Development 118, 1003-1012 (1993)
Printed in Great Britain © The Company of Biologists Limited 1993

Molecular characterization of the *lethal of scute* genetic function

María D. Martín-Bermudo†, Francisco González, María Domínguez, Isabel Rodríguez, Mar Ruiz-Gómez‡, Susana Romani§, Juan Modolell and Fernando Jiménez*

Centro de Biología Molecular "Severo Ochoa" (CSIC-UAM), Universidad Autónoma, 28049 Madrid, Spain

*Author for correspondence
†Present address: Wellcome/CRC Institute, Cambridge CB2 1QR, UK
‡Present address: Department of Zoology, University of Cambridge, Cambridge CB2 3E, UK
§Present address: Max-Planck Institut für Biophysikalische Chemie, D-3400 Göttingen, Germany

SUMMARY

The *lethal of scute* (*l'sc*) genetic function, which plays an essential role in the early development of the central nervous system of the *Drosophila* embryo, is localized within the *achaete-scute* complex (AS-C). Several lines of evidence have suggested that the AS-C T3 transcription unit corresponds to the *l'sc* function. We demonstrate that short fragments of DNA, containing the T3 transcribed region and a few kilobases of flanking sequences, rescue, albeit partially, the lethality and neural phenotype of *l'sc* deletions. Still, the complex wild-type pattern of expression of T3 is not reproduced by the transduced genes. This depends on *cis*-control elements scattered within the entire AS-C DNA and intermingled with regulatory elements specific for other AS-C transcription units. These elements are necessary for the initial activation of T3 in the neuroectoderm, probably mediated by axis-patterning genes. The presence of a cluster of E-boxes, upstream of the T3 transcribed region, suggests another level of control of T3 expression by basic-helix-loop-helix proteins, among them its own gene product.

Key words: *Drosophila*, *lethal of scute* gene, *cis*-regulatory elements, central nervous system, neurogenesis

INTRODUCTION

Subdivision 1B of the X chromosome of *Drosophila melanogaster* contains at least six genetic functions involved in the development of the central (CNS) and peripheral (PNS) nervous systems. Three of these functions affect major aspects of CNS development (Jiménez and Campos-Ortega, 1979, 1987; White, 1980). One of them is *lethal of scute* (*l'sc*), which has been localized within the *achaete-scute* complex (AS-C) (García-Bellido, 1979).

Developmental studies and mosaic analysis had suggested that the *l'sc* function is required for CNS development (García-Bellido and Santamaría, 1978). More precisely, a deletion of the *l'sc* region of the AS-C causes slight defects in neuroblast (NB) segregation and cell death in the developing embryonic CNS, which lead to the loss of neurons, gross anatomical malformation of the ventral cord and embryonic lethality (Jiménez and Campos-Ortega, 1987; Martín-Bermudo et al., 1991). Other genes of the AS-C, namely *achaete* (*ac*), *scute* (*sc*) and *asense* (*ase*), were shown to be mainly involved in PNS development, both in the embryo and in the adult (García-Bellido, 1979; Dambly-Chaudière and Ghysen, 1987; Marí-Beffa et al., 1991; Domínguez and Campuzano, 1993), although their deletion also affects CNS development, when *l'sc* is simultaneously absent (Jiménez and Campos-Ortega, 1987).

Genetically, the *l'sc* function is defined by the breakpoints of $In(1)sc^4$ and $In(1)sc^9$ (see Fig. 5 and Muller, 1955; García-Bellido, 1979), for it is uncovered by the synthetic chromosomal deletion obtained after recombination of both inversions. However, an extensive search for EMS-induced lethal mutations at the distal region of the X chromosome failed to demonstrate a lethal function within the AS-C (J. Lim and K. White, personal communication). This negative result may indicate that *l'sc* is a relatively small target for mutagenesis or that *l'sc* is, in itself, a complex locus.

The molecular cloning of the AS-C has shown that the breakpoints of $In(1)sc^4$ and $In(1)sc^9$ are separated by 20 kb. Within this stretch of DNA, one transcription unit has been detected (Campuzano et al., 1985); it gives rise to the T3 RNA which, similarly to other genes of the AS-C, encodes a basic-helix-loop-helix protein (bHLH) (Villares and Cabrera, 1987; Alonso and Cabrera, 1988; González et al., 1989; and the present work). This RNA has been assumed to correspond to the *l'sc* function (Campuzano et al., 1985). Thus, the T3 RNA accumulates in all known neurogenic regions of the embryo (Cabrera et al., 1987; Romani et al., 1987), and the corresponding protein is found in proneural clusters of the neuroectoderm and in the segregating NBs (Cabrera, 1990; Martín-Bermudo et al., 1991). However, conclusive evidence for T3 RNA involvement in the *l'sc* function is still lacking. It is also unknown how much of the T3 gene flanking regions, or of more distant regions of the AS-C, are necessary for this function.

To resolve the above issues, we have first improved on the molecular characterization of the T3 transcription unit, by defining its transcriptional origin and termination and by sequencing both a near to full-length cDNA and a genomic fragment with nearly one kilobase of upstream sequences. Secondly, we have shown, by P-mediated transformation experiments, that short fragments of DNA containing the T3 transcribed sequences rescue, although in only a fraction of the mutant embryos, the lethality and the CNS mutant phenotype associated with *l'sc* deletions. However, these transgenes, in a *l'sc*$^-$ background, are expressed in a simpler pattern and at lower levels than the wild-type T3 gene, suggesting that additional sequences are necessary for wild-type expression. Indeed, the analysis of T3 expression in a series of AS-C mutations has revealed the existence of multiple *cis*-regulatory regions, distributed along the entire complex, required for the proper activation of T3 in specific proneural clusters of the embryonic neuroectoderm. These results unambiguosly define the T3 transcription unit as the main, or exclusive, component of the *l'sc* function, and demonstrate that this gene, like the *ac* and *sc* genes, has highly complex *cis*-regulatory regions.

MATERIALS AND METHODS

DNA sequencing

Subclones of the 3.2 kb *Eco*RI fragment (coordinates +18.1 to +21.3 of the AS-C physical map Campuzano et al., 1985) that contains the T3 structural gene and of a near to full-length cDNA (Campuzano et al., 1985) were obtained in M13 (Messing et al., 1977) and pBluescript KS(+) (Stratagene) vectors. Progressive deletions of these clones were prepared by exonuclease III treatment (Henikoff, 1984). Either single- or double-stranded templates were sequenced by the dideoxy chain termination method (Sanger et al., 1977). Part of the sequence was obtained using the modification of Tabor and Richardson (1987) to use chemically modified T7 DNA polymerase. Synthetic oligonucleotides of defined sequence were used to extend sequences and fill in gaps. Consensus sequences were assembled and analyzed with the help of the Staden (1980) and University of Wisconsin GCG software packages (Devereux et al., 1984). Sequences were determined in both strands.

RNAse protection and primer extension analysis

An RNAse protection experiment was performed, as described in Melton et al. (1984) and Campuzano et al. (1986), with a single-stranded RNA probe, homologous to the DNA contained between the *Hin*dIII and *Pst*I sites located at positions 307 and 1189 of the T3 sequence (Fig. 1). This indicated that the T3 RNA and the probe were complementary for approximately 221 nucleotides towards the left of the *Pst*I site and suggested that the origin of transcription was in the vicinity of T_{973}. This was confirmed with primer extension experiments, which showed that transcription starts on A_{972}. These experiments were performed according to Bensi et al. (1985). A synthetic oligonucleotide complementary to the sequence extending from T_{999} to C_{1019} was labeled with T4 polynucleotide kinase and [γ-^{32}P]ATP (3000 Ci/mmole) and hybridized at 60°C for 30 minutes in 25 mM Tris, pH 7.7, 75 mM KCl, 3 mM $MgCl_2$ and 5 mM dithiothreitol with 50 mg of poly(A)$^+$ RNA from 0-12 hour old embryos. The hybridized primer was extended with 10 units of AMV reverse transcriptase and deoxynucleotide triphosphates by incubation at 42°C for 1 hour. The length of the synthesized DNA was 48 nucleotides.

