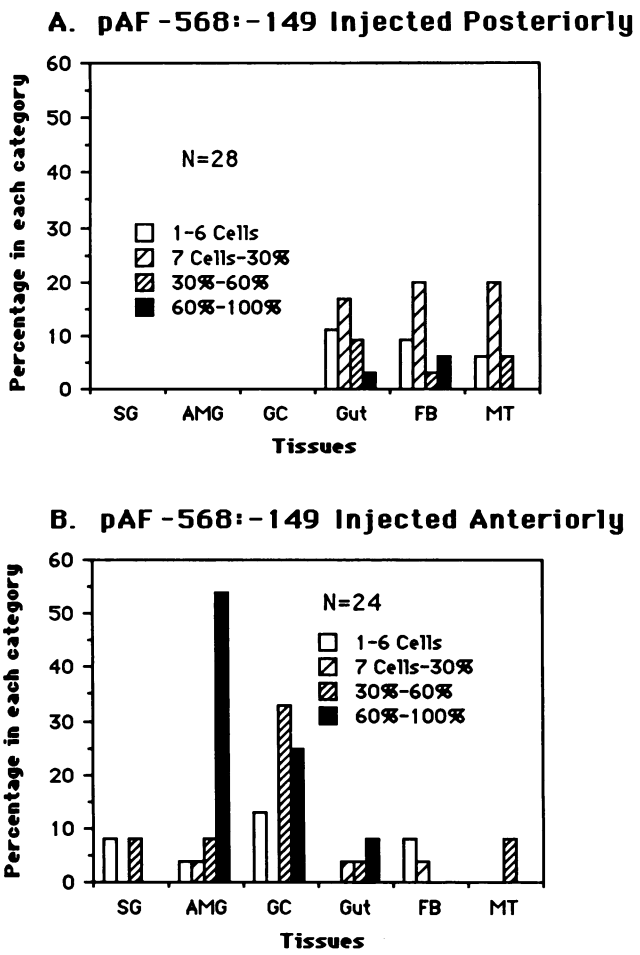

Fig. 3. The distribution of stained tissues in individuals injected anteriorally or posteriorally with pAF −568:−149. Plasmid DNAs were injected into either the anterior or posterior ends of preblastoderm embryos, and the resultant third instar larvae histochemically stained as described in Materials and methods. Because staining occurred in mosaic patches, the amount of staining in each tissue was scored in four categories: 1−6 cells, 7 cells−30% of the tissue, 30−60% or 60−100%. The frequency of each category is indicated by the height of the bars. N, total number of larvae stained; SG, salivary glands; AMG, anterior mid gut; GC, gastric caecae; Gut, rest of the mid and hind gut; FB, fat body; MT, malpighian tubules.

construct had been shown to direct salivary gland expression of *Adh* in stable transformants (Shermoen *et al.*, 1987).

To test whether the tissue specificity of the *Sgs-4* enhancer would be observed in the transient assay, we injected *Adh* null embryos either with this pAF −568:−149 plasmid or with the control pAF 0 plasmid which is identical except that it lacks the *Sgs-4* sequence (Figure 1). About 20% of the larvae which developed from embryos injected with pAF −568:−149 expressed *Adh* in their salivary glands (Figures 2 and 3). As expected from previous results with stably transformed flies (Shermoen *et al.*, 1987), *Adh* staining was also detected in the tissues which express *Adh* in wild-type larvae: the anterior midgut and its associated gastric caecae, parts of the mid- and hindgut, Malpighian tubules and fat body (Ursprung *et al.*, 1970). Unregulated *Adh* expression or expression in unexpected tissues was never detected. Instead, salivary gland expression directed by the *Sgs-4* sequence is simply added to the normal pattern of *Adh* expression directed by the sequences near the *Adh* proximal promoter (Posakony *et al.*, 1985). In contrast, larvae resulting from pAF 0 injections stained for *Adh* activity only in the normal *Adh* tissues. None of these larvae had any *Adh* expression in salivary glands (Table I).

In these and all subsequent transient assays, *Adh* expression occurred in mosaic patterns, apparently reflecting unequal distribution of the injected DNA among the cells of the larvae (Figure 2). The patterns varied from just a few or no staining cells in a particular tissue to large patches that might include the entire tissue (Figure 3). This observation, originally made by Martin *et al.* (1986), made it necessary to screen a large number of injected individuals before concluding that a particular construct was unable to direct salivary gland expression. Presence of the active *Adh* regulatory sequences, however, provided a valuable internal control. For any particular injected animal, expression of *Adh* in the normal *Adh* tissues indicated that the DNA had been successfully introduced.

As Martin *et al.* (1986) had noted, the spectrum of tissues which express *Adh* depends greatly on the site of injection (Figure 3). The salivary glands, as well as the anteriorly derived tissues which normally express *Adh*, stained for *Adh*

