MOLECULAR AND CELLULAR BIOLOGY, June 1993, p. 3514–3521
0270-7306/93/063514-08$02.00/0

Vol. 13, No. 6

Regulation of the Proneural Gene *achaete* by Helix-Loop-Helix Proteins

CARMEN MARTÍNEZ,† JUAN MODOLELL,* AND JOAN GARRELL‡

Centro de Biología Molecular "Severo Ochoa," Consejo Superior de Investigaciones Científicas and Universidad Autónoma de Madrid, Canto Blanco, 28049 Madrid, Spain

Received 8 January 1993/Returned for modification 3 March 1993/Accepted 24 March 1993

The Achaete (Ac) protein, a transcriptional regulator of the basic-helix-loop-helix (bHLH) type, confers upon ectodermal cells the ability to become neural precursors. Its temporally and spatially regulated expression, together with that of the related Scute (Sc) protein, helps define the pattern of *Drosophila melanogaster* sensory organs. We have examined the transcriptional control of the *ac* gene and shown, using in vivo assays, that several E-boxes, putative interacting sites for bHLH proteins, present in the *ac* promoter are most important for *ac* regulation. They most likely mediate *ac* self-stimulation and *sc trans*-activation. We also demonstrate that *ac* transcription is negatively regulated in vivo by the gene *extramacrochaetae* (*emc*) in a manner dependent on Ac and Sc products. *emc* encodes an HLH protein that lacks the basic region and presumably antagonizes Ac and Sc function by sequestering these proteins in complexes unable to interact with DNA. Our results strongly support the model of negative regulation of *emc* on *ac* and *sc* transcription through titration of their products. As currently thought, this seems accomplished by heterodimerization via the HLH domain, because an amino acid substitution in this region abolishes the *emc* antagonistic effect both in vitro and in vivo.

The *achaete-scute* complex (AS-C) plays a key role in the development of the central and peripheral nervous systems of *Drosophila melanogaster* (reviewed in references 5 and 16). Its protein products confer to ectodermal cells the ability to become neural precursors. The AS-C encodes four of these proneural proteins. Two of them, Achaete (Ac) and Scute (Sc), are most important for the generation of the peripheral nervous system of the adult fly; in their absence, most of its cuticular sensory organs (SOs) do not develop (11). These proteins and Daughterless (Da), another proneural protein encoded outside of the AS-C, contain the basic-helix-loop-helix (bHLH) domain characteristic of a family of transcriptional regulators (22). The HLH domain mediates heterodimerization between proteins containing it (21), and the presence of the basic region in both members of the heterodimer allows its binding to DNA (8). On the basis of in vitro binding assays of combinations of proneural proteins to synthetic oligonucleotides (4, 22, 28, and our unpublished observations), it has been proposed that the AS-C proteins form heterodimers with Da and that these dimers regulate the transcription of a set of genes required for a cell to become a neural precursor (reviewed in reference 14). Indeed, L'sc, another proneural protein encoded in the AS-C, in combination with Da acts as transcriptional activator in a *Saccharomyces cerevisiae* model system (4).

Development of most of the adult SOs requires expression of the *ac* and *sc* genes in the third larval instar and early pupal stages, when ectodermal cells of the imaginal discs and abdominal histoblasts, the primordia of most of the adult epidermis, become SO precursors (13). Analyses of wing imaginal discs, the anlagen of most of the thorax and wings, have shown that *ac* and *sc* are coexpressed in groups of cells, the proneural clusters, that occupy highly specific positions. SO precursors are selected among the cells of these clusters (6, 23, 27). The spatial distribution of proneural clusters is one of the constraints that help define the position of SOs. A 0.8-kb DNA fragment obtained from the region immediately upstream from the *ac* coding sequences directs, in transformed flies, *lacZ* expression in a pattern of clusters very similar to that of the wild-type *ac* gene (19). This fragment mediates both *ac* self-activation and activation by *sc* and also responds to other local cues. These regulatory interactions are essential to construct the wild-type expression pattern of *ac* (19). The *ac* promoter fragment contains three sequences matching the HLH protein consensus binding site (E-boxes), which are capable of binding in vitro AS-C proteins as heterodimers with Da (2, 4, 28). This suggests that part of the *ac* transcriptional regulation may be mediated directly by the Ac and Sc proteins. However, the functional significance of these E-boxes has not been demonstrated. Moreover, the AS-C proneural proteins have been shown to bind in vitro to mammalian bHLH binding sites (4, 22, and our unpublished data), suggesting they have relatively lax binding requirements. Here, we have assessed the role of the *ac* E-boxes in *ac* regulation by examining, in transformant flies, the expression of *ac* promoter-*lacZ* constructs (*ac-lacZ*) lacking some or all of these sequences.