Plasmid constructions

p3.2T3contains a genomic DNA fragment of 3.2 kb *Eco*RI-*Eco*RI, (coordinates +21.3 and +18.1 from the AS-C molecular map, Campuzano et al., 1985). It was inserted in the *Eco*RI cloning site of the transformation vector pPSXD1A*Adh*$^+$.

p5.9T3 contains a genomic fragment of 5.9 kb *Xba*I-*Eco*RI (coordinates +23.3 to +18.1). It comprises 2.7 additional kilobases in the 5′ region of the former construct. The genomic fragment was cloned into the *Xba*I site of the transformation vector pC20.1*ry*$^+$.

Drosophila transformation

P-element-mediated transformation into *Adh*^{fn6} *cn*; *ry*^{506} and *ry*^{506} embryos was performed according to Rubin and Spradling (1982), using 0.5 mg/ml of either p3.2T3or p5.9T3 plasmid DNA and 0.15 mg/ml of pπ25.7wc DNA, to provide for transposase (Karess and Rubin, 1984).

Drosophila strains

All AS-C mutant strains used in this work are from the collection of A. García-Bellido and are described in Lindsley and Zimm (1992). The *Df(1)sc*^{4L}*sc*^{9R}, labelled with *y* and *w*, was synthesized anew, by recombination of viable *In(1)sc*^{4} and *In(1)sc*^{9} chromosomes, to avoid modifiers that might have accumulated in older stocks of the deficiency. In spite of this precaution, the viability of *Df(1)sc*^{4L}*sc*^{9R}*/y*^{2}*Y61l* males was only 7%, as compared to that of *+/y*^{2}*Y61l* males, even though the duplication *y*^{2}*Y61l* covers the entire AS-C and completely rescues the lethality of *Df(1)sc*^{19} (*ac*$^-$, *sc*$^-$, *l'sc*$^-$) flies (data not shown). The impaired viability of *Df(1)sc*^{4L}*sc*^{9R}*/y*^{2}*Y61l* flies should, therefore, be attributed to deleterious factors mapping outside of the deficiency, most probably to the proximal breakpoints of the two inversions used to synthesize the deficiency. For this reason, the estimated percentage of adult viability of *Df(1)sc*^{4L}*sc*^{9R} flies rescued with T3 transgenes (Table 1) has been corrected by a factor that takes into consideration the reduced viability of the *Df(1)sc*^{4L}*sc*^{9R}*/y*^{2}*Y61l* control flies.

Immunocytochemistry

Dechorionated embryos were fixed for 20 minutes in heptane saturated with 4% formaldehyde in PBS, and vitelline membranes were removed by methanol treatment. All subsequent incubations were performed in PBS, 0.1% Tween 20, 0.1% bovine serum albumin. For the analysis of CNS phenotypes in AS-C deficiencies rescued with T3 transgenes, embryos were first treated with a rabbit anti-horseradish peroxidase (HRP) antibody (Cappel) (Jan and Jan, 1982), then with a HRP-labelled secondary antibody, and finally stained with DAB and H_2O_2. For the analysis of *l'sc* expression in AS-C mutants, embryos were incubated with a rat anti-l'sc antibody (Martín-Bermudo et al., 1991), followed by a biotin-labelled secondary antibody and the Vectastain Elite ABC kit (Vectorlabs). Before the staining reaction, the embryos were incubated with 1% glutaraldehyde in PBS for 5 minutes and washed for 20 minutes, treatment that substantially reduces the background. DAB staining was performed in the presence of Ni and Co ions.

Other procedures

Plasmid DNA, fly RNA and single-stranded RNA probe preparations were performed as described (Maniatis et al., 1982; Melton et al., 1984; Campuzano et al., 1985).

RESULTS

The T3 transcription unit

The T3 gene gives rise to a 1.1 kb RNA (Campuzano et al., 1985; and Fig. 5). The sequence of a genomic fragment of

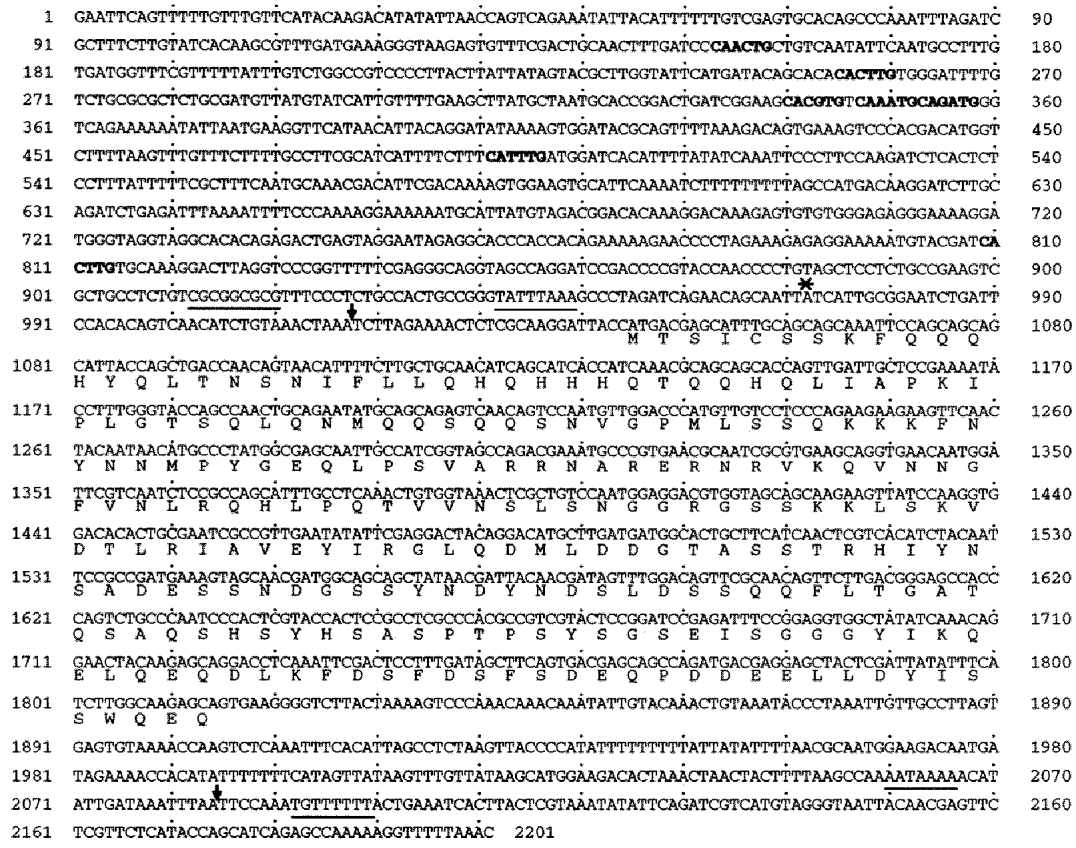

Fig. 1. Sequences of the T3 transcription unit and its deduced protein product. The nucleotide sequence corresponds to the genomic DNA. Vertical arrows preceding T_{1018} and following A_{2084} mark the start and end of the colinearity region (1067 nucleotides long) between genomic and cDNA clones. The only difference between the two sequences was at the 5′ end of the cDNA. This contained 16 nucleotides with a sequence identical to that contained between T_{1890} and G_{1905} and should correspond to a cloning artifact. The long region of colinearity, which starts 46 nucleotides 3′ from the origin of transcription (marked with an asterisk) together with RNAse protection and primer extension experiments, show that the T3 transcription unit, similarly to those of the other AS-C proneural genes, does not have introns. A polyadenylation signal (underlined) is found 25 nucleotides before the 3′ end of the cDNA (at A_{2084}) and a TG motif (underlined) (Birnstiel et al., 1985) is located 8 nucleotides after its end. Most likely, transcription termination occurs very near this nucleotide. In a previous report (Alonso and Cabrera, 1988), the colinearity region for this same cDNA clone was mistakenly considered to be 1183 nucleotides long, starting at the same point as in our sequence (T_{1018}) but ending at C_{2201}, the end of our genomic sequence. Other differences between that report and ours are: an extra G is included following T_{935}; G_{1033} is changed to an A; and A_{1637} is changed to a G, which changes a histidine into an arginine in the deduced T3 protein sequence. We think it most unlikely that these differencies are due to polymorphisms since, to our knowledge, both sequencing projects have been conducted on the same genomic subclone. Putative GC and TATA boxes are underlined. Similar GC-rich stretches are found in comparable locations of the upstream regions of the other three AS-C proneural genes (*ac*, *sc* and *ase*) (Alonso and Cabrera, 1988; González et al., 1989; Villares and Cabrera, 1987). In the upstream nontranscribed region, putative binding sites for bHLH proteins (E-boxes) are indicated in bold letters.