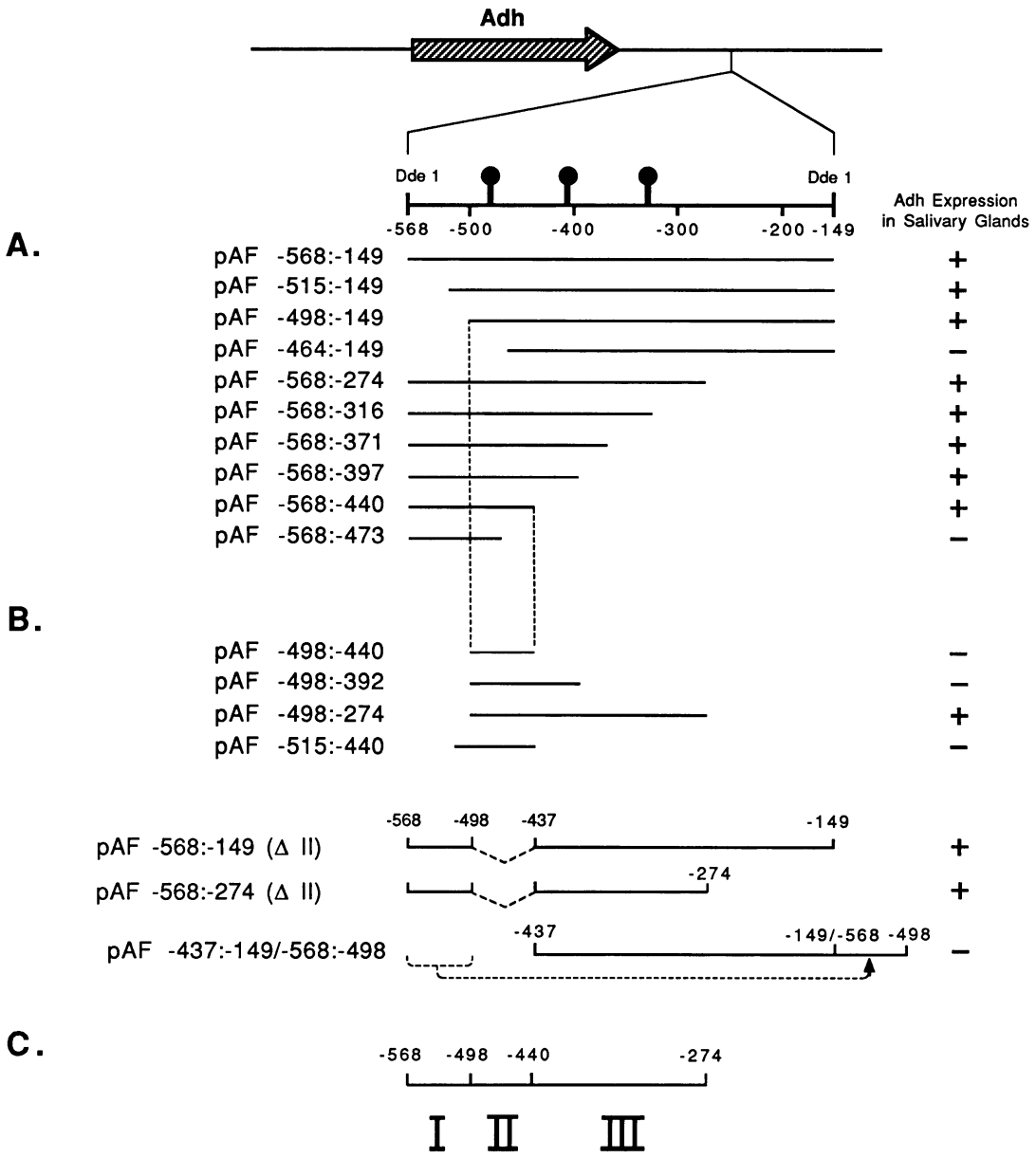

Fig. 4. Truncations and rearrangements of the *Sgs-4* enhancer used to locate its functional elements. Diagram at the top shows the arrangement of pAF −568:−149 with the *Sgs-4* sequences expanded. Below are indicated the *Sgs-4* sequences remaining in plasmids carrying deletions or more complicated rearrangements of the enhancer sequence. Construction of these plasmids is detailed in Materials and methods. (**A**) Plasmids derived from 5′ or 3′ deletion of the *Sgs-4* DNA by *Bal*31. Endpoints of the deletions were determined by sequencing and are indicated in the names of the plasmids. (**B**) Doubly truncated plasmids and those with more complex rearrangements. Dotted lines in the two internally deleted plasmids connect sequences that are normally separated but are contiguous in these plasmids. Each plasmid was tested for its ability to direct salivary gland expression of *Adh*. The results are shown at the right and in more detail in Table I. (**C**) The three functional, tissue-specifying elements deduced from these experiments (see text for explanation).

activity more frequently when DNA was injected into the anterior end of the embryo. Likewise, the posteriorly derived *Adh*-expressing tissues stained more frequently following posterior injections. Therefore, all of the remaining constructs we tested were injected into the anterior pole of the embryo.

From the pAF −568:−149 and pAF 0 injections we concluded that the transient assay would be useful for analyzing deletions of the *Sgs-4* enhancer region. The assay is relatively rapid and faithfully reproduces the tissue specificities of the injected DNA.

External deletions of the Sgs-4 enhancer region

Next, we used the transient assay to determine which of the *Sgs-4* sequences in pAF −568:−149 specified salivary gland expression. Two series of deletions were constructed, one from the 5′ end and the other from the 3′ end of the 420-bp *Sgs-4* enhancer region (for details see Materials and methods). Selected deletions were subcloned into pAF 0 at the same location relative to the *Adh* promoter as the entire *Sgs-4* fragment in pAF −568:−149. The resulting plasmids were then tested in the transient assay.

Tests of the 5′ deletions revealed that the sequences from −568 to −498 can be removed without any effect on salivary gland tissue specificity (Figure 4A). However, when another 34 bp were removed (pAF −464:−149) no salivary gland expression of *Adh* was detected. Table I shows that a similar percentage of non-salivary gland tissues stained