The *extramacrochaetae* (*emc*) gene is a negative regulator of the *ac* and *sc* functions. *emc* hypomorphic mutations promote the generation of extra SOs in an AS-C dosage-dependent way (3, 20). *emc* encodes a product that contains the HLH domain but lacks the adjacent basic region (9, 15). This has suggested that *emc* negatively regulates the *ac* and *sc* proneural functions by sequestering their products and preventing the efficient binding of Ac and Sc to their DNA target sequences located in genes implementing the neural developmental fate (9, 14, 15, 28). Moreover, as suggested on genetic grounds (10, 20), *emc* downregulates *sc* (and probably *ac*) transcription (6). This effect is easily explainable, at least in part, by *emc* interfering with *ac* and *sc* cross-stimulation and with *ac* self-activation (19). We have tested this proposal by analyzing, first, the effect in vivo of *emc* insufficiency on the expression of an *ac-lacZ* construct

* Corresponding author.
† Present address: Laboratorios Cusí, 08320 El Masnou, Spain.
‡ Present address: Ingelheim Diagnóstica y Tecnología, 08017 Barcelona, Spain.

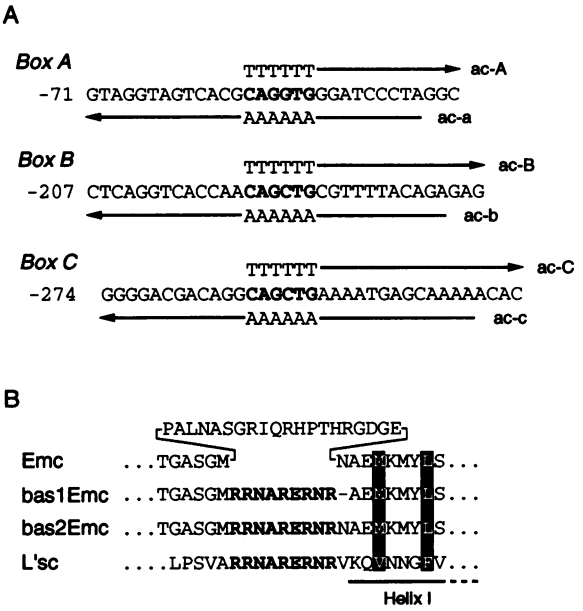

FIG. 1. Oligonucleotides used for site-directed mutagenesis of the *ac* E-boxes (A) and amino acid sequences of the chimeric bas1Emc and bas2Emc proteins in the regions that are different from those of the wild-type Emc protein (B). Sequences in panel A show the three *ac* E-boxes in boldface type and flanking nucleotides. Above and below each sequence, the oligonucleotides used for mutagenesis are indicated. Oligonucleotides starting by $(T)_6$ continued with sequences identical to the genomic sequence below up to the end of the arrow. Oligonucleotides containing the $(A)_6$ stretch had sequences identical to the complementary genomic sequence throughout the extent of the arrows. (B) The basic domain of the L'sc protein (in boldface) that was introduced into the Emc protein. Two different chimeric proteins, bas1Emc and bas2Emc, were prepared. They differed by the presence in bas2Emc of a single amino acid between the basic domain and the putative helix I. The start of this helix and some of its conserved hydrophobic residues (solid boxes) are indicated. Note that the introduction of the basic L'sc domain in Emc was accompanied by the deletion of the 21 or 20 amino acids flanking the amino side of the HLH domain. The decision to remove these amino acids was based on the observations that similar "domain swapping" strategies have worked for MyoD (8) and that fragments of other HLH proteins containing little more than the basic and HLH domains can still specifically bind to DNA (21). In addition, it was thought that two proline residues present in the deleted stretch might interfere with DNA binding (8).

in the presence and absence of Ac and Sc proteins, and second, the effect in vitro of wild-type and mutagenized Emc protein on the specific binding of proneural proteins to DNA.