T3 DNA that comprises an open reading frame has previously been reported (Alonso and Cabrera, 1988). We have further characterized the T3 transcription unit by independently sequencing a larger segment of genomic DNA and a near to full-length cDNA, and by determining its origin of transcription. The results are shown in Fig. 1. RNAse protection and primer extension experiments showed that transcription starts on A_{972}. Thirty nucleotides upstream from this point, a rather unconventional TATA box (TATTTAAA) is found. Seven CANNTG motifs, putative binding sites for bHLH transcriptional regulators (Blackwell and Weintraub, 1990), are located in the upstream region, which suggests that T3 transcription may be regulated by its own protein and/or by other bHLH proteins (see below). The sequences comprising and surrounding the E-boxes at positions 255 and 809 are reminiscent of *snail* protein-binding sites (Ip et al., 1992). *snail* represses *l'sc* transcription in the presumptive embryonic mesoderm (Kosman et al., 1991).

Rescue of the *lethal of scute* phenotype by transduced T3 genes

Genetic and molecular data suggest that the T3 transcription unit, the only one detected within the 20 kb of DNA uncovered by the synthetic chromosomal deletion $Df(1)sc^{4L}sc^{9R}$, is responsible for the *l'sc* function (García-Bellido, 1979; Campuzano et al., 1985; Jiménez and Campos-Ortega, 1987). This is the smallest available deficiency of the *l'sc* gene for which no point mutations have yet been isolated. To obtain more conclusive evidence for

Table 1. Rescue of the *lethal of scute* lethality by transduced T3 genes

	Cross*		Hatched larvae†			Adults	
Nº	Females	Males	y^+	y^-‡	% Viability	y^-	% Viability§
1	*Df(1)sc^{4L}sc^{9R}/+; 3.2T3.1/+; 3.2T3.2/+*	*3.2T3.1; 3.2T3.2*	2089	337 (696)	48.4	4	8.2
2	*Df(1)sc^{19}/+; 3.2T3.1/+; 3.2T3.2/+*	*3.2T3.1; 3.2T3.2*	1531	244 (510)	47.8	0	0
3	*Df(1)sc^{4L}sc^{9R}/+; 5.9T3.1/+*	*5.9T3.1*	1081	171 (360)	44.7	5	19.8

*The *3.2T3.1* transduced gene is inserted in the 2nd chromosome and *3.2T3.2* and *5.9T3.1* in the 3rd chromosome.
†Larvae with a wild-type first chromosome are y^+ and those with a mutant chromosome are y^-.
‡ In parenthesis, the number of y^- larvae corresponding to 100% viability, that is one third of the number of hatched y^+ larvae.
§These viability values have been corrected by a factor that takes into consideration the low viability of *Df(1)sc^{4L}sc^{9R}/y^{2}Y61l* control flies, where the *y^{2}Y61l* chromosome carries a duplication that contains the entire ASC (see Materials and methods).

the identity between the T3 and the *l'sc* genes, we prepared, by P-element-mediated transformation, *Drosophila* lines carrying a 3.2 kb DNA fragment that contains the T3 transcribed region, 0.9 kb of the upstream region and 1.2 kb of the downstream region. Four of such transformants (lines *3.2T3.1-4*) were tested for their ability to rescue the embryonic lethality associated with the *Df(1)sc^{4L}sc^{9R}*. One copy of any of the transgenes rescued the embryonic lethality of a fraction of the *l'sc*$^-$ embryos, up to 26% with line *3.2T3.1* (not shown). Two to four copies improved rescuing to 48% (Table 1, cross 1). However, only a few of the *l'sc*$^-$ larvae reached adulthood, being the adult viability 8% when compared to the viability of *Df(1)sc^{4L}sc^{9R}/y^{2}Y61l* control flies (see Materials and Methods). The rescue of the embryonic lethality of the larger deletion *Df(1)sc^{19}* (*ac*$^-$, *sc*$^-$, *l'sc*$^-$) with two to four copies of the transgene (Table 1, cross 2) was 47%. In this case, all larvae died before the third instar, probably because of the absence of the *ac* and *sc* genes, which should aggravate the effects of the inefficient rescue of *l'sc*.

To attempt improving the rescue achieved with the *3.2T3* transformants, we prepared fifteen new transformant lines (*5.9T3.1-15*) with a fragment that provided 2.7 kb more DNA in the T3 upstream region, that is, a total of 3.6 kb. With one copy of this transgene, line *5.9T3.1* was the most efficient in yielding *Df(1)sc^{4L}sc^{9R}* viable adults (not shown). One to two copies of this *5.9T3.1* gene allowed hatching of 48% of *l'sc*$^-$ embryos (Table 1, cross 3). Some of them