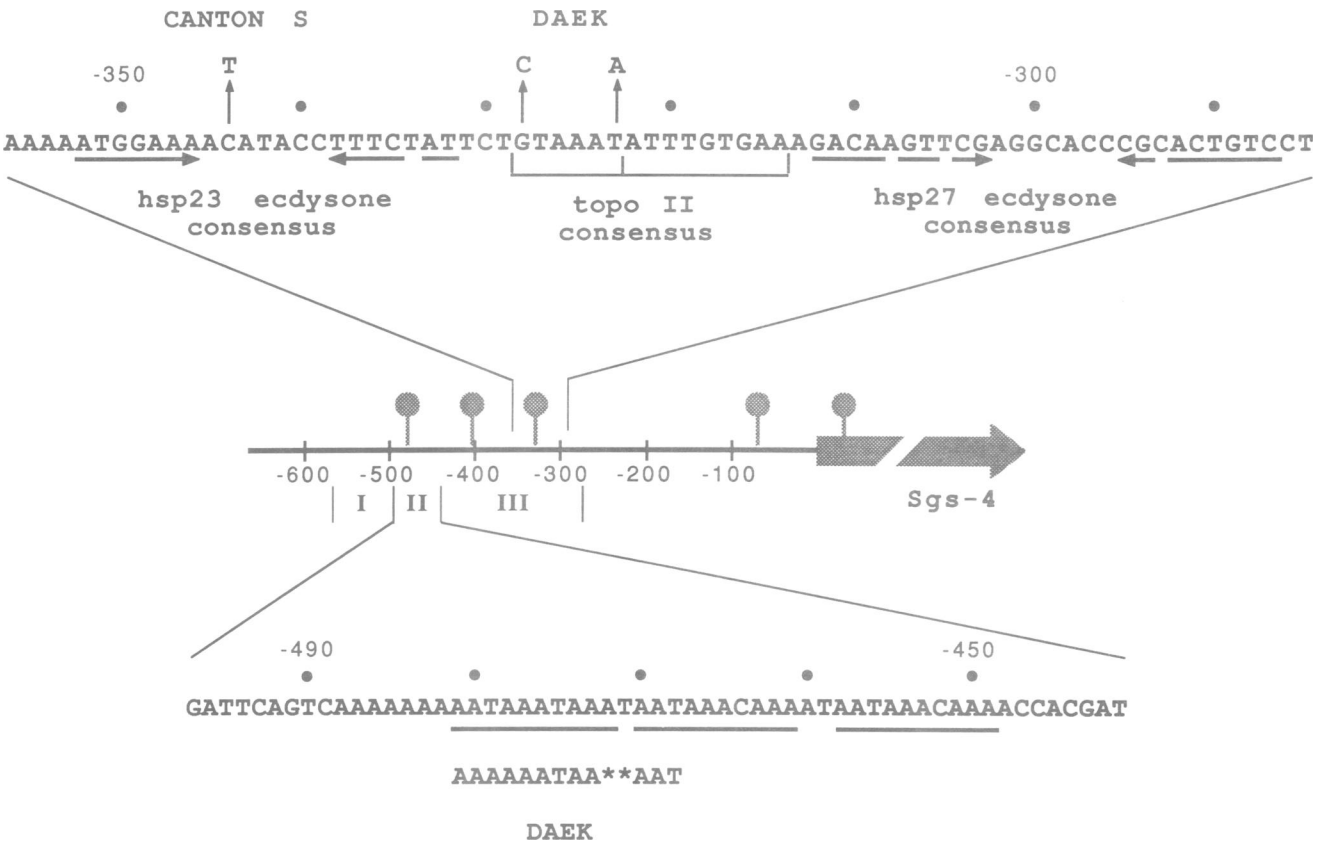

Fig. 5. Consensus sequences and mutations in element II and part of element III. The sequence of element II is indicated in the lower part of the figure, and the sequence of part of element III is indicated near the top. A repeating sequence in element II is underlined and below this region is indicated the corresponding sequence from the underproducing strain Daek. Missing nucleotides are shown as *. Below the partial element III sequence are shown three consensus sequences discussed in the text. The vertical line near the center of the topoisomerase II consensus indicates the cleavage site. Nucleotide substitutions present in the Canton S and Daek strains are indicated at the top.

following injection with this pAF −464:−149 construct as with those constructs that retained salivary gland expression. These non-salivary gland tissues serve as a control to indicate that the tested larvae had received the DNA and that it contained an intact *Adh* gene. The similar frequency of staining for these tissues, along with the large number of pAF −464:−149 larvae tested, strongly suggest that this construct is incapable of specifying salivary gland expression.

From the 3′ end it was possible to remove all the sequence from −149 to −440 without affecting salivary gland tissue specificity (Figure 4A). However, in embryos injected with the construct pAF −568:−440 the frequency of salivary gland staining was low: only 8% (Table I). This result may reflect experimental variation or may indicate that the pAF −568:−440 construct expresses *Adh* in salivary glands at a lower level than the other constructs. We plan to test this possibility by making stable transformants containing this construct.

Individuals injected with the next 3′ deletion, pAF −568:−473 did not express *Adh* in salivary glands, yet the frequencies of staining in other tissues were near their respective averages and a large number of larvae were stained. Therefore, we are confident that this construct does not contain a functional salivary gland tissue specifier.

Multiple tissue-specifying elements

Comparison of the results from the 5′ and 3′ deletion series shows that sequences from −498 to −440 are present in all of the functional 5′ or 3′ deletions, suggesting that this region might be able to act alone as a salivary gland tissue specifier. This simple hypothesis was tested with a construct containing just the −498:−440 region inserted into pAF 0 at the same location as the 5′ and 3′ deletions. To our surprise this pAF −498:−440 construct was incapable of directing salivary gland expression (Figure 4B and Table I). This result indicates not only that the −498:−440 region is insufficient for salivary gland expression, but also that two other functional regions exist in the enhancer, one on either side of the −498:−440 region. This conclusion comes from the salivary gland expression of constructs pAF −498:−149 and pAF −568:−440, which contain only the sequences from −498 to −440 in common. Since this common region is insufficient, each of these constructs must contain an additional element capable of complementing the −498 to −440 sequence to form a functional salivary gland tissue specifier. To simplify further discussion of these elements, we will refer to the region from −568 to −498 as element I and −498 to −440 as element II (Figure 4c). At this stage of the analysis a relatively large region, −440 to −149, was defined as element III.

To map the functional part of element III more precisely, we added increasing amounts of element III sequence to element II and tested for tissue-specifying activity. When the sequence was extended to −274, activity appeared (pAF −498:−274, Figure 4B), demonstrating that element III is contained between −440 and −274 (Figure 4C).