MATERIALS AND METHODS

Transformation of *Drosophila* embryos. P-element-mediated transformation (ry^{506} stock) was performed as described previously (24) by using 0.8 to 1 mg of transformation plasmid per ml and 0.2 mg of the helper plasmid pUChsπΔ2-3 per ml.

Mutagenesis. Sequences in the *ac* promoter matching the consensus for HLH protein binding sites (E-boxes) are located at positions −58, −194, and −262 (Fig. 1A; start of transcription is taken as +1) (29). They have been named T5E1, T5E2, and T5E3, in order of increasing distance from the transcription start point (28). In this work, for the sake of simplicity, we shall call them A, B, and C, respectively. The core sequence of the E-boxes (CANNTG [2]) was replaced by the sequence $(T:A)_6$ by using the splicing by overlapping extension method (17). A pHSS7 plasmid containing a 0.8-kb fragment that includes the *ac* promoter and transcription initiation site and extends from an *Hae*II site to an *Eco*RI site (47 and 941 nucleotides upstream from the start of the *ac* coding sequence, respectively [29]) was used as template for the first polymerase chain reaction.

The following oligonucleotides were used in the primary polymerase chain reactions (see also Fig. 1A) (primers labeled with lowercase letters were used to amplify the antisense strand): ac-A, TTTTTTGGATCCCTAGGC; ac-a, TAGGGATCCAAAAAACGTGACTACCTAC; ac-B, TTTTTTCGTTTTACAGAGAG; ac-b, TCTGTAAAACGAAAAAATTGGTGACCTGAG; ac-C, TTTTTTAAAATGAGCAAAAACAC; ac-c, TTTTGCTCATTTTAAAAAACCTGTCGTCCCC.

All of the outer primers contained a *Not*I restriction site. Our sequencing of the nonmutated *ac* 0.8-kb genomic fragment used to construct the *ac-lacZ* gene (which is a subclone of the 2.2-kb *Eco*RI fragment originally used to sequence the *ac* gene [29]) showed that A-627 of the published *ac* sequence is, in our subclone, a G. This change has been introduced in primers ac-C and ac-c (underlined nucleotides). A mutated fragment without box A was used as template for removing box B. The resulting product was used as template for substituting box C. Each DNA fragment was subcloned in a pBS plasmid, modified as described previously (18). The resulting construct was digested with *Not*I, and the insert was subcloned in the *Drosophila* transformation plasmid pLac20 (26). Plasmids and the corresponding transformed lines were named with the letter(s) of the deleted E-box(es).

Emc^1 mutant protein was prepared by mutagenizing in vitro the wild-type *emc* cDNA 33 (15) with an Amersham mutagenesis kit according to the manufacturer's instructions. The oligo-AGGACCTCGATCCGTTCAT was used to introduce the mutation. bas1Emc was obtained from cDNA 33 by the splicing by overlapping extension method with the following mutation-containing oligonucleotides: *c*, GAGAACGAAATCGTGCCGAGATGAAG; *d*, GGCACGATTTCGTTCTCTAGCATTGCGCCTCATTCCGGAGGC. As external oligonucleotides, *a*, CCGCTCGCAGCAAAAGATATA (containing an *Sfu*I site), and *b*, CGTTGAAGAGATGCGCCTGGG (with a *Bgl*II site), were used. The amplified and mutagenized fragment was simultaneously digested with *Sfu*I and *Bgl*II and ligated to the *emc* cDNA 33-containing plasmid previously digested with the same enzymes. bas2Emc was also obtained by the splicing by overlapping extension method with the bas1Emc plasmid as the template. The oligonucleotides *e*, GAGAACGAAATCGTAATGCCGAGATGAAG, and *f*, ATTACGATTTCGTTCTC, were used. As external oligonucleotides, *a* and *b* were used. The mutagenized fragment was cloned as in the bas1Emc mutagenesis. All mutagenized sequences were confirmed by DNA sequencing of appropriate clones.