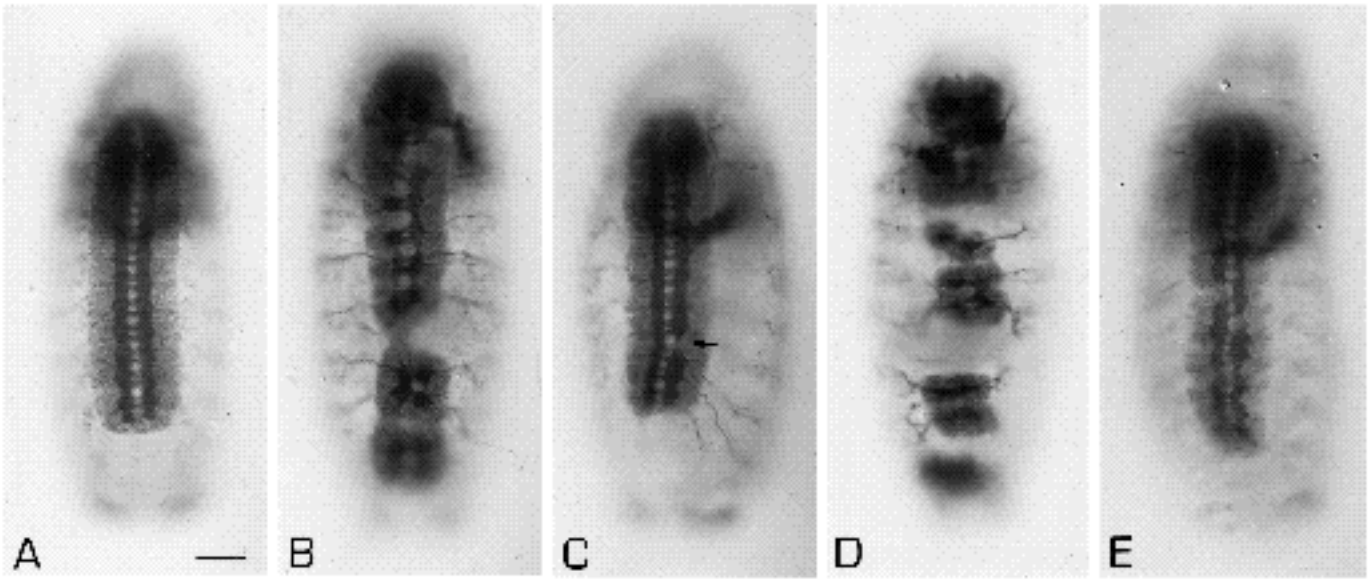

Fig. 2. Rescue of the neural phenotype of AS-C deficiencies by a transduced T3 gene. The panels show ventral views of different late embryos stained with anti-HRP antibodies. (A) Wild-type embryo with the characteristic morphology of the condensed ventral cord and the ladder-like organization of the axonal tracts. (B) *Df(1)sc^{19}* embryo. Note the incomplete condensation and interruptions of the ventral cord, as well as the alterations in the pattern of connectives and comissures. (C) A *Df(1)sc^{19}* embryo, progeny from cross 2 (Table 1), bearing between two and four copies of the *3.2T3* gene. The ventral cord is essentially normal, except for the interruption of a connective at the level of one abdominal neuromere (arrow). (D) *Df(1)260.1* embryo. The severe hypoplasy of the CNS has reduced the ventral cord to a series of isolated clusters of neurons. (E) A *Df(1)260.1* embryo, bearing between two and four copies of *3.2T3*, obtained from a cross similar to crosses 1 and 2 (Table 1). The mutant phenotype has been substantially rescued, and the ventral cord is very similar to that of the wild type. Still, some defects in the axonal pattern are evident. It should be noted that the absence of labeled chromosomes hindered the unambiguous identification of all mutant embryos. Therefore, it might well be that some wild-type-looking embryos were in fact mutant embryos that had been phenotypically rescued to an extent larger than those shown in C and E. Scale bar in A: 50 μm.

reached adulthood, the adult viability being nearly 20% when compared to that of *Df(1)sc^{4L}sc^{9R}/y^{2}Y61l* control flies. This indicates that, although the *5.9T3.1* insertion rescues *l'sc*$^{-}$ lethality more efficiently than any of the *3.2T3* insertions, its rescuing capacity is still far from complete.

We also investigated the capacity of the *3.2T3* transgene to rescue the neural phenotype of *l'sc* deletions. Since a fraction of the rescued *Df(1)sc^{4L}sc^{9R}* embryos reached adulthood, we assumed that the gross CNS morphology of at least that fraction of embryos should be fairly normal. Therefore, they were not further analyzed. *Df(1)sc^{19}* embryos show a partial hypoplasy of the CNS (compare Fig. 2A and B), which is largely due to the lack of *l'sc* (Jiménez and Campos-Ortega, 1987). Two to four doses of the *3.2T3* gene substantially improved this mutant phenotype (Fig. 2C). Similarly, the severe neural hypoplasy of *Df(1)260.1* embryos (Fig. 2D), which lack the entire AS-C, was also considerably rescued (Fig. 2E).

We also analyzed the expression of the *3.2T3* and *5.9T3* transgenes in *Df(1)sc^{4L}sc^{9R}* embryos. In the wild type, T3 starts to be expressed in the blastoderm (stage 5), in a pattern that changes during gastrulation (Cabrera, 1990; Martín-Bermudo et al., 1991). By the end of stage 8, a well-defined pattern of groups of ectodermal cells (proneural clusters) and the S1 NBs that delaminate from them, express T3. Later, in stage 9, T3 expression occurs in new proneural clusters that give rise to the S2 NBs. In contrast, during these stages, the transgenes were apparently not expressed (not shown). T3 protein accumulation from the transgenes was first detectable in stage 9, both in NBs and, to a lesser extent, in the overlaying neuroectoderm (Fig. 3A,B,D). Moreover, this late expression was much weaker than that of the wild-type gene and it was difficult to appreciate its characteristically defined spatial pattern (Fig. 3C,E).

In summary, we believe that the rescue of the lethality and of the neural phenotype (although incomplete, most likely because of the imperfect expression of the transduced genes), together with previous genetic and molecular evidence, unequivocally confirm that the T3 gene is responsible for the *l'sc* function. Accordingly, we shall refer from now on to the AS-C T3 gene as the *l'sc* gene.

Mapping of *cis*-regulatory regions of the *l'sc* gene within the AS-C

The low level and simplified patterns of expression of the *l'sc* transgenes suggest that they do not contain all the regulatory elements necessary for wild-type expression. These elements should lie somewhere else within the 90 kb that spans the AS-C. In fact, expression of the *sc* gene in imaginal discs appears to be driven by discrete *cis*-regulatory elements, scattered through a length of 50 kb (Ruiz-Gómez and Modolell, 1987). Similar *cis*-regulatory sites have been proposed to direct embryonic *ac* expression both in the CNS (Skeath et al., 1992) and in the PNS (Ruiz-Gómez and Ghysen, 1993). To test for the presence of regulatory elements for *l'sc*, we analyzed its expression in AS-C mutants. Most of the mutations studied are small deletions or chromosomal rearrangements (inversions or translocations) that either remove portions of the AS-C or physically separate them from the *l'sc* coding region, thus uncoupling presumptive *cis*-regulatory elements from the basal

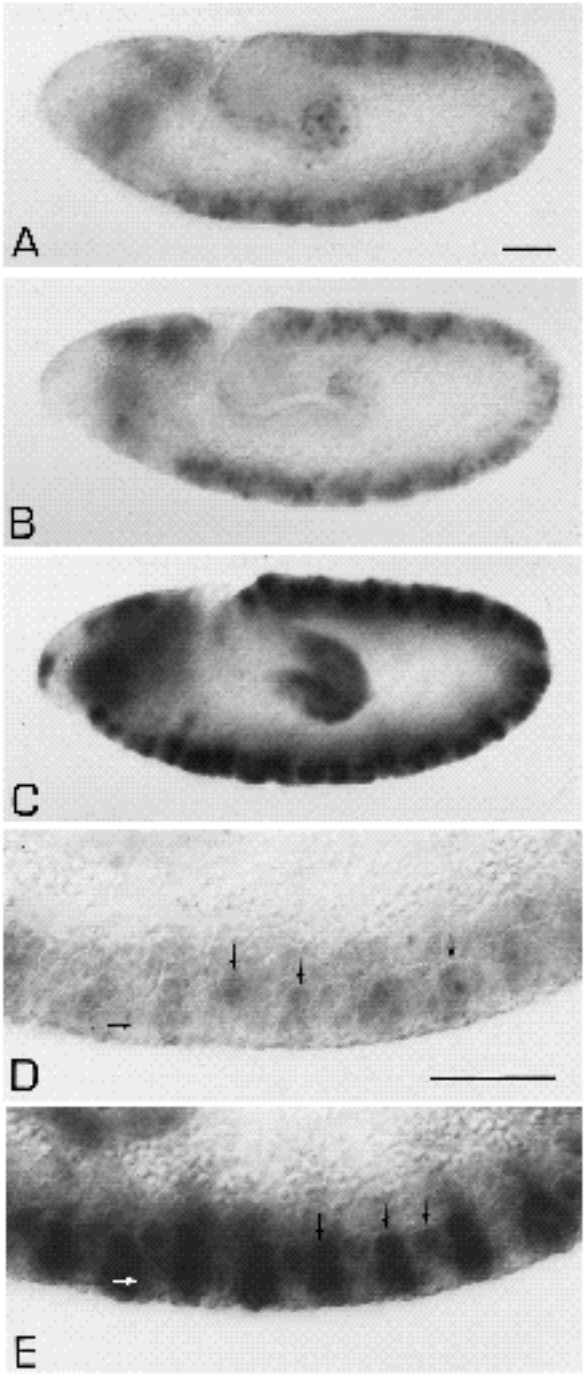

Fig. 3. Expression of transduced T3 genes in *Df(1)sc^{4R}sc^{9L}* embryos. Embryos from crosses 1 and 3 (Table 1) were stained with a l'sc antibody. (A) Lateral view of a *Df(1)sc^{4L}sc^{9R}* embryo bearing between two and four copies of the *3.2T3* transduced gene. (B) A *Df(1)sc^{4L}sc^{9R}* embryo bearing one to two copies of the *5.9T3* transduced gene. Accumulation of the L'SC protein is mainly observed in NBs, and also in the neuroectoderm. Both kind of embryos show comparable levels of staining, but these are significantly lower than in a wild-type embryo (C). This latter embryo is a sibling of that shown in B. (D,E) Lateral views, at higher magnification, of embryos similar to those shown in B and C, respectively. Vertical arrows point to some of the NBs and horizontal arrows indicate the position of the neuroctoderm. In all cases, anterior is to the left and dorsal to the top. Scale bars in A and D, 50 μm.