In a similar attempt to localize the functional part of element I, we tested a construct containing the proximal part

Table II. Sequence motif found near salivary gland genes

Gene	Location	Sequence	Reference
Sgs-4	−480	T A A A T A A A	Muskavitch and Hogness
	−469	T A A A C A A A	(1982)
	−457	T A A A C A A A	
	−219	T A A A C A A A	
Pig-1	−61	T A A A T A A A	B.Rogers (personal
	−273	T A A A C A A A	communication)
Sgs-7/8	−169/−303	T A A A T A A A	Garfinkel *et al.* (1983)
	−220/−352	T A A A T A A A	
Sgs-3	−537	T A A A T A A T	Garfinkel *et al.* (1983)
	−566	T A A A T A T A	
	−580	T A A A T A A G	
	−611	T A A A T T A A	
Sgs-5	−73	T C A A T T A A	Shore and Guild (1986)

Table III. Comparison of *Sgs-4* sequences with two *hsp* ecdysone-responsive sequences

hsp 23	−192	ATCTGCCAT	−19−	ATGGAAAAT	−228
Sgs-4	−352	ATGGAAA	− 7−	TTTCTAT	−332
hsp 27	−553	GACAAGGGTTC	− 3−	GCACTTGTC	−527
Sgs-4	−312	GACAAG TTC	− 9−	GCAC TGTC	−287

The *hsp 23* sequence and its putative *Sgs-4* counterpart are both inverted repeats, but one is oriented head to head and the other tail to tail. Thus the left half of the *Sgs-4* sequence should be compared with the right half of the *hsp 23* sequence. [References: *hsp 23*, Mestril *et al.* (1986); **hsp 27**, Riddihough and Pelham (1987).]

of element I along with element II. This construct, pAF −515:−440, however, was not capable of directing salivary gland *Adh* expression (Figure 4B and Table I). Thus the functional part of element I probably includes the more distal sequences from −568 to −515.

Since all of the active constructs contained element II, we then made two internally deleted constructs to test whether this element must be present for salivary gland expression [pAF −568:−149 (ΔII) and pAF −568:−274 (ΔII), Figure 4B]. Despite the elimination of the entire element II,both constructs were able to specify salivary gland expression. These results show that the sequences between −498 and −440 are, in fact, not necessary for activity. In addition, they show that elements I and III can co-operate to specify salivary gland expression. In fact, the three elements now appear to be functionally interchangeable. Any pairwise combination of elements produces an active salivary gland tissue specifier.

In these internally deleted constructs, elements I and III are 61 bp closer to each other than they are normally. The fact that these constructs are active suggests that the exact spacing between elements I and III is not important. It might be that the two elements work independently of each other. To test this hypothesis, we constructed a plasmid which interchanged the positions of elements I and III (pAF −437:−149/−568:−498, Figure 4B). It contains the same sequences as the internally deleted plasmid, pAF −568:−149 (ΔII) described above, but position −149 is now adjacent to −498 (see Figure 4B). This construct was unable to confer salivary expression on the *Adh* gene, suggesting that elements I and III do not act independently. Their arrangement relative to each other is important for function.

Discussion

Elements in the Sgs-4 enhancer

The most important result of this work is that the *Sgs-4* enhancer contains three separable elements that are involved in salivary gland tissue specificity. As tested by the transient transformation assay, combination of any two of the three elements gives expression that is indistinguishable from the initial 420-bp fragment, yet none of the separable elements is by itself sufficient for salivary gland expression. These results imply that there must be functionally redundant regions within the enhancer that can substitute for each other.

What clues do we have about the location within elements I, II and III of these functional regions? Although we cannot yet identify them precisely, several sequences which may be important have been suggested by previous analyses of *Sgs-4* expression.

There are several clues for element II. This very AT-rich 59-bp sequence includes three copies of the sequence AATAAA$^{T}/_{C}$AAA (Figure 5). If this sequence were actually involved in tissue-specific expression, one might expect to find exact or nearly exact copies of it near other salivary gland genes. Table II shows that part of this sequence, TAAA$^{T}/_{C}$AAA, is found in the upstream regions of all five sequenced salivary gland genes, usually in multiple copies. Mutations in the *Sgs-4* underproducing strain Daek are also consistent with an important role for this sequence motif. Daek produces about 0.1% as much *Sgs-4* RNA as wild-type and has multiple sequence alterations distributed throughout the region 5′ to *Sgs-4*. Two of these changes are within element II and they convert the TAAATAAAT sequence at −480 to AAAATAAAA (Figure 5; McGinnis *et al.*, 1983) and may affect the chromatin structure of the region. In salivary gland nuclei from wild-type strains, a strong, tissue-specific DNase I-hypersensitive site is located within this repeating sequence at about −480 (Figure 5; Shermoen and Beckendorf, 1982), but this hypersensitive site is barely detectable in salivary gland nuclei from Daek (McGinnis *et al.*, 1983). We have recently found that a barrier to exonuclease III digestion, which probably represents a tightly bound protein, is located within the repeating region at about −470 (T.Jongens and S.K.Beckendorf, in preparation). Thus, it appears likely that there is tissue-specific binding of one or more proteins to the repeating TAAA$^{T}/_{C}$AAA sequence.

Despite these indications that the TAAA$^{T}/_{C}$AAA repeat is important for tissue-specific expression of *Sgs-4*, it cannot be the only sequence with these properties. Neither element I nor III includes this sequence, but constructs possessing just those two elements do specify salivary gland expression.