Analysis of *lacZ* expression. Imaginal discs were dissected in phosphate-buffered saline (PBS) and fixed in 0.5% glutaraldehyde for 2 min in ice. After two 5-min washes, discs were stained for β-galactosidase activity by incubation at 37°C in 0.2% X-Gal (5-bromo-4-chloro-3-indolyl-β-D-galactopyranoside)–0.5 mM $K_4[Fe(CN)_6]$–0.5 mM $K_3[Fe(CN)_6]$–0.3% Triton X-100 in PBS. Discs were dehydrated and mounted in Canada balsam.

In vitro transcription and translation. *l'sc* mRNA was

obtained from a pBluescript KS(+) derivative containing an *l'sc* cDNA. After linearization with *Xba*I, RNA was synthesized with T3 RNA polymerase under the usual conditions (25). *sc* mRNA was synthesized with T3 RNA polymerase after *Xba*I digestion of a pBluescript KS(+) derivative containing *sc* genomic DNA (29). E47S mRNA was transcribed with T3 RNA polymerase from a previously linearized (*Eco*RI) plasmid from a cDNA-containing plasmid (22) kindly provided by Harald Vaessin. *emc* mRNA was transcribed with T7 RNA polymerase from the *emc* cDNA 33-containing plasmid (15). Emc[1], bas1Emc, and bas2Emc mRNAs were obtained with the adequate RNA polymerase from the respective plasmids as described above. Each transcription reaction contained approximately 100 U of RNA polymerase. After phenolization, the RNA was precipitated with ethanol, redissolved in 50 μl of DEPC (pyrocarbonic acid diethyl ester)-treated water, and stored at −70°C. An aliquot was analyzed with an agarose gel to assess yield and size.

Approximately one-tenth of the RNA obtained in each transcription reaction was translated with rabbit reticulocyte lysates treated with micrococcal nuclease (Promega) in a 50-μl volume reaction according to the manufacturer's directions. Reaction yields were estimated in parallel translations carried out in the presence of [^{35}S]methionine (>1,000 Ci/mmol; Amersham) and were analyzed with sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE) gels (25).

DNA binding assay. Oligonucleotides containing E-boxes were 5'-labeled with T4 polynucleotide kinase and [γ-^{32}P] ATP. Oligonucleotides were denatured and annealed with a 10-fold excess of complementary unlabeled oligonucleotide. About 3 to 10 μl of reaction lysates containing, except where indicated, similar amounts of synthesized proteins was mixed and adjusted to a final 15-μl volume with mock lysate. These mixes were incubated at 37°C for 20 min. Fifteen microliters of a DNA binding cocktail was then added to yield the following final concentrations: 20 mM HEPES (*N*-2-hydroxyethylpiperazine-*N'*-2-ethanesulfonic acid [pH 7.6]), 50 mM KCl, 1 mM dithiothreitol, 1 mM EDTA, 5% glycerol, 1 μg of poly(dI-dC), and about 0.1 ng of labeled, double-stranded probe. This mixture was incubated for 15 min at room temperature and was loaded on a nondenaturing PAGE gel as described previously (1). Gels were dried and autoradiographed.

RESULTS

***ac* regulation requires E-boxes present in its promoter.** A 0.8-kb *ac* promoter fragment mediates both *ac* self-activation and activation by *sc* (19). Within 262 nucleotides from the origin of transcription, this fragment contains three E-boxes that bind Ac and Sc proteins in vitro and may mediate these regulatory interactions (4, 28). Accordingly, we have examined the functional significance of these E-boxes by obtaining transformed flies carrying variants of the 0.8-kb *ac-lacZ* construct in which either one (A), two (AB), or three (ABC) of the E-boxes had been replaced by $(T:A)_6$ sequences. *lacZ* expression in wing imaginal discs was examined. As previously shown (19), the unmodified *ac-lacZ* construct was expressed in discrete groups of cells (Fig. 2A) that very closely match the distribution of proneural clusters of *ac* gene expression (6, 27). Elimination of the A box, the one closest to the origin of transcription, almost did not modify the pattern of *lacZ* expression in proneural clusters (compare Fig. 2A and B). However, it greatly intensified expression in a large region covering the border between presumptive notum and wing (Fig. 2B, wn) and, to a lesser extent, in two bands parallel and posterior to the presumptive third vein (Fig. 2B, pc; the position of this vein is defined by the strong proneural cluster that gives rise to its sensilla campaniformia; Fig. 2A, v3). Expression in these regions and in the anterior-lateral region of the presumptive prescutum (an) also occurred in discs carrying constructs lacking two or three of the *ac* E-boxes (Fig. 2C and D). Most importantly, removal of these additional E-boxes reduced *lacZ* expression in proneural clusters, and these became ill-defined. These effects were always most noticeable in lines lacking the three E-boxes (Fig. 2D). It is thus concluded that proper expression of the *ac* gene requires at least the B and C boxes. This, together with the ability of the Ac and Sc proteins to bind in vitro as heterodimers with Da to these E-boxes (4, 28, and our unpublished observations), strongly supports direct stimulation of the *ac* promoter by the Ac and Sc proteins.