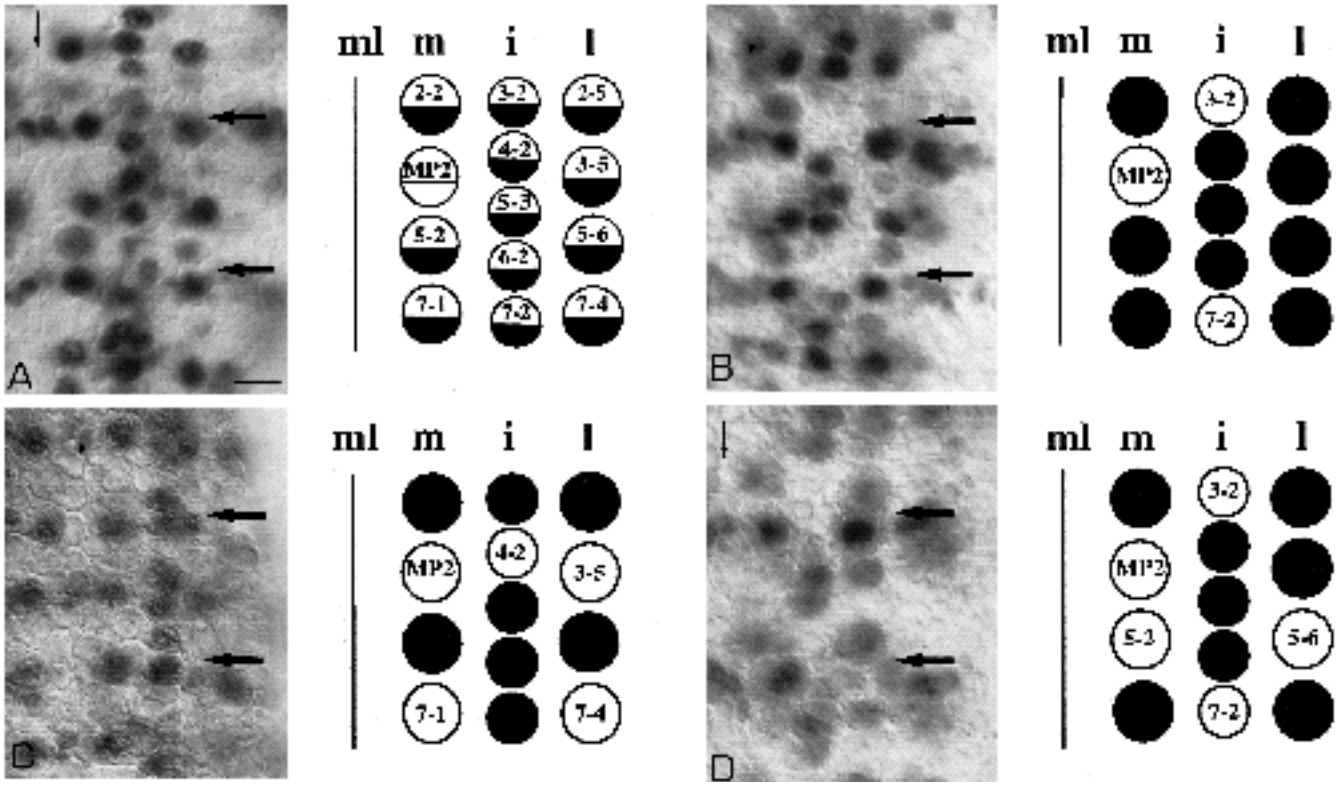

Fig. 4. Expression of *l'sc* in the NBs of AS-C mutants. (A) *wt*; (B) *T(1;2)sc*^{19}; (C) *In(1)sc*^{4}; (D) *T(1;4)sc*^{H}. For each genotype, the corresponding photograph shows a ventral view of a portion of the NB layer, after segregation of S1 and S2 NBs, stained with a l'sc antibody. Thick arrows indicate the limits of one segment. A thin arrow marks the position of the midline. To the right of each photograph, a diagram of circles represents the S1 and S2 NBs. For the wild type, the upper part of each circle shows the number of the corresponding NB, named according to Doe (1992) after their approximate NB in grasshopper. The lower part indicates whether the NB expresses (full) or not (empty) *l'sc*, disregarding the level of expression, which is actually different and somewhat variable for each particular NB (Martín-Bermudo et al., 1991). All NBs, except MP2, express *l'sc*. For the mutants, empty circles indicate those NBs that do not express *l'sc*, except for an occasional faint staining first detected at stage 9 (see text). All other NBs (full circles) show approximate normal levels of L'SC accumulation. The identification of NBs, particularly those of the intermediate row whose position is somewhat variable, has been aided in several instances by simultaneously staining with engrailed antibodies that labels NBs 6-2, 7-1, 7-2 and 7-4 (Doe, 1992). Anterior is to the top. ml, midline; m, i, l, medial, intermediate and lateral rows of NBs, respectively. Scale bar in A, 20 μm.

promoter (Fig. 5A). For the sake of simplicity, we have only examined L'SC protein accumulation in the NBs that delaminate during the first two waves of segregation and in their corresponding proneural clusters. All these NBs, except one, accumulate L'SC to different extents in the wild type (Martín-Bermudo et al., 1991; and Fig. 4A). Fig. 4B-D shows L'SC accumulation in a few of the cases analyzed. In most mutants, *l'sc* expression is abolished in a particular subset of proneural clusters and corresponding NBs. Some light staining in these NBs and overlying ectoderm is first detected by late stage 9, but always at much lower levels than in the wild type. This late expression is similar to that found in *Df(1)sc*^{4L}*sc*^{9R} embryos carrying *l'sc* transgenes (Fig. 3A-B). Table 2 summarizes the analysis of 22 AS-C mutations. Two main features stand out: first, mutations that map proximal or distal to the *l'sc* structural gene suppress its expression in different subsets of NBs; and second, NBs affected by any given breakpoint are also affected by all other breakpoints that map between it and the structural gene. These results strongly suggest the presence, at both sides of the transcribed sequence, of discrete *cis*-regulatory regions that activate *l'sc* in specific proneural clusters and, as a consequence, in their corresponding NBs.