Element III is the largest, and we expect that it has several functional regions within it. Consistent with this expectation, there are two distinct DNase I-hypersensitive sites in element III, one at about −405, the other at about −330 (Shermoen and Beckendorf, 1982; McGinnis *et al.*, 1983). The *Sgs-4* underproducing strain Hikone R focuses attention on sequences in the proximal half of element III. It has a 52-bp deletion extending from −305 to −356 and its expression

is reduced about 20-fold (Muskavitch and Hogness, 1982). Several interesting consensus sequences lie within or immediately adjacent to the Hikone R deletion (top of Figure 5). First, there are sequences implicated in the ecdysone-mediated induction of two small heat-shock genes. Presence of these ecdysone-responsive sequences is particularly interesting because ecdysone is required to induce *Sgs-4* expression in mid third instar (Hansson and Lambertsson, 1983; Streck, 1986) and because ecdysone binds to polytene chromosomes at the 3C11-12 puff that contains *Sgs-4* (Dworniczak *et al.*, 1983). The region between −352 and −332 includes a dyad repeat that is similar to a dyad required for induction of the *hps 23* gene by ecdysone (Mestril *et al.*, 1986; see Table III). Examples of this consensus have been found near several other ecdysone-inducible genes, often in inverted pairs (Mestril *et al.*, 1986; Streck, 1986). Moreover, in the Canton S strain a C to T change at a position between the two copies of this sequence is correlated with a 2-fold decrease in *Sgs-4* expression (McGinnis *et al.*, 1983; Hofmann *et al.*, 1987). The sequence between −312 and −287 includes an imperfect dyad that is similar to another ecdysone-responsive sequence, this one required for ecdysone induction of the *hps 27* gene (Riddihough and Pelham, 1987; see Table III). This sequence is probably important for *Sgs-4* expression because insertion of a *Cla*I linker at −301 reduces *Sgs-4* expression >50-fold (T.A.Jongens, unpublished). Both kinds of *hsp* sequence and their *Sgs-4* counterparts consist of inverted repeats, a feature they share with characterized binding sites for vertebrate steroid hormone receptors (Evans and Hollenberg, 1988). The *hsp 23* and *hsp 27* sequences are not obviously related, suggesting either that the ecdysone−receptor complex can bind to diverse DNA sequences or that there are multiple ecdysone receptors with different specificities.

Another region that is important for *Sgs-4* expression is located between these two possible ecdysone-responsive sequences. The Daek strain, which underproduces *Sgs-4*, has two sequence changes in this region and substitution of these two for the corresponding wild-type sequence reduces expression 10-fold (J.Nourse, K.Flynn and S.K.Beckendorf, unpublished). Both of these Daek changes lie within a copy of the consensus topoisomerase II cleavage site, one of them immediately adjacent to the expected point of cleavage (Figure 5; Sander and Hsieh, 1985). Although a role has yet to be established for topoisomerase II in gene expression, it has been shown to be a major component of the nuclear scaffolding (Earnshaw *et al.*, 1985), and a scaffolding attachment site has been located 5′ of *Sgs-4* between −392 and −1 (Gasser and Laemmli, 1986).

So far we know little about which sequences are important for element I. However, it contains several types of short, repeated sequence motif, the most striking of which are four copies of $CAAA^{T}/_{A}$ between −557 and −534. In addition, exonuclease III digestion of salivary gland nuclei detects a bound protein in this same region, at about −550 (T.A.Jongens and S.K.Beckendorf, in preparation).

Co-operation and redundancy in the enhancer

This analysis suggests sequences within each individual element which might be important for expression. But our evidence shows that pairs of elements must interact to promote salivary gland expression. These interactions might be required for several reasons. First, if proteins binding to multiple sequence elements interact with each other and bind co-operatively, their effective binding affinities will be increased, and the resulting complexes are likely to be more stable than any of the individual DNA−protein interactions (Brown, 1984; Ptashne, 1986; Schleif, 1987). This stability might be required to allow proteins bound to the enhancer sufficient time to locate and interact with proteins bound to the promoter. This sort of requirement for multiple binding sites has been demonstrated for sequences in the SV40 enhancer. Short parts of the enhancer that are known to bind protein and to be essential for full enhancer activity cannot function individually as enhancers (Herr and Clarke, 1986; Ondek *et al.*, 1987). However, when such a sequence element is polymerized, it gains enhancer activity, even though no additional specificities have been added (Ondek *et al.*, 1987). It seems likely that such a repeating motif gains activity because it promotes high-affinity, co-operative protein binding. A second reason for multiple, interacting elements may be to provide the specificity needed to pick out a particular sequence in a particular tissue as appropriate for expression. As suggested by a number of authors (Yamamoto, 1985; Echols, 1986; McKnight and Tjian, 1986), specificity could be generated by the presence in a particular tissue of a unique combination of transcription factors, even though each factor is present in several tissues. Finally, this co-operation between sequence elements might be necessary if more than one function of the enhancer is required before salivary gland expression is detected using the transient assay. Recent experiments show that very low level, salivary gland-specific expression of *Sgs-4* is established early in development and that a large increase in expression occurs during the last half of third instar (S.W.Barnett, M.Webster and S.K.Beckendorf, in preparation). It may well be that the production of detectable *Adh* activity in the transient assay requires the co-operation of elements directing early, tissue-specific activation and those directing late third instar induction.

Comparison with other Drosophila tissue specifiers

Tissue-specifying sequences have been defined for several *Drosophila* genes, among the best studied being the segmentation gene *ftz* (Hiromi and Gehring, 1987), the yolk protein genes *yp1* and *yp2* (Garabedian *et al.*, 1985, 1986), the chorion gene *S18* (Orr-Weaver and Spradling, 1986), the *Adh* gene (Posakony *et al.*, 1985; Fischer and Maniatis, 1986) and the glue genes *Sgs-3* (Vijay Raghavan *et al.*, 1986) and *Sgs-4* (Shermoen *et al.*, 1987). For some of these, tissue-specifying sequences are found, as they are for *Sgs-4*, in enhancer-like elements a few hundred base pairs away from the genes they regulate. For example, the divergently transcribed *yp1* and *yp2* genes are separated by 1225 bp and are expressed in both fat body and in ovarian follicle cells (Shepherd *et al.*, 1985). In the intergenic region there are distinct tissue specifiers for fat body and ovary and each specifier acts on both genes (Garabedian *et al.*, 1985). The fat body specifier has been limited to the region between 196 and 321 bp upstream of *yp1* and has been shown to work as an enhancer element which retains at least partial activity when inverted or placed downstream of a reporter gene (Garabedian *et al.*, 1986).