It is of interest that Ac and Sc proteins accumulate to large extents in SO precursor cells (6, 27). This appears to be accomplished by a specific mechanism that activates *ac* and *sc* in these cells (19). In lines carrying the triple mutant promoter, we have not detected *lacZ* expression in SO precursors, suggesting that the *ac* E-boxes are also involved in this activation.

***emc* downregulates the *ac* promoter.** The Emc protein, an antagonist of proneural function, interacts with the proneural proteins and prevents their binding in vitro to oligonucleotide probes containing the *ac* E-boxes plus short flanking sequences (4, 28 [also see Fig. 4]). If in vivo Ac and/or Sc bind to and activate the *ac* promoter, it should be expected that Emc inhibited such activation and that this effect depended on the presence of Ac and/or Sc. To test these predictions, we compared the spatial pattern of *lacZ* transcription in wing imaginal discs from *ac-lacZ* larvae carrying either wild-type *emc* alleles or one of the phenotypically strongest, but still viable, hypomorphic combinations of *emc* mutations [$Df(3)emc^{E12}/emc^{Pel}$] (10). Figure 3A and C show that the *emc* insufficiency caused a general *lacZ* overexpression, which was strongest in the center (DC and PSA areas) and anterior parts of the presumptive notum (PS), thus fulfilling the first expectation. When attempting to test the second one, it was found that larvae containing these *emc* mutations, together with the deficiency of *ac* and *sc* and the chromosome carrying the *ac-lacZ* insertion, were not viable. However, larvae of a similar genotype but deficient for only *ac* were viable. Because most *sc* transcription in the central part of the presumptive notum (mainly in the DC and PSA proneural clusters) depends on the presence of Ac protein (19, 27), these areas of the *ac* mutant discs are in fact devoid of Ac and most of the Sc protein. Figure 3C and D show that the depletion of these proteins largely reversed the *lacZ* overexpression caused by the *emc* insufficiency in the DC and PSA areas, while it did not appreciably modify the *lacZ* overexpression in other sites, like the tegula (TG), anterior notopleural (ANP), anterior postalar (APA), and PS areas where *sc* is still strongly transcribed (19). This indicates that overactivation of the *ac* promoter by the *emc* insufficiency requires the presence of at least the Sc protein. Note, however, that the level of *lacZ* expression in the DC and PSA areas of the *ac emc* mutant disc is slightly higher than in the *ac* mutant emc^+ disc (compare Fig. 3D and B). This is consistent with residual HLH activators, like the Sc protein, being present in these areas (see Discussion).