The data allow definition of at least five regions containing *cis*-regulatory elements that drive *l'sc* expression in twelve proneural clusters (Fig. 5). Three *l'sc*-expressing NBs, NBs 3-5, 7-1 and 7-4, also express *ac* and *sc* (Martín-Bermudo et al., 1991; Skeath et al., 1992; Ruiz-Gómez and Ghysen, 1993). Interestingly, control regions A and B, required for wild-type expression of *l'sc* in these NBs, also seem to contain enhancer elements for *ac* (and probably *sc*) expression in the same NBs (Skeath et al., 1992, for region A; and Ruiz-Gómez, unpublished observations, for region B). The possibility thus arises that, in these NBs, either *l'sc* activates *ac* and/or *sc*, or vice versa. In fact, in imaginal discs, *ac* and *sc* can self-activate and cross-activate each other (Martínez and Modolell, 1991; Skeath and Carroll, 1991). *Df(1)sc*^{4L}*sc*^{9R}, which removes *l'sc*, still allows *ac/sc* expression in these NBs, demonstrating that their initial activation is *l'sc*-independent, although *l'sc* appears to enhance the level of *ac* expression in NBs 7-1 and 7-4 (Skeath et al., 1992; and our own observations). Moreover, NBs 7-1 and

Table 2. Expression of *lethal of scute* in neuroblasts of ASC mutations

	Neuroblasts				
Mutation	3-5	4-2, 7-1, 7-4	2-2, 2-5, 5-3, 6-2	5-2, 5-6	3-2, 7-2
In(1)y^{3P} *In(1)sc^{V2}* *sc$^{10.1}$* *sc^{L3}* *Df(1)sc^{5}*	+	+	+	+	+
In(1)sc^{8} *Df(1)y^{3PL}sc^{8R}* *In(1)sc^{L8}* *In(1)sc^{S1}*	−	+	+	+	+
In(1)sc^{4} *T(1;3)sc$^{260.15}$* *Df(1)y^{3PL}sc^{4R}* *Df(1)sc^{8L}sc^{4R}*	−	−	+	+	+
In(1)sc$^{260.14}$	+	+	−	−	−
T(1;4)sc^{H} *In(1)sc^{9}* *T(1;2)sc^{S2}* *In(1)sc^{7}* *In(1)sc^{29}*	+	+	+	−	−
In(1)sc$^{260.22}$ *T(1;2)sc^{19}* *Df(1)sc^{2}*	+	+	+	+	−

Summary of the analysis of *l'sc* expression in NBs, and corresponding proneural clusters, in different ASC mutations. The molecular mapping of the mutations is shown in Fig. 5. Mutants that display the same pattern of expression are grouped together.

7-4 express *l'sc*, but not *ac*, in *In(1)sc^{S1}* mutants (Table 2 and data not shown), indicating that *l'sc* expression in these two NBs is at least *ac* independent. The same can be suggested for NB 3-5, for, in *In(1)sc^{V2}*, it expresses *l'sc* normally (Table 2), but not *ac* (Skeath et al., 1992). Therefore, regions A and B may contain independent enhancers for *ac/sc* and *l'sc*, or, alternatively, the same enhancers act simultaneously on the promoters of the three genes.

DISCUSSION

We have shown that two transduced AS-C T3 genes, containing the transcribed region and relatively short flanking sequences, rescue the lethality of a *l'sc* deficiency, albeit with low penetrance. Likewise, we have shown that the smaller transgene rescues to a great extent the neural phenotype associated with larger deletions of the AS-C. These phenotypes are largely due to the lack of the *l'sc* function (Jiménez and Campos-Ortega, 1987). Therefore, our results unequivocally demonstrate, confirming previous indirect evidence (see Introduction), that the AS-C T3 gene and the genetically defined *l'sc* function are the same entity.

Remarkably, *l'sc* mutants can reach adulthood, even though the amounts of L'SC protein provided by the transgenes are substantially lower than normal and its distribution does not conform to the wild-type pattern. This indicates that the deployment of specific quantities of L'SC is not essential for its function. Moreover, in many AS-C mutations *l'sc* expression is practically abolished in specific NBs and flies are still viable. These data are consistent with the view that *l'sc* function is partially dispensable and that it can be largely replaced by that of other proneural genes (Martín-Bermudo et al., 1991). This in turn implies that the function of the different proneural genes is essentially the same, namely, to confer upon groups of ectodermal cells a capacity to become neural precursors. Only the extent of expression of each gene, as measured both by the number of proneural clusters that express it and by the level of its expression in each cluster, would establish its relative importance for the proneural function. On the other hand, if the *l'sc* function is to a great extent replaceable, this argues against its participation, together with other AS-C genes, in the specification of particular NB fates, according to a proposed model based in a combinatorial mechanism of action (Cabrera et al., 1987; Cabrera, 1992).

cis-control of *l'sc*

The wild-type expression of *l'sc* undergoes dynamic changes during embryonic development, suggesting that its control is complex (Cabrera, 1990; Martín-Bermudo et al., 1991). Indeed, the analysis of *l'sc* expression in AS-C mutations has revealed at least five *cis*-controlling regions scattered throughout the gene complex. These regions contain the presumptive enhancers that activate *l'sc* in twelve proneural clusters that give rise to S1 and S2 NBs. Presumably, other enhancers exist to drive *l'sc* expression in the proneural clusters that give rise to later segregating NBs and to sensory mother cells (SMCs) of the peripheral nervous system, where *l'sc* is also expressed (Romani et al., 1987; and our observations). What is the need for so many *cis*-regulatory elements acting on the *l'sc* promotor? The reason may be the positional and temporal accuracy with which neural precursors delaminate from the neuroectoderm. Each precursor arises from a small, proneural group of cells defined by the activation, in response to positional cues, of one or more proneural genes. These positional cues should be different, at least in part, for every particular cluster. As *l'sc* expression helps define most of the proneural clusters within the neuroectoderm, its promotor should respond to a large variety of cues and, hence, the need for a substantial number of enhancer elements.

It seems paradoxical that, despite such complex regulation, most NBs can still segregate in the absence of *l'sc* (Martín-Bermudo et al., 1991). A most likely explanation is the expression of other proneural genes in the same proneural clusters. For example, *ac* and *sc* may partially substitute for the absence of *l'sc* in the proneural clusters of NBs 3-5, 7-1 and 7-4, where these three genes are normally coexpressed. However, there is no correspondence between the frequency with which a particular NB is absent in a *l'sc* deficiency and the extent of AS-C gene expression in that NB. Again as an example, NB 7-1 is the one most commonly absent in *Df(1)sc^{4L}sc^{9R}* embryos (Martín-Bermudo et al., 1991), even though it coexpresses *ac/sc*, whereas NB 5-2, which expresses *l'sc*, but neither *ac* nor *sc*, segregates normally in the same mutant embryos. This indicates that the requirement for *l'sc* is not necessarily higher in those NBs where *ac* and *sc* are not expressed. It further suggests,