From current evidence, the regulatory regions of the two glue protein genes that have been analyzed, *Sgs-3* and *Sgs-4*, appear to be organized very differently. For *Sgs-3* the region between −106 and +12, in the absence of any more distal sequences, is sufficient for tissue specificity, even though

the level of expression is reduced to ~10% of wild-type (Meyerowitz *et al.*, 1987). In addition, *Sgs-3* has a distal region, within the next 2.3 kb upstream from the gene, that is needed for normal, high-level expression. Location of the active sequences in this region is controversial (between −2.3 and −1.6 kb, Giangrande *et al.*, 1987; between −650 and −220 bp, Meyerowitz *et al.*, 1987) as is the ability of these sequences to act when inverted or placed 3′ to *Sgs-3*. In any case, it appears that the *Sgs-3* tissue specifier is part of the promoter, while that of *Sgs-4* is located not in its promoter but in an upstream enhancer. However, this apparent dissimilarity may be misleading. Because the *Sgs-4* enhancer is almost absolutely required for expression, it has not yet been possible to test whether the promoter itself is tissue specific. Similarly, the upstream *Sgs-3* sequences have not yet been tested for tissue specificity by asking them to alter the expression pattern of a heterologous gene. Thus, it is possible that both genes have tissue-specifying sequences in both their upstream and their promoter regions. Such a functional duplication has been found for immunoglobulin genes, in which both the enhancer and promoter are specific for B cells (Banerji *et al.*, 1983; Gillies *et al.*, 1983; Grosschedl and Baltimore, 1985; Mason *et al.*, 1985). A similar organization is found for the *Drosophila ftz* gene. The striped pattern of *ftz* expression in early embryos is governed by an enhancer element located ~2.6 kb upstream and by a 'zebra' element located within the first 620 bp upstream from the transcription start site (Hiromi and Gehring, 1987).

Can the three functions of the enhancer be uncoupled?

Previous analysis identified three activities of the *Sgs-4* regulatory region—tissue specificity, developmental timing and high-level expression—and showed that each of them possessed the positional flexibility of classical enhancer elements (Shermoen *et al.*, 1987). By placing the *Sgs-4* sequences 3′ to *Adh*, we have demanded this same positional flexibility of the constructs used here to dissect the enhancer region. Thus, all of the pairwise combinations of elements I, II and III can act from 1.7 kb 3′ to *Adh* to direct salivary gland expression. However, we have assayed for just one of the three regulatory activities of the enhancer, tissue specificity. While the I−II, I−III and II−III constructs appear indistinguishable by this criterion, they may well direct different levels of salivary gland expression, or they may result in differential timing of expression. To test these temporal and quantitative aspects of enhancer function and to see whether the three activities can be separated from each other, we are now using several of the constructs described here to produce stably transformed lines which will be suitable for more detailed analysis.

Materials and methods

Fly strains

The recipient strain used for all injections was ACR. As described previously, it is homozygous for Adh^{fn6},*cn*,ry^{506} (Shermoen *et al.*, 1987). Oregon R-B (McGinnis *et al.*, 1983) was used as the wild-type control.

Transient assay for Adh expression

This procedure is adapted from Martin *et al.* (1986). The embryos were dechorionated by washing in 0.5 × Chlorox bleach for 60 s immediately followed by two rinses with 0.1% Triton X-100, 0.7% NaCl and then with distilled water. Embryos were mounted, desiccated and injected as described by Rubin and Spradling (1982). All DNA in the transient assay was at 1μg/μl in injection buffer (5 mM KCl, 0.1 mM sodium phosphate buffer, pH 6.8). Upon reaching third instar the larvae were dissected and their internal tissues stained for *Adh* activity.

Histochemical assay

Tissues were dissected in *Drosophila* Ringers (7.5 g/l NaCl, 0.35 g/l KCl, 0.28 g/l $CaCl_2 \cdot 2H_2O$) and stored on ice for no more than 30 min. The tissue samples were then fixed in fresh 1% glutaraldehyde (Sigma Grade 1, dissolved in 10 mM Tris−HCl, pH 8.2) for 15 min on ice, washed with ice-cold Ringers (3 × 10 min), and then stained for 5−15 min with *Adh*-specific stain (0.2 M Tris−HCl, pH 8.2, 0.5 mg/ml nitroblue tetrazolium, 0.02 mg/ml phenazine methosulfate, 0.5 mg/ml NAD, 0.27 M 2-butanol; Ursprung *et al.*, 1970). It was important to use fresh, high-quality glutaraldehyde, as older or non-EM grade reagent gave faint blue staining in the salivary glands of known *Adh*-negative controls. Control larvae were stained with every new batch of larvae to ensure the same quality of staining with each experiment.

DNA constructions

The pAF 0 plasmid contains the *Hpa*I/*Eco*RI fragment of Adh^F (Posakony *et al.*, 1985). This fragment includes the *Adh* structural gene along with its proximal but not its distal promoter. The *Hpa*I site is 380 bp 5′ to the proximal cap site; the *Eco*RI site is ~2.5 kb 3′ to the cap site. To construct pAF 0, the *Hpa*I site was converted to a *Sal*I site and the resulting *Sal*I/*Eco*RI fragment inserted into the polylinker of pUC 13 (Figure 6).