Modifications of the Emc protein and its interference with

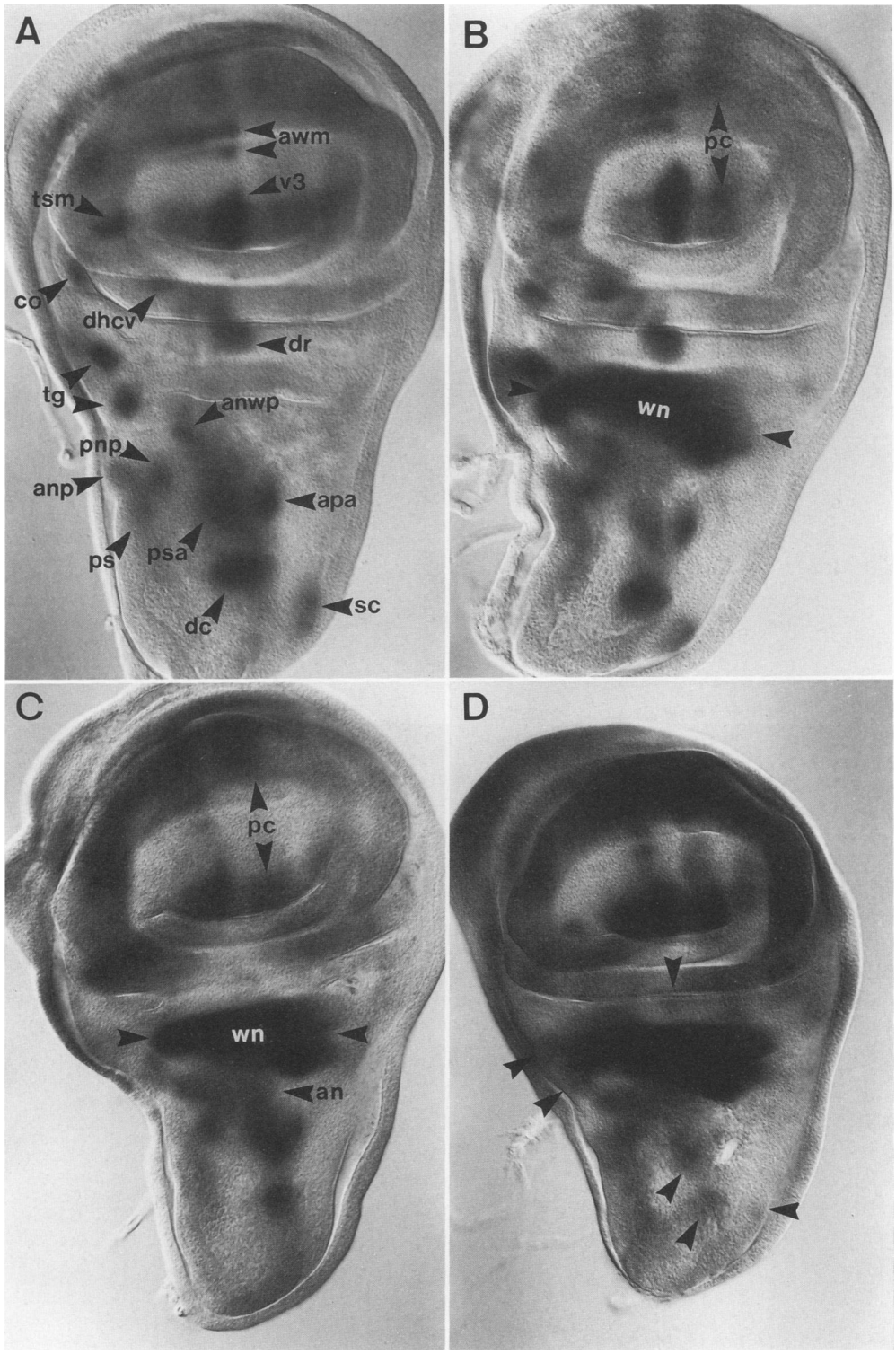

FIG. 2. Elimination of E-boxes in the *ac* promoter decreases *ac-lacZ* transcription in proneural clusters. *lacZ* expression was driven by a wild-type *ac* promoter (A) or an *ac* promoter lacking one (B), two (C), or three (D) E-boxes. The pattern of β-galactosidase accumulation in different regions of the disc in panel A faithfully reproduces the wild-type pattern of proneural clusters of *ac* and *sc* expression (6, 19, 27). Many proneural clusters are shown by arrowheads and are labeled according to the corresponding region of the adult fly: awm, anterior wing margin; v3, vein 3; tsm, twin sensilla of the wing margin; co, costa; dhcv, dorsal humeral crossvein; dr, dorsal radius; tg, tegula; anwp, anterior notal wing process; anp and pnp, anterior and posterior notopleural, respectively; ps, presutural; psa, posterior supraalar; apa, anterior postalar; dc, dorsocentral; sc, scutelar. Note that the stepwise elimination of the E-boxes (B to D) results in a progressive decrease of β-galactosidase accumulation in the sites that correspond to the proneural clusters. In panel D, the residual β-galactosidase accumulation in the dr, tg, psa, dc, and sc sites is shown by arrowheads. Elimination