The EMBO Journal vol.7 no.8 pp.2585 – 2591, 1988

The *achaete – scute* gene complex of *Drosophila melanogaster* comprises four homologous genes

Maria C.Alonso and Carlos V.Cabrera

MRC Laboratory of Molecular Biology, Hills Road, Cambridge CB2 2QH, UK

Communicated by A.Travers

We have determined the nucleotide sequence of two genes of the *achaete – scute* complex (AS-C) and show that they are homologous to two previously sequenced members of the same locus. These four genes are interspersed with other transcription units of unknown function. We also study the expression of one of these genes by *in situ* hybridization and compare it with the other three genes. We suggest that the complete function ascribed to the AS-C by genetic experiments is carried out by the four homologous genes. We discuss the possible biochemical function of the AS-C in neurogenesis in the light of the homologies of the four genes with the mammalian *myc* family.

Key words: achaete – scute/neurogenesis/homologies/*Drosophila*

Introduction

Neurogenesis in insects is initiated shortly after gastrulation when a subpopulation of ectodermal cells segregates within the neuroepithelium. These cells, the neuroblasts (Wheeler, 1891), are assigned a precise identity by their position. Each cell can be identified because it generates a unique family of neurones through an invariant cell lineage (Taghert *et al.*, 1984; Doe and Goodman, 1985). We have proposed that the assignment of each neuroblast identity is under the control of a family of genes, some of whose members are contained within the *achaete – scute* gene complex (AS-C) of *Drosophila* (Cabrera *et al.*, 1987; Villares and Cabrera, 1987). This view is supported by three lines of evidence. Firstly, the loss of function alleles of the AS-C produce the lack of neural elements in both the central and peripheral nervous systems (Muller, 1955; Garcia-Bellido, 1979; Jiménez and Campos-Ortega, 1979, 1987; White, 1980; Cabrera *et al.*, 1987; Dambly-Chaudière and Ghysen, 1987). Conversely, the dominant alleles of the AS-C provoke the ectopic appearance of sensory organs (Garcia-Bellido, 1981; Garcia-Alonso and Garcia-Bellido, 1986). Secondly, the expression of three members (T_3, T_4 and T_5) of the AS-C precedes and parallels the segregation of neuroblasts in the early embryo (Cabrera *et al.*, 1987; Romani *et al.*, 1987). Finally, a deficiency for the AS-C prevents the appearance of at least one identified class of neuroblast (Cabrera *et al.*, 1987).

Sequence analysis of two genes of the AS-C revealed the existence of two highly conserved domains in their protein-coding regions (Villares and Cabrera, 1987). In addition, cross-hybridization experiments suggested the existence of other regions of homology within the AS-C. This study provided an explanation for the complementary phenotypes obtained with non-overlapping deficiencies of the locus (Garcia-Bellido, 1979; Dambly-Chaudière and Ghysen, 1987). Further, it led to the proposition that the AS-C encodes several homologous polypeptides, representing a subset of a complex gene family (Villares and Cabrera, 1987). The general significance of this family of proteins has been recognized by the discovery of its homology with *myc* (Villares and Cabrera, 1987). Recently this has been emphasized by the finding of a new homologous mouse gene which can convert fibroblasts to myoblasts (Davis *et al.*, 1987).

In this paper we complete our initial study on the homologies in the AS-C by determining the nucleotide sequence of the remaining cross-hybridizing regions. In addition, we study the spatial patterns of expression of a new member of this family and compare these results with those of previously reported examples.

Results

The AS-C has been defined by means of a pseudo-allelic series of mutations affecting adult bristle development (Garcia-Bellido, 1979). Within the cloned region of the AS-C six transcription units $T_1 – T_5$ and an embryonic cluster (EC) have been described (Figure 1). These transcripts are scattered throughout the 110 kb of DNA proximal to the *yellow* locus (Campuzano *et al.*, 1985). We have reinvestigated the transcriptional activity of the AS-C by means of single-stranded RNA probes to maximize the sensitivity of the method. In this survey we found three new transcription units, T_7, T_8 and T_9 (Figure 1), whose developmental profiles, with peaks at embryonic and pupal stages, are shown in Figure 2.

We have recently obtained the complete nucleotide sequence for the T_4 and T_5 genes (Villares and Cabrera, 1987). This resulted in the finding that these two genes share two regions of homology. It further allowed us to prepare an oligonucleotide mix against those regions of the putative translation products of T_4 and T_5 showing the highest sequence conservation (10 amino acids of the basic domain, Figure 1). With these oligonucleotides we mapped further regions of homology within the cloned AS-C DNA (Villares and Cabrera, 1987). In that experiment two new regions of strong cross-hybridization were found within the T_3 and T_8 transcription units. In addition, weaker cross-hybridization was found in two other places, one lying within the T_9 transcription unit and another mapped to a 1-kb restriction fragment at position +57 (Figure 1), for which no transcript has been found [note that in Figure 6B of Villares and Cabrera (1987) the arrowheads pointing at T_8 and T_9 in the autoradiogram were mistakenly switched].