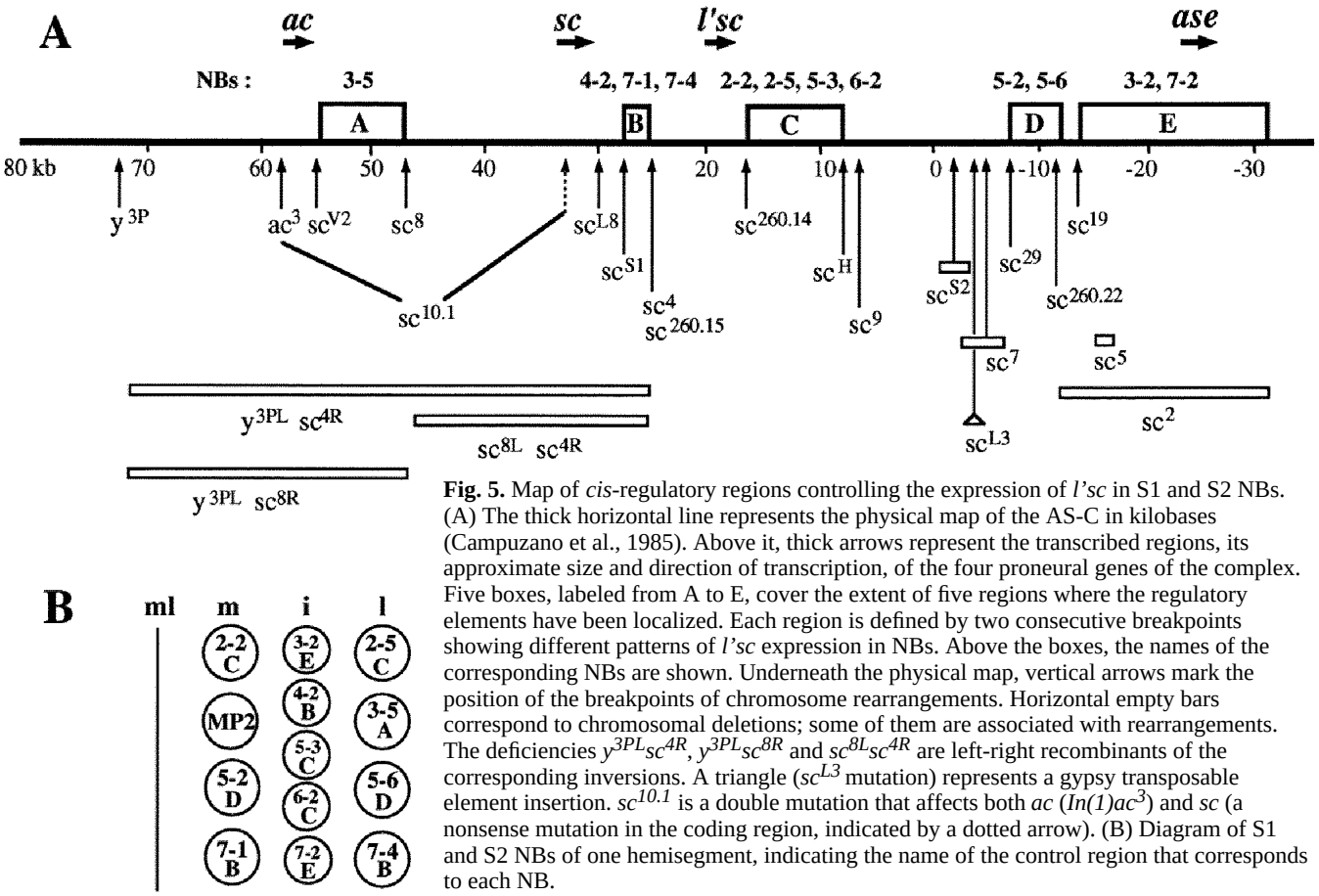

Fig. 5. Map of *cis*-regulatory regions controlling the expression of *l'sc* in S1 and S2 NBs. (A) The thick horizontal line represents the physical map of the AS-C in kilobases (Campuzano et al., 1985). Above it, thick arrows represent the transcribed regions, its approximate size and direction of transcription, of the four proneural genes of the complex. Five boxes, labeled from A to E, cover the extent of five regions where the regulatory elements have been localized. Each region is defined by two consecutive breakpoints showing different patterns of *l'sc* expression in NBs. Above the boxes, the names of the corresponding NBs are shown. Underneath the physical map, vertical arrows mark the position of the breakpoints of chromosome rearrangements. Horizontal empty bars correspond to chromosomal deletions; some of them are associated with rearrangements. The deficiencies $y^{3PL}sc^{4R}$, $y^{3PL}sc^{8R}$ and $sc^{8L}sc^{4R}$ are left-right recombinants of the corresponding inversions. A triangle (sc^{L3} mutation) represents a gypsy transposable element insertion. $sc^{10.1}$ is a double mutation that affects both *ac* ($In(1)ac^{3}$) and *sc* (a nonsense mutation in the coding region, indicated by a dotted arrow). (B) Diagram of S1 and S2 NBs of one hemisegment, indicating the name of the control region that corresponds to each NB.

as previously noted (Jiménez and Campos-Ortega, 1990), that the AS-C genes do not account for all the proneural activity in the CNS. If *l'sc* can be replaced in many sites by other proneural genes, why is it still expressed in most proneural clusters? We suggest that the phenotypic criteria so far used to characterize AS-C mutations, namely, neuroblast segregation and viability, probably do not reveal the full extent of the *l'sc* function. It has been shown for the imaginal SMCs that the levels of proneural products and, in some cases, the specific combination of these proteins can be critical for the correct implementation of the neural differentiation programme (Domínguez and Campuzano, 1993). Proneural insufficiency may be compatible with SMC emergence, but it can cause malformations of the resulting sensory organ. Similarly, *l'sc* insufficiency may subtly damage the NBs progeny and still be compatible with fly viability.

On the other hand, there may be sites where *l'sc* expression is truly redundant. Expression in these sites could be evolutionarily maintained by constraints imposed by the specificity and organization of the enhancers within the AS-C. If an enhancer directs expression in more than one site or if it interacts with other enhancers, so that the final expression is not merely the sum of the individual expressions, it may be impossible to delete expression in sites where *l'sc* is entirely redundant without removing it from places where it is necessary. Thus, these enhancers and the patterns of expression they direct would have been conserved in evolution despite the true redundant character of the expression at some sites.

It is likely that the same positional cues activate different genes in the same proneural cluster. In agreement with this, the *cis*-acting elements that control *ac* and *l'sc* expression in NBs 3-5, 7-1 and 7-4 are found in the same AS-C regions and most probably are the same for both genes. It is interesting that other *l'sc*-expressing NBs (for example, NBs 5-2 and 5-6), whose corresponding enhancers are located proximal to the *l'sc* transcription unit, do not express *sc* (Skeath et al., 1992), although this gene is activated, in the wing disc, by enhancers located in the same region. And, conversely, *l'sc* is not activated in the proneural clusters of the wing disc by those *sc*-specific enhancers (S. Romani, unpublished observations). This indicates that the *sc* and *l'sc* promoters have different responsiveness to the set of enhancers that are intermingled within the proximal region of the AS-C, or that the same enhancers alternatively activate one or the other promoter depending on the tissue. Strong candidates to activate the *cis*-regulatory elements that drive expression of AS-C genes are the axis-patterning genes that operate early in embryonic development. Thus, the pattern of *l'sc* expression in the neuroectoderm undergoes specific modifications in mutants for the pair-rule genes (Martín-Bermudo et al., 1991) and that of *ac* is set up by these and the dorsoventral coordinate genes (Skeath et al., 1992).

Another level of regulation of *l'sc* expression is suggested by the presence, upstream from the origin of transcription,

of a cluster of E-boxes, putative binding sites for bHLH transcription factors. In the wild type, the initial *l'sc* expression, most likely mediated by the enhancers referred to above, is followed by a progressive accumulation of the protein in the proneural clusters and in the delaminating NBs (Martín-Bermudo et al., 1991). Such accumulation could be explained by a positive feedback regulation of the L'SC protein on its own promoter or by cross-activation by other bHLH proteins. Indeed, E-boxes participate in the *ac* self-stimulation and its activation by *sc* that occur in the imaginal discs (Martínez and Modolell, 1991; Van Doren et al., 1992; Martínez et al., 1993). The late and weak L'SC accumulation in NBs and in neuroectodermal cells observed in *l'sc*⁻ embryos carrying a transgene, or in NBs of AS-C mutants, might also be explained by similar positive regulation involving the E-boxes that would amplify a basal *l'sc* expression.

We thank our laboratory colleagues for constructive criticism of the manuscript and to A. González-Reyes for help in computer drawing of Fig. 5. MDMB, FG and MD were predoctoral fellows from the Ministerio de Educación y Ciencia, SR was a predoctoral fellow of Instituto de Cooperación Iberoamericana and MRG was a postdoctoral fellow of the Consejo Superior de Investigaciones Científicas. This work was supported by grants from Dirección General de Investigación Científica y Técnica (PB90-0082 to FJ, and PR84-0212 and PB90-0085 to JM), Comunidad Autónoma of Madrid to JM, and an institutional grant from Fundación Ramón Areces to the Centro de Biología Molecular.

REFERENCES

Alonso, M. C. and Cabrera, C. V. (1988). The *achaete-scute* gene complex of Drosophila melanogaster comprises four homologous genes. *EMBO J.* **7**, 2585-2591.

Bensi, G., Rangei, G., Klefenz, H. and Cortese, R. (1985). Structure and expression of the human haptoglobin locus. *EMBO J.* **4**, 119-126.

Birnstiel, M. L., Busslinger, M. and Strub, K. (1985). Transcription termination and 3′ processing: the end is in sight! *Cell* **41**, 349-359.

Blackwell, K. and Weintraub, H. (1990). Differences and similarities in DNA-binding preferences of MyoD and E2A protein complexes revealed by binding site selection. *Science* **250**, 1104-1110.