To allow production of 5′ and 3′ deletions of the *Sgs-4* enhancer, two derivatives of pEMBL8(+) (Dente *et al.*, 1983) were prepared. Each contained the 420-bp *Dde*I fragment of *Sgs-4* along with a unique stuffer fragment. pEMBL8(Δ5′) was constructed by inserting the 622-bp *Sal*I/*Hin*dIII fragment of pBR322 (containing a unique *Sph*I site near the *Sal*I site) into the polylinker of pEMBL8(+) (see Figure 6). pEMBL8(Δ3′) has the 998-bp *Sal*I/*Eco*RI fragment from lambda inserted into the polylinker (containing a unique *Bcl*I site near the *Sal*I site). To insure that the *Bcl*I site could be cut, pEMBL8(Δ5′) was grown in the *dam*⁻ strain GM113 (*F*⁻, *dam*⁻, *thy*⁻).

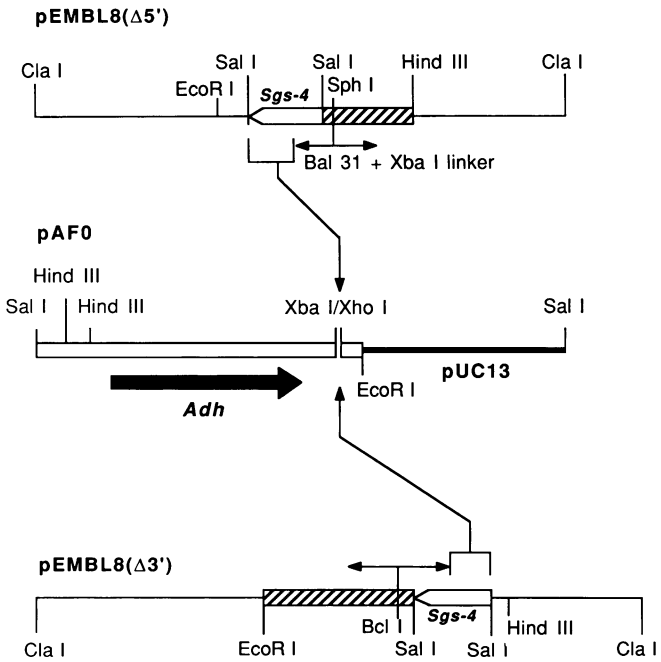

Fig. 6. Construction of 5′ and 3′ *Sgs-4* enhancer deletions for use in the transient assay. *Sgs-4* regulatory sequences, shown as an open bar whose point indicates orientation, were derived from the *Dde*I/*Dde*I fragment of *Sgs-4* that is normally located between −149 and −568 (Shermoen *et al.*, 1987). The cross-hatched boxes indicate 'stuffer' fragment DNA. pEMBL8(+) DNA is indicated by the thin line. Following 5′ or 3′ *Bal*31 deletion and attachment of *Xba*I linkers, the *Sgs-4* sequence was excised as a *Sal*I/*Xba*I fragment and inserted into the *Adh*-containing plasmid, pAF 0, at the *Xba*I and *Xho*I sites 3′ to the *Adh* gene. It should be noted that the *Sgs-4* regulatory sequences end up in opposite orientations in the 5′ and 3′ series.

The *Sgs-4 DdeI/DdeI* fragment (−568 to −149) came originally from the Canton S stock and was obtained for the current experiments from pCS 0.85 (Shermoen *et al.*, 1987). The *DdeI* ends were filled in using the large Klenow fragment from DNA polymerase I, and *SalI* linkers were attached. The resulting fragment was then inserted into the *SalI* site of the pEMBL constructs and oriented such that *Bal*31 exonuclease digestion from the unique restriction sites within the stuffer fragment created 3′ deletions into the *Sgs-4* DNA in pEMBL8(Δ3′) and 5′ deletions in pEMBL8(Δ5′) (see Figure 6). Following *Bal*31 deletion, *XbaI* linkers were attached and the plasmids reclosed.

The partially deleted *Sgs-4* sequences were then transferred to pAF 0 by excising the *SalI/XbaI* fragment of the pEMBL plasmids and inserting it into the *XbaI* and *XhoI* sites 3′ to the *Adh* gene in pAF 0. The internal 700-bp *XbaI/XhoI* fragment from pAF 0 was removed.

All of the constructs in Figure 4B except the last were derived from the 5′ or 3′ deletions either by truncation at a convenient restriction site, by joining two deleted plasmids at an internal restriction site, or by a combination of the two. The last plasmid, pAF −437:−149/−568:−498, was constructed from a plasmid that contained two tandemly repeated copies of the sequence −568:−149 with the internal sequence from −498 to −437 deleted. Since an *XbaI* site was present at the point of the internal deletion in both tandem copies, digestion with *XbaI* released the pAF −437:−149/−568:−498 sequence which could then be inserted in pAF 0.

DNA sequencing

All deletion endpoints were confirmed by DNA sequencing (Sanger *et al.*, 1977) using an NEBL dideoxynucleotide sequencing kit modified for sequencing supercoiled plasmids (Agellon and Chen, 1986). Primer hybridization was done at 50°C and extension with Klenow fragment at 37°C. The −20 and −40 primers for sequencing from the pEMBL polylinkers were obtained from NEBL. To facilitate sequencing those deletions whose endpoints were >200 bp from commercially available primer sites, we had primers made which were complementary to *Sgs-4* DNA at positions 5′ −405 to −387 3′ and 5′ −283 to −300 3′.

Acknowledgements

We thank Jim Posakony for donating the p13E3 plasmid from which Adh^F sequences were extracted. We are grateful to Jill Jongens and Kathy Flynn for critically reading the manuscript. This work was supported by grants from the Department of Energy and the National Institutes of Health.