We have now further investigated these cross-hybridizations using subcloned fragments from the five different

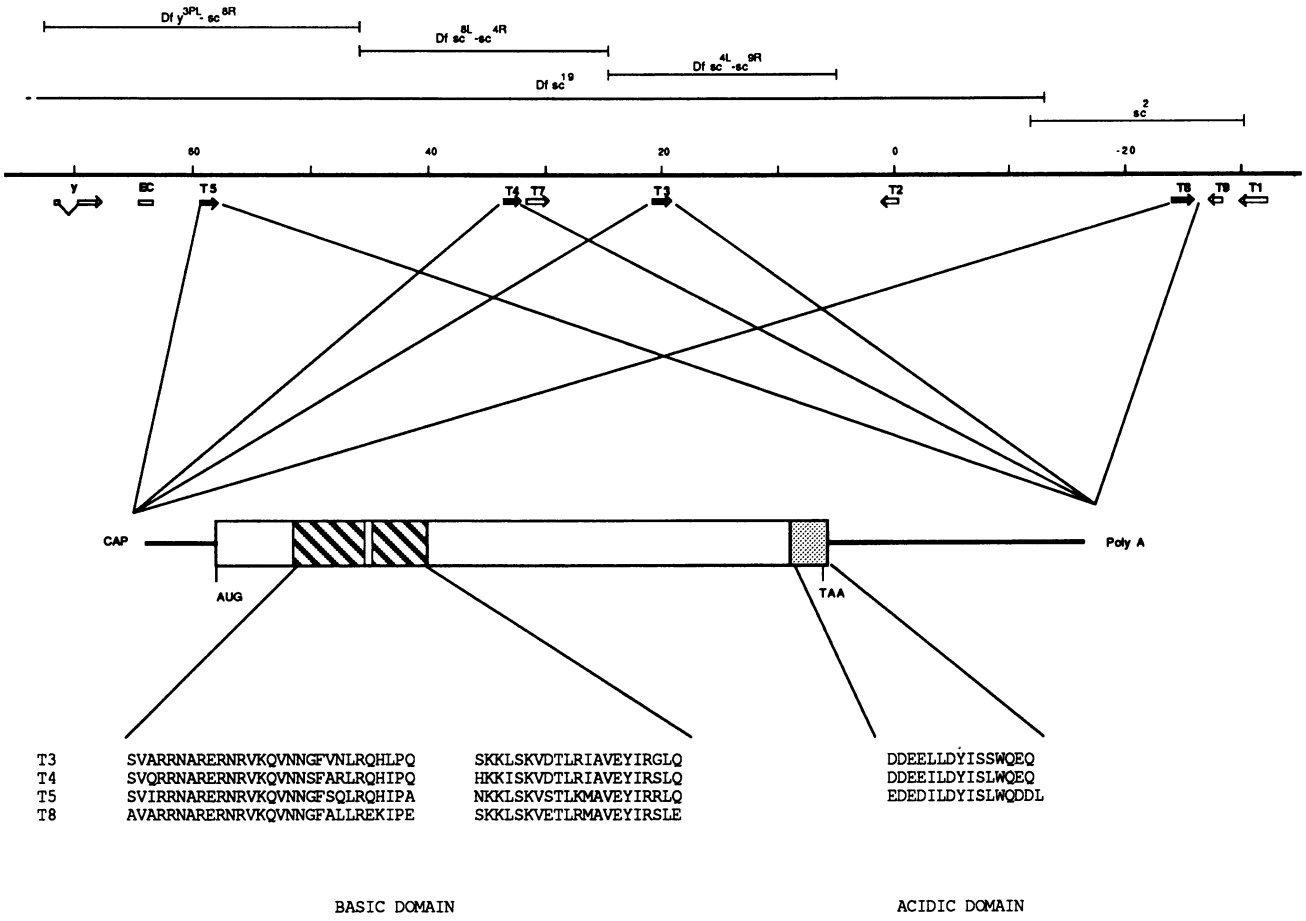

Fig. 1. Global structure of the AS-C. The uppermost lines depict the extent of the genetic deficiencies discussed in the text and later in this legend. Below, the physical map of the DNA is shown in kilobases. Underneath the map the different transcription units are represented by filled arrows (those sharing homology to one another) and empty arrows (the non-homologous). The transcripts are named T_1-T_9 and EC for embryonic cluster. T_4, T_5, T_7 and T_9 have peaks of expression in embryos and pupae; T_3, T_8 and the EC in embryos; T_2 in larvae and adults and T_1 in adults. *yellow* (y) represents the transcription unit codifying for this gene (Chia *et al.*, 1986). The typical structure of the homologous transcription units is depicted below. It shows the position of a domain rich (average 28%) in basic residues (striped), which is subdivided into two subregions of 29 and 21 amino acids joined by 13–19 residues with less or no homology (not shown). The overall conservation in these subregions is 93 and 95% respectively. The previously described homologies with the *myc* (Villares and Cabrera, 1987) and the MyOD1 proteins (Davis *et al.*, 1987) lie within this basic domain. At their C termini a region of 15 residues (shaded) appears in three out of the four cases. It shows a highly acidic environment (43%) surrounding a central tyrosine. This residue is likely to serve as a substrate for protein tyrosine kinase phosphorylation (Hunter and Cooper, 1985). The identity of each of these conserved domains is outlined on the left. The set of deficiencies depicted on the top have been used to define the functional components of the AS-C. Each deficiency removes a single homologous gene and on this basis T_5 might be responsible for the *achaete* function, T_4 for *scute* and T_3 for the *lethal of scute* (Garcia-Bellido, 1979; Jiménez and Campos-Ortega, 1979 and 1987; White, 1980; Dambly-Chaudière and Ghysen, 1987). The effect of removing T_8 has not been assessed yet with the sc^2 deletion. The combined effect of removing T_3, T_4 and T_5 with $Df(1)sc^{19}$ was studied at the single-cell level and shown to prevent the appearance of, at least, the single posterior median neuroblast (Cabrera *et al.*, 1987).

regions and have been able to reproduce only those belonging to the T_3, T_4, T_5 and T_8 transcripts (not shown). We therefore conclude that the weaker cross-reactions were possibly due to partial digests of the λ clones and that the only regions of homology are those stated above and summarized in Figure 1. In what follows we show by direct sequencing that this conclusion is true.