Cabrera, C. V. (1990). Lateral inhibition and cell fate during neurogenesis in *Drosophila* - The interactions between *Scute*, *Notch* and *Delta*. *Development* **110**, 733-742.

Cabrera, C. V. (1992). The generation of cell diversity during early neurogenesis in *Drosophila*. *Development* **115**, 893-901.

Cabrera, C. V., Martínez-Arias, A. and Bate, M. (1987). The expression of three members of the *achaete-scute* gene complex correlates with neuroblast segregation in *Drosophila*. *Cell* **50**, 425-433.

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

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

Dambly-Chaudière, C. and Ghysen, A. (1987). Independent subpatterns of sense organs require independent genes of the *achaete-scute* complex in *Drosophila* larvae. *Genes Dev.* **1**, 297-306.

Devereux, J., Haeverli, P. and Smithies, O. (1984). A comprehensive set of sequence analysis programs for the VAX. *Nucl. Acids Res.* **12**, 387-395.

Doe, C. Q. (1992). Molecular markers for identified neuroblasts and ganglion mother cells in the *Drosophila* central nervous system. *Development* **116**, 855-863.

Domínguez, M. and Campuzano, S. (1993). *asense*, a member of the *Drosophila achaete-scute* complex, is a proneural and neural differentiation gen. *EMBO J. (in press).*

García-Bellido, A. (1979). Genetic analysis of the *achaete-scute* system of *Drosophilamelanogaster*. *Genetics* **91**, 491-520.

García-Bellido, A. and Santamaría, P. (1978). Developmental analysis of the *achaete-scute* system of *Drosophila melanogaster*. *Genetics* **88**, 469-486.

González, F., Romani, S., Cubas, P., Modolell, J. and Campuzano, S. (1989). Molecular analysis of the *asense* gene, a member of the *achaete-scute* complex of *Drosophila melanogaster*, and its novel role in optic lobe development. *EMBO J.* **8**, 3553-3562.

Henikoff, S. (1984). Unidirectional digestion with exonuclease III creates breakpoints for DNA sequencing. *Gene* **28**, 351-359.

Ip, Y. T., Park, R. E., Kosman, D., Bier, E. and Levine, M. (1992). The *dorsal* gradient morphogen regulates stripes of *rhomboid* expression in the presumptive neuroectoderm of the *Drosophila* embryo. *Genes Dev.* **6**, 1728-1739.

Jan, L. Y. and Jan, Y. N. (1982). Antibodies to horseradish peroxidase as specific neuronal markers in *Drosophila* and grasshoper embryos. *Proc. Natl. Acad. Sci. USA* **79**, 2700-2704.

Jiménez, F. and Campos-Ortega, J. A. (1979). A region of the Drosophila genome necessary for central nervous development. *Nature* **282**, 1398-1402.

Jiménez, F. and Campos-Ortega, J. A. (1987). Genes of the subdivision 1B of the genome of *Drosophila melanogaster* and their participation in neural development. *J. Neurogenet.* **4**, 179-200.

Jiménez, F. and Campos-Ortega, J. A. (1990). Defective neuroblast commitment in mutants of the achaete-scute complex and adjacent genes of D. melanogaster. *Neuron* **5**, 81-89.

Karess, R. E. and Rubin, G. M. (1984). Analyses of P transposable elements functions on *Drosophila. Cell* **38**, 135-146.

Kosman, D., Ip, Y. T., Levine, M. and Arora, K. (1991). Establishment of the mesoderm-neuroectoderm boundary in the *Drosophila* embryo. *Science* **254**, 118-122.

Lindsley, D. L. and Zimm, G. G. (1992). *The Genome of* Drosophila melanogaster. San Diego, USA: Academic Press.

Maniatis, T., Fritsch, E. F. and Sambrook, J. (1982). *Molecular Cloning. A Laboratory Manual.* New York: Cold Spring Harbor Laboratory.

Marí-Beffa, M., de Celis, J. F. and García-Bellido, A. (1991). Genetic and developmental analyses of chaetae pattern formation in *Drosophila* tergites. *Roux's Arch. Dev. Biol.* **200**, 132-142.

Martín-Bermudo, M. D., Martínez, C., Rodríguez, A. and Jiménez, F. (1991). Distribution and function of the *lethal of scute* gene product during early neurogenesis in *Drosophila*. *Development* **113**, 445-454.

Martínez, C. and Modolell, J. (1991). Cross-regulatory interactions between the proneural *achaete* and *scute* genes of *Drosophila*. *Science* **251**, 1485-1487.

Martínez, C., Modolell, J. and Garrell, J. (1993). Regulation of the proneural gene *achaete* by HLH proteins. *Molec. Cell. Biol.in press*.

Melton, D. A., Krieg, P. A., Rebagliati, M. R., Maniatis, T., Zinn, K. and Green, M. R. (1984). Efficient in vitro synthesis of biologically active RNA and RNA hybridization probes from plasmids containing a bacteriophage SP6 promoter. *Nucl. Acids Res.* **12**, 7035-7056.

Messing, J., Gronenborn, B., Muller-Hill, B. and Hofschneider, P. (1977). Filamentous coliphage M13 as a cloning vehicle: Insertion of a *Hin*dII fragment of the lac regulatory region in M13 replicative form in vitro. *Proc. Natn. Acad. Sci. USA* **74**, 3642-3646.

Muller, H. J. (1955). On the relation between chromosome changes and gene mutations. *Brookhaven Symp.* **8**, 126-147.

Romani, S., Campuzano, S. and Modolell, J. (1987). The *achaete-scute* complex is expressed in neurogenic regions of *Drosophila* embryos. *EMBO J.* **6**, 2085-2092.

Rubin, G. M. and Spradling, A. C. (1982). Genetic transformation of *Drosophila* with transposable element vectors. *Science* **218**, 348-353.

Ruiz-Gómez, M. and Ghysen, A. (1993). The expression and role of a proneural gene, *achaete*, in the development of the larval *Drosophila* nervous system. *EMBO J.* **12**, 1121-1130.

Ruiz-Gómez, M. and Modolell, J. (1987). Deletion analysis of the *achaete-scute* locus of D. melanogaster. *Genes Dev.* **1**, 1238-1246.

Sanger, F., Nickleen, S. and Coulson, A. R. (1977). DNA sequencing with chain-terminating inhibitors. *Proc. Natl. Acad. Sci. USA* **74**, 5463-5467.

Skeath, J. B. and Carroll, S. B. (1991). Regulation of *achaete-scute* gene expression and sensory organ pattern formation in the *Drosophila* wing. *Genes Dev.* **5**, 984-995.

Skeath, J. B., Panganiban, G., Selegue, J. and Carroll, S. B. (1992). Gene regulation in two dimensions: the proneural *achaete* and *scute* genes are controlled by combinations of axis-patterning genes through a common intergenic control region. *Genes Dev.* **6**, 2606-2619.

Staden, R. (1980). A new computer method for the storage and manipulation of DNA gel reading data. *Nucl. Acids Res.* **8**, 3673-3694.

Tabor, S. and Richardson, C. C. (1987). DNA sequence analysis with a modified bacteriophage T7 DNA polymerase. *Proc. Natl. Acad. Sci*. **84**, 4767-4771.

Van Doren, M., Powell, P. A., Pasternak, D., Singson, A. and Posakony, J. W. (1992). Spatial regulation of proneural gene activity: auto- and cross-activation of *achaete* is antagonized by *extramacrochaete*. *Genes Dev*. **6**, 2592-2605.

Villares, R. and Cabrera, C. V. (1987). The *achaete-scute* gene complex of *D. melanogaster*: conserved domains in a subset of genes required for neurogenesis and their homology to *myc*. *Cell* **50**, 415-424.

White, K. (1980). Defective neural development in *Drosophila* embryos deficient for the tip of the X chromosome. *Dev. Biol.* **80**, 332-344.

(Accepted 6 April 1993)