References

Agellon,L.B. and Chen,T.T. (1986) *Gene Anal. Technol.*, **3**, 86–89.
Beckendorf,S.K. and Kafatos,F.C. (1976) *Cell*, **9**, 365–373.
Banerji,J., Olson,L. and Schaffner,W. (1983) *Cell*, **33**, 729–740.
Brown,D.D. (1984) *Cell*, **37**, 359–365.
Dente,L., Cesareni,G. and Cortese,R. (1983) *Nucleic Acids Res.*, **11**, 1645–1655.
Dworniczak,B., Seidel,R. and Pongs,O. (1983) *EMBO J.*, **2**, 1323–1330.
Earnshaw,W.C., Halligan,B., Cooke,C.A., Heck,M.M.S. and Liu,L.F. (1985) *J. Cell Biol.*, **100**, 1706–1715.
Echols,H. (1986) *Science*, **233**, 1050–1056.
Evans,R.M. and Hollenberg,S.M. (1988) *Cell*, **52**, 1–3.
Fischer,J.A. and Maniatis,T. (1986) *EMBO J.*, **5**, 1275–1289.
Fraenkel,G. and Brookes,V.J. (1953) *Biol. Bull.*, **105**, 442–449.
Garabedian,M.J., Hung,M.C. and Wensink,P.C. (1985) *Proc. Natl. Acad. Sci. USA*, **82**, 1396–1400.
Garabedian,M.J., Shepherd,B.M. and Wensink,P.C. (1986) *Cell*, **45**, 859–867.
Garfinkel,M.D., Pruitt,R.E. and Meyerowitz,E.M. (1983) *J. Mol. Biol.*, **168**, 765–789.
Gasser,S.M. and Laemmli,U.K. (1986) *Cell*, **46**, 521–530.
Giangrande,A., Mettling,C. and Richards,G. (1987) *EMBO J.*, **6**, 3079–3084.
Gillies,S.D, Morrison,S.L., Oi,V.T. and Tonegawa,S. (1983) *Cell*, **33**, 717–728.
Grosschedl,R. and Baltimore,D. (1985) *Cell*, **41**, 885–897.
Hansson,L. and Lambertson,A. (1983) *Mol. Gen. Genet.*, **192**, 395–401.
Herr,W. and Clarke,J. (1986) *Cell*, **45**, 461–470.
Hofmann,A., Keinhorst,A., Krumm,A. and Korge,G. (1987) *Chromosoma*, **96**, 8–17.
Hiromi,Y. and Gehring,W.J. (1987) *Cell*, **50**, 963–974.
Korge,G. (1975) *Proc. Natl. Acad. Sci. USA*, **72**, 4550–4554.
Mason,J.O., Williams,G.T. and Neuberger,M.S. (1985) *Cell*, **41**, 479–487.
McGinnis,W., Shermoen,A.W. Heemskerk,J. and Beckendorf,S.K. (1983) *Proc. Natl. Acad. Sci. USA*, **80**, 1063–1067.
McKnight,S. and Tjian,R. (1986) *Cell*, **46**, 795–805.
McNabb,S.L. and Beckendorf,S.K. (1986) *EMBO J.*, **5**, 2331–2340.
Martin,P., Martin,A., Osmani,A. and Sofer,W. (1986) *Dev. Biol.*, **117**, 574–580.
Mestril,R., Schiller,P., Amin,J., Klapper,H., Ananthan,J. and Voellmy,R. (1986) *EMBO J.*, **5**, 1667–1673.
Meyerowitz,E.M., Vijay Raghavan,K., Mathers,P.H. and Roark,M. (1987) *Trends Genet.*, **3**, 288–293.
Muskavitch,M.A.T. and Hogness,D. (1980) *Proc. Natl. Acad. Sci. USA*, **77**, 7362–7366.
Muskavitch,M.A.T. and Hogness,D. (1982) *Cell*, **29**, 1041–1051.
Ondek,B., Shepard,A. and Herr,W. (1987) *EMBO J.*, **6**, 1017–1025.
Orr-Weaver,T.L. and Spradling,A.C. (1986) *Mol. Cell. Biol.*, **6**, 4624–4633.
Posakony,J.W., Fischer,J.A. and Maniatis,T. (1985) *Cold Spring Harbor Symp. Quant. Biol.*, **50**, 515–520.
Ptashne,M. (1986) *Nature*, **322**, 697–701.
Riddihough,G. and Pelham,H.R.B. (1987) *EMBO J.*, **6**, 3729–3734.
Rubin,G.M. and Spradling,A.C. (1982) *Science*, **218**, 348–353.
Sander,M. and Hsieh,T. (1985) *Nucleic Acids Res.*, **13**, 1057–1072.
Sanger,F. Nicklen,S. and Coulson,A.R. (1977) *Proc. Natl. Acad. Sci. USA*, **74**, 5463–5467.
Schleif,R. (1987) *Nature*, **327**, 369–370.
Shepherd,B., Garabedian,M.J., Hung,M.-C. and Wensink,P.C. (1986) *Cold Spring Harbor Symp. Quant. Biol.*, **50**, 521–526.
Shermoen,A.W. and Beckendorf,S.K. (1982) *Cell*, **29**, 601–607.
Shermoen,A., Jongens,J., Barnett,S., Flynn,K. and Beckendorf,S.K. (1987) *EMBO J.*, **6**, 207–214.
Shore,E.M. and Guild,G.M. (1986) *J. Mol. Biol.*, **190**, 149–158.
Steller,H. and Pirrotta,V. (1984) *EMBO J.*, **3**, 165–173.
Streck,R.D. (1986) Ph.D. thesis, University of California, Berkeley.
Ursprung,H.I., Sofer,W. and Burroughs,N. (1970) *Wilhelm Roux's Arch. Dev. Biol.*, **164**, 201–208.
Vijay Raghavan,K., Crosby,M.A., Mathers,P.H. and Meyerowitz,E.M. (1986) *EMBO J.*, **5**, 3321–3326.
Yamamoto,K. (1985) *Annu. Rev. Genet.*, **19**, 209–252.

Received on March 15, 1988