The sequence of the T_3 gene

A 3.2-kb *Eco*RI fragment (coordinates +19.9 to +17.8 of Figure 1) was shown to contain the T_3 transcription unit by cross-hybridization with a cDNA (Campuzano *et al.*, 1985). We prepared progressive deletions from this genomic fragment to obtain overlapping sequences of the T_3 gene. Figure 3A shows the 1845 nucleotides (nt) of sequence from the central region of this fragment. These sequences span the complete T_3 cDNA as both its ends are contained within. Further, the 1183 nt of the corresponding cDNA region seem to span the full length of the corresponding RNA product, as S1 protection experiments originated a fragment of that length (Figure 4A).

The genomic sequences obtained include 275 nt upstream and 385 nt downstream of the cDNA. At position 41 a putative TATA box motif is found (Breathnach and Chambon, 1981). At position 1319 the consensus polyadenylation signal (Proudfoot and Brownlee, 1976) precedes several G/T clusters, possibly involved in transcription termination (Birnstiel *et al.*, 1985).

The longest open reading frame obtained from the putative translation of this sequence is shown in Figure 3A. The first in-frame AUG fits the consensus for translation initiation (Kozak, 1984, 1986). The polypeptide translated from this

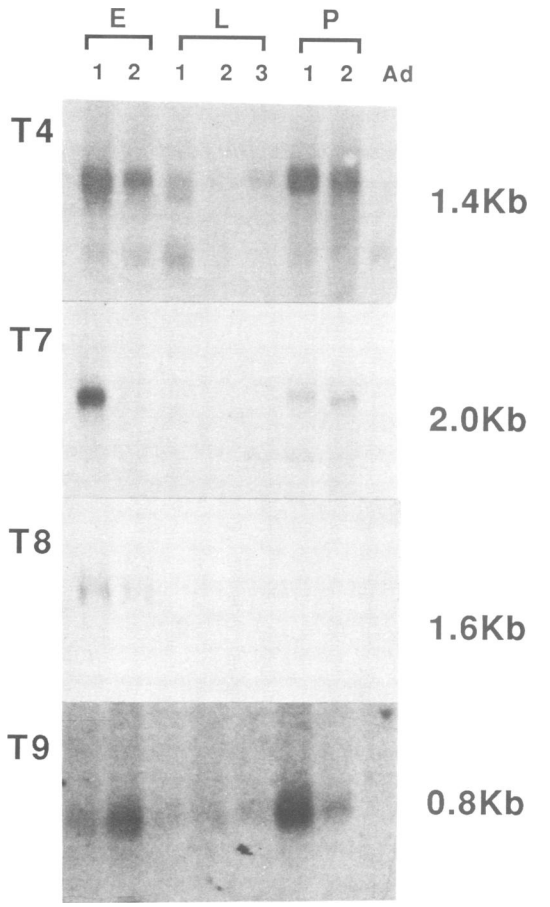

Fig. 2. RNA blot analysis. Poly(A)$^+$ RNA was prepared from (**E**) two embryonic stages, 0−12 h (**lane 1**) and 12−24 h (**lane 2**); (**L**) the three larval stages (1−3); (**P**) two pupal stages, 0- to 1-day-old pupae (**lane 1**) and 2- to 4-day old pupae (**lane 2**); (**Ad**) adults. Blots were hybridized with asymmetric RNA probes prepared from a cDNA (T_4) or diverse genomic fragments as follows: T_7 a 3.2-kb *Bam*HI−*Eco*RI fragment (around coordinate +30); T_8 a 1.65-kb *Hin*dIII fragment (coordinate −25); and T_9, a 2-kb *Bam*HI−*Eco*RI fragment (coordinate −28). The T_4 blot is shown as an example of a previously studied AS-C gene (Campuzano *et al.*, 1985). The relative mobilities of the different RNAs are indicated on the right side. The minor bands in the T_4 and T_7 blots are likely an artefact of the RNA probes, as they do not show when DNA probes are used.

point is 258 amino acids long and 29 kd in mol. wt.

From comparisons with the other two AS-C genes previously sequenced (Villares and Cabrera, 1987) we confirmed the existence of the homologies detected by cross-hybridization. Figure 1 summarizes these findings and shows that T_3 has the characteristic structure previously described for the T_4 and T_5 genes. That is, an N-terminal basic domain which we have subdivided into two subregions of 29 and 20 residues respectively, and a C-terminal acidic domain spanning 15 residues.

The sequence of the T_8 gene

The cross-hybridizations with mixed oligonucleotides described above allowed mapping of the potential coding region of the T_8 transcription unit around coordinate −25 (Figure 1). We therefore used a 1.65-kb *Hin*dIII fragment spanning this coordinate to obtain deletions for sequencing. In Figure 3B the 1600 nt of sequence from this genomic fragment are shown. The putative translation from the first in-frame AUG (Kozak, 1984, 1986) yields an open reading frame of 397 amino acids, which would produce a polypeptide of apparent mol. wt of 43.5 kd. Within this translated region lies the conserved basic domain, shown in Figure 1 in comparison with the others. We have not found, however, sequences homologous to the acidic domain.

As no cDNA is yet available for this transcription unit we performed S1 protection experiments (Figure 4B) to show that the sequence in Figure 3B contains most of the transcribed region. We therefore conclude that the T_8 gene, unlike T_3, T_4 and T_5 bears only one of the conserved domains which characterize this family (see Figure 1).

The pattern of expression of the T_8 gene

We have studied the pattern of expression of T_8 using as a probe the above described genomic fragment. RNA homologous to the T_8 probe becomes first detectable during extending germ band embryos (late stage 8, Figure 5a). At this time conspicuous segmentally arranged groups of cells accumulate label in a medio-lateral region of the metameric germ band between ectoderm and mesoderm. This region of the developing embryo has just engaged in neurogenesis, as the first neuroblasts segregate about this time (Hartenstein and Campos-Ortega, 1984). The principal characteristic of this initial expression is that the T_8 RNA appears in presumptive neural precursors, once they have segregated from the ectoderm. This is in clear contrast with the other homologous genes within the AS-C, whose expression is first detected earlier in the ectoderm and then also in the internalized neuroblasts (Cabrera *et al.*, 1987; Romani *et al.*, 1987).

Subsequently, expression of the T_8 gene builds up, as more intense label accumulates throughout the metameric germ band (Figure 5b). Stages 10 and 11 represent the peak of this gene's expression, the label always appearing between ectoderm and mesoderm (Figure 5c−e). As neuroblast segregation takes place in three waves (Hartenstein and Campos-Ortega, 1984), this suggests that the onset of T_8 transcription takes place reiteratively in neuroblasts after their segregation. As described before for the other homologous genes (Cabrera *et al.*, 1987), the pattern of T_8 expression in the metameric germ band is blotchy, as adjacent sections with strong label alternate with unlabelled ones.

In addition to the metameric germ band, at stage 11 conspicuous hybridization is detected in the primordia of stomatogastric and optic lobes nervous systems (Figure 5e,f). This signal prevails in subsequent stages, whereas that of the metameric germ band declines. Indeed, in Figure 5f an embryo at full retraction of the germ band shows a drastically reduced signal in the newly formed neuromeres in comparison with previous stages. At this time (stage 13, Figure 5f) the other homologous genes of the AS-C are no longer expressed in this region, but, as with T_8, they are clearly detected in the presumptive optic lobe region. In addition, T_8 is more prominent in the presumptive procephalic lobe, whereas T_3, T_4 and T_5 label strongly the presumptive stomatogastric nervous system.

In summary, the initiation of T_8 transcription is shifted both in time and space in relation to the other genes of the AS-C. This difference implies that the T8 promoter responds to specific positional cues for its expression, as previously suggested for the other genes of the AS-C (Cabrera *et al.*, 1987; Villares and Cabrera, 1987). Other details of the

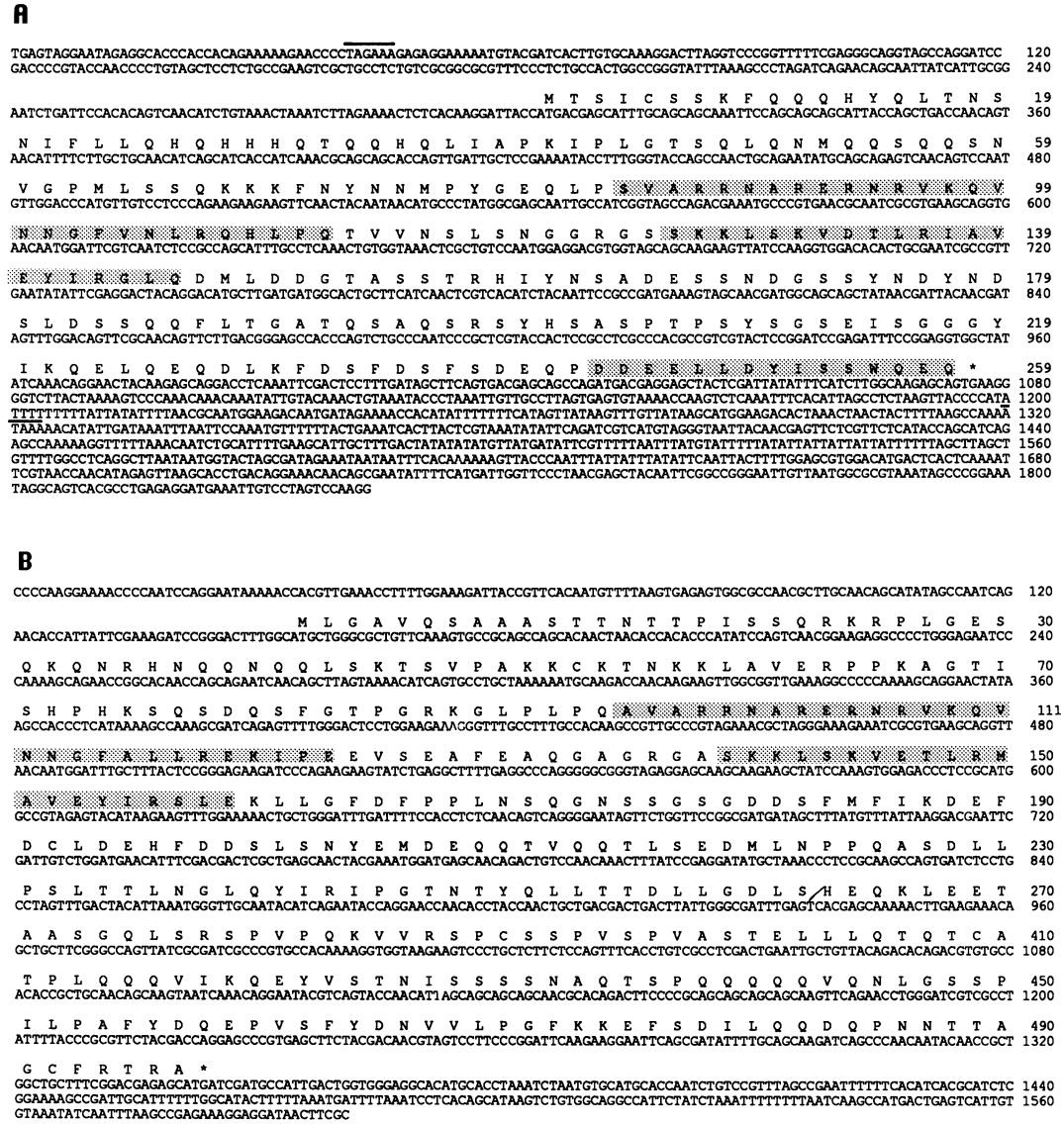

Fig. 3. Nucleotide sequences of the T_3 and T_8 regions and their conceptual translation products. (**A**) Genomic sequences of T_3. The regions of homology are shaded. The putative TATA box at position 41 and polyadenylation signal at position 1319 are overlined. The ends of the cDNA map at 277 and 1460 nt. (**B**) Genomic sequences of T_8. The shaded areas represent the regions homologous to domain I.

pattern of expression of this gene, like the cephalic and posterior midgut region and presumptive precursors of the peripheral nervous system, are outlined in the legend to Figure 5.

Discussion

The results presented in this report together with a previous one (Villares and Cabrera, 1987) show that the AS-C is composed of four homologous genes: T_3, T_4, T_5 and T_8. These genes span 90 kb of DNA of the locus and are intermingled with other non-homologous transcription units (Figure 1).

Two lines of evidence suggest that only the homologous genes are responsible for the complete AS-C function. Firstly, a point mutation in the T_4 gene has been correlated with the drastic phenotype of the $sc^{10\text{-}1}$ allele (Villares and Cabrera, 1987). As the AS-C is located at the tip of the X-chromosome (Garcia-Bellido, 1979), classical genetic analysis is hampered by the low recombination frequencies. Phenotypic analysis has therefore proceeded by the use of synthetic deficiencies created by the combination of two nearby breakpoints (Muller, 1955; Garcia-Bellido, 1979). As these deficiencies remove several transcription units at the time (see Figure 1), the correlation of a point mutation with a given phenotype was the first supporting evidence to associate the homologous genes with the AS-C function (Villares and Cabrera, 1987). Secondly, only the homologous genes have patterns of expression which correlate with neurogenesis (Cabrera *et al.*, 1987; Romani *et al.*, 1987 and Figure 5). We have investigated by *in situ* hybridization the expression of two of the non-homologous genes (T_7 and T_9, both of which show embryonic RNAs) and found no signal between blastoderm and retraction of the germ band (our unpublished observations).

The finding of the homologies between T_4 and T_5, together with the phenotypic analysis of their deficiencies, led to the proposal that both genes have identical biological

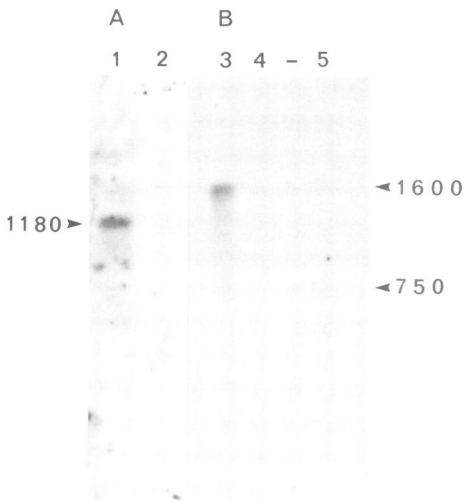

Fig. 4. S1 analysis of T_3 and T_8. The genomic fragments whose sequence is shown in Figure 2 were used in S1 protection experiments to test the structure of the T_3 and T_8 RNAs. (**A**) the 3.2-kb *Eco*RI fragment protects 1180 nt when hybridized to RNA from *Drosophila* embryos (**lane 1**) but it does not do so in the presence of *E.coli* tRNA (**lane 2**). (**B**) the 1.6-kb *Hin*dIII fragment is fully protected in the presence of embryonic RNA (**lane 3**). As trace amounts of putative self-hybridizing products are detected in the presence of *E.coli* tRNA (**lane 4**), a control experiment was carried out in parallel in which a 743-nt-long sense RNA synthesized with phage RNA polymerase was used for the protection. In **lane 5** it is shown that a fragment of this size is protected in addition to trace amounts of the 1.6-kb fragment.

function, the difference between them being the spatial domains in which they become active (Villares and Cabrera, 1987). This proposal was substantiated by the analysis of their patterns of expression. Indeed promoter-specific patterns in their spatial and temporal aspects were shown to be the rule for expression of those two genes, as well as for T_3, in the early embryo (Cabrera *et al.*, 1987). The present data support and extend these conclusions. In contrast to the other three homologous genes, T_8 is never expressed in epidermal cells, but the onset of its expression reiteratively occurs in neural precursors, as they become segregated from the ectoderm (Figure 5).

This finding suggests that the AS-C genes respond to two different sets of underlying signals. The early expressed genes (T_3, T_4 and T_5), whose patterns of expression develop in the ectoderm, might be controlled by the segmentation genes (Cabrera *et al.*, 1987). The late gene (T_8) might respond to the activity of the earlier acting AS-C genes.

Despite the spatial and temporal difference in their patterns of expression, the four AS-C genes share the common feature of being expressed in intimate association with neuroblast segregation and the subsequent development of the neuroblast lineage. This correlation has led to the proposal that these genes contribute in some way to elements of neuroblast identity (Cabrera *et al.*, 1987). This identity results in each neuroblast producing a characteristic family of neurones by an invariant cell lineage (Taghert *et al.*, 1984; Doe *et al.*,

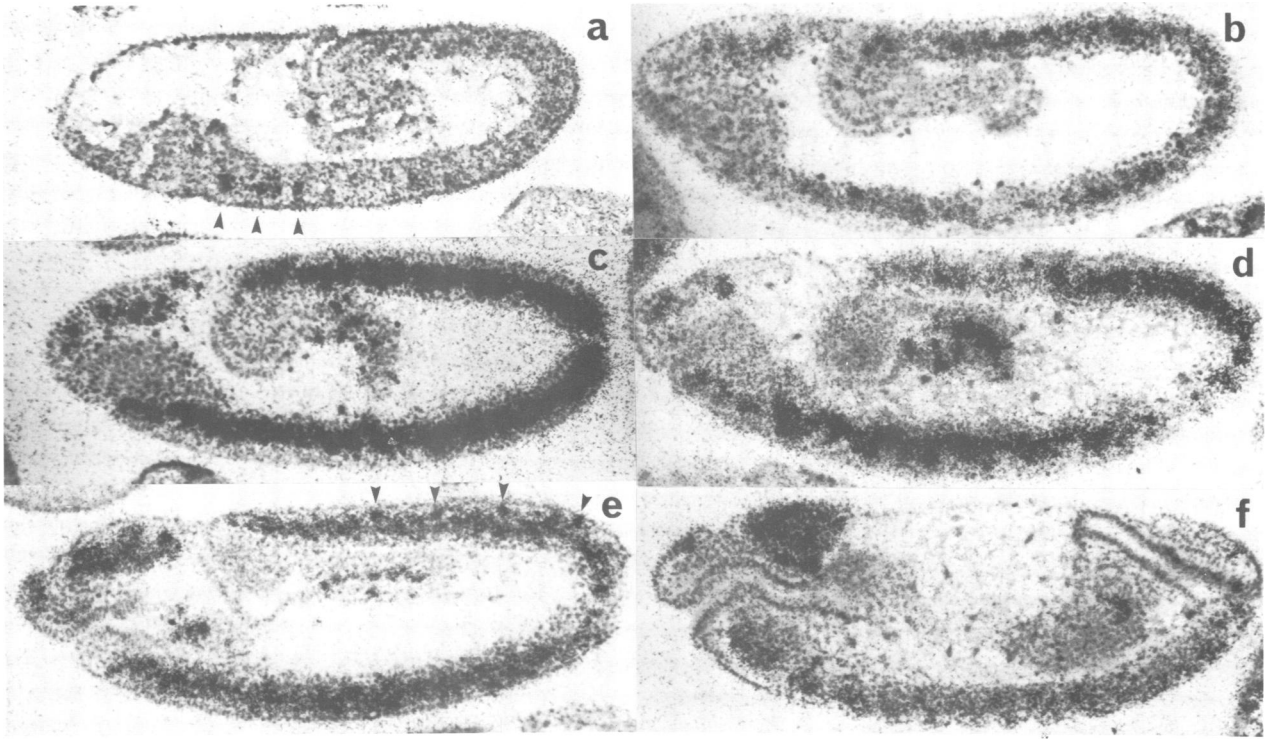

Fig. 5. The patterns of expression of T_8. Parasagittal sections through embryos at (**a**) mid-stage 8. Arrowheads point to groups of segmentally arranged cells which represent the onset of T_8 expression. Note the position of the silver grains between the ectoderm and mesoderm in the metameric germ band. (**b**) Stage 9 embryo showing the loss of segmental organization of the initial pattern. The label accumulates in the metameric germ band and certain areas of the cephalic region. (**c**) Stage 10 embryo, when expression reaches its peak. The distribution of silver grains is as in (b). (**d**) Late stage 10 embryo. The transient segmental bulges of the mesoderm affect the morhology of the underlying developing ventral cord and thus the distribution of the label in the metameric germ band. In addition, label appears in the posterior midgut region. (**e**) Stage 11 embryo. In addition to the previous accumulation of label in the metameric germ band and posterior midgut, new hybridization is detected in the presumptive procephalic and optic lobe regions, as well as in discrete groups of segmentally organized cells (arrowheads), possibly precursors of the peripheral nervous system. (**f**) Stage 13 embryo showing the declining accumulation of label in the newly formed neuromeres and its persistence in the procephalic and optic lobes presumptive regions.

1988). Specific events of gene activation, of the segmentation and selector gene families have been shown to occur at different steps of the neuroblast linear (Doe *et al.*, 1988). An attractive possibility is to consider this differential gene expression through the neuroblast lineage as a consequence of the activity of the *scute* gene family.

Testing this hypothesis is hampered at the moment by the lack of specific markers for the neuroblast lineage and by the inherent complexity of the *scute* gene family (Villares and Cabrera, 1987). However, two lines of evidence lend some support to it. Firstly, $Dfsc^{19}$ (see Figure 1) precludes expression of the *fushi tarazu* gene in, at least, precursors of the MP2 neurones (our unpublished observations). Secondly, the conserved basic domain of the AS-C genes is homologous to the *myc* protein family (Villares and Cabrera, 1987). The *myc* proteins have been shown to be capable of activating transcription in a yeast hybrid protein assay (Lech *et al.*, 1988). It is therefore possible that the *scute* proteins are transcription factors which act in a similar manner to *myc*. By this analogy we suggest that the basic domain would act in DNA binding and the acidic domain would play a role in transcription activation (Ma and Ptashne, 1987).

Materials and methods

The wild-type strains of *Drosophila* Canton S and Oregon R were used throughout.

RNA blotting

RNA was prepared from staged collections by homogenization in guanidinium isothiocyanate (Chirgwin *et al.*, 1979). The extract was clarified by low-speed centrifugation and the RNA pelleted through 5.7 M CsCl cushions (Glisin *et al.*, 1979). The pellet was dissolved in water and precipitated in 3 M sodium acetate at −20°C, dissolved again in water and precipitated with ethanol. The poly(A)$^+$ fraction of this RNA, obtained by chromatography in oligo(dT)−cellulose, was electrophoresed in formaldehyde−agarose gels and blotted in Hybond N (Thomas, 1983). Subclones of cDNAs or genomic fragments in the Bluescribe plasmid (Vector Labs) were used to prepare ^{32}P-labelled single-stranded probes with phage RNA polymerases (Melton *et al.*, 1984). The blots were hybridized with 10^7 d.p.m. of probe in 50% formamide/4 × SSC/0.1% SDS/1 × Denhardt's/10% dextran sulphate 50 μg/ml denatured calf thymus DNA and 50 μg/ml yeast RNA at 65°C for 16 h. Washes were in three changes of 0.1 × SSC/0.1% SDS at the same temperature. The asymmetry of the probes permitted the determination of the transcriptional orientation.

DNA sequencing

The genomic restriction fragments described in the text were subcloned in M13 vectors (Messing, 1983). Single-stranded DNA from these recombinant phages was purified and subjected to the manipulations described in Dale *et al.* (1985) to obtain progressive deletions. These clones were sequenced by the chain termination method (Sanger *et al.*, 1977) as described in Bankier and Barrel (1983). The sequences were assembled and analysed by the computer programs of Staden (1986).

S1 analysis

S1 protection experiments were performed as described by Maniatis *et al.* (1982). Briefly, 10 ng of DNA (the 3.2-kb *Eco*RI fragment for T_3 protections and the 1.6-kb *Hin*dIII fragment for T_8) were hybridized with 5 μg of *Drosophila* embryonic Poly(A)$^+$ RNA or *Escherichia coli* tRNA in 80% formamide/0.4 M NaCl/40 mM Pipes (pH 7.0)/1 mM EDTA at 54°C for 16 h (Cassey and Davidson, 1977). The samples were then diluted with 30 vol of 280 mM NaCl/30 mM sodium acetate (pH 4.5)/5 mM $ZnSO_4$/20 μg/ml denatured calf thymus DNA/250 units of S1 and incubated at 37°C for 30 min. Reactions were stopped by addition of one-quarter the volume of 2.5 M ammonium acetate/50 mM EDTA/20 μg *E.coli* tRNA and precipitated with 1 vol of isopropanol. The precipitates were run in alkaline agarose gels and blotted to nylon membranes. The blots were probed with the respective genomic fragments described above labelled with [α-^{32}P]dCTP by the oligo-priming method (Feinberg and Vogelstein, 1984).

In situ hybridization

Embryo wax sections were prepared and processed as described before (Ingham *et al.*, 1985) except that the SSC heating step was omitted. [^{35}S]RNA probes from the 1.6-kb *Hin*dIII fragment homologous to the T_8 transcription unit were prepared by standard methods (Melton *et al.*, 1984). Size reduction of the probe was acheived as described by Cox *et al.* (1984). Hybridization conditions and washes were as before (Cabrera *et al.*, 1987). The staging of the embryos throughout the text follows Campos-Ortega and Hartenstein (1985).

Acknowledgements

We are most indebted to A.A.Travers for constant encouragement and laboratory facilities. We thank P.A.Lawrence and A.A.Travers for critical reading of the manuscript. C.V.C. is on leave from the Spanish CSIC supported by an EMBO long-term fellowship.

References

Bankier,A.T. and Barrell,B.G. (1983) In Flavell,R.A. (ed.), *Techniques in the Life Sciences B5, Nucleic Acid Biochemistry B508*. Elsevier, Ireland, pp. 1−34.
Birnstiel,M.L., Busslinger,M. and Strub,K. (1985) *Cell*, **41**, 349−359.
Breathnach,R. and Chambon,P. (1981) *Annu. Rev. Biochem.*, **50**, 349−383.
Cabrera,C.V., Martinez-Arias,A. and Bate,M. (1987) *Cell*, **50**, 425−433.
Campos-Ortega,J.A. and Hartenstein,V. (1985) *The Embryonic Development of Drosophila melanogaster*. Springer-Verlag, Berlin.
Campuzano,S., Carramolino,L., Cabrera,C.V., Ruiz-Gómez,M., Villares,R., Boronat,A. and Modolell,J. (1985) *Cell*, **40**, 327−338.
Cassey,J. and Davidson,N. (1977) *Nucleic Acids Res.*, **4**, 1539−1552.
Chia,W., Howes,G., Martin,M., Meng,Y., Moses,K. and Tsubota,S. (1986) *EMBO J.*, **5**, 3597−3605.
Chirgwin,J., Acyble,A., McDonald,R. and Ruther,W. (1979) *Biochemistry*, **18**, 5294−5299.
Cox,K.H., De Leon,D.V., Angerer,L.M. and Angerer,R.C. (1984) *Dev. Biol.*, **101**, 485−502.
Dale,R.M.K., McClure,B.A. and Houchins,J.P. (1985) *Plasmid*, 13, 31−40.
Dambly-Chaudière,Ch. and Ghysen,A. (1987) *Genes Dev.*, **1**, 297−306.
Davis,R.L., Weintraub,H. and Lasser,A.B. (1987) *Cell*, **51**, 987−1000.
Doe,Ch.Q. and Goodman,C.S. (1985) *Dev. Biol.*, **111**, 206−219.
Doe,Ch.Q., Hiromi,., Gehring,W.J. and Goodman,C. (1988) *Science*, **239**, 170−175.
Feinberg,A.P. and Vogelstein,B. (1984) *Anal. Biochem.*, **137**, 266−267.
Garcia-Alonso,L. and Garcia-Bellido,A. (1986) *Roux's Arch. Dev. Biol.*, **195**, 259−264.
Garcia-Bellido,A. (1979) *Genetics*, **91**, 491−520.
Garcia-Bellido,A. (1981) In Lloyd,C.W. and Rees,D.E. (eds), *Cellular Controls in Differentiation*. Academic Press, New York, pp. 281−304.
Glisin,V., Crkvenjakov,R. and Bys,C. (1979) *Biochemistry*, **13**, 2633−2637.
Hartenstein,V. and Campos-Ortega,J. (1984) *Roux's Arch. Dev. Biol.*, **193**, 308−325.
Hunter,T. and Cooper,J.A. (1985) *Annu. Rev. Biochem.*, **54**, 897−930.
Ingham,P.W., Howard,K.R. and Ish-Horowicz,D. (1985) *Nature*, **318**, 439−445.
Jiménez,F. and Campos-Ortega,J. (1979) *Nature*, **282**, 310−312.
Jiménez,F. and Campos-Ortega,J. (1987) *J. Neurogenet.*, **4**, 179−200.
Kozak,M. (1984) *Nucleic Acids Res.*, **12**, 857−871.
Kozak,M. (1986) *Cell*, **44**, 282−292.
Lech,K., Anderson,K. and Brent,R. (1988) *Cell*, **52**, 179−184.
Ma,J. and Ptashne,M. (1987) *Cell*, **51**, 113−119.
Maniatis,T., Fritsch,E.F. and Sambrook,J. (1982) *Molecular Cloning: A Laboratory Manual*. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.
Melton,D.A., Krieg,P.A., Rebagliati,M.R., Maniatis,T., Zimm,K. and Green,M.R. (1984) *Nucleic Acids Res.*, **12**, 7035−7056.
Messing,J. (1983) *Methods Enzymol.*, **101**, 20−78.
Muller,H.J., (1955) *Brookhaven Symp.*, **8**, 126−147.
Proudfoot,N. and Brownlee,G. (1976) *Nature*, **263**, 211−213.
Romani,S., Campuzano,S. and Modolell,J. (1987) *EMBO J.*, **6**, 2085−2092.

Sanger,F., Nicklen,S. and Coulson,A.R. (1977) *Proc. Natl. Acad. Sci. USA*, **74**, 5463–5467.
Staden,R. (1986) *Nucleic Acids Res.*, **14**, 217–231.
Taghert,P.H., Doe,Ch.Q. and Goodman,C.S. (1984) *Nature*, **307**, 163–165.
Thomas,P. (1983) *Methods Enzymol.*, **100B**, 255–266.
Villares,R. and Cabrera,C.V. (1987) *Cell*, **50**, 415–424.
Wheeler,W.M. (1891) *J. Morphol.*, **4**, 337–343.
White,K. (1980) *Dev. Biol.*, **80**, 332–344.

Received on March 18, 1988; revised on May 9, 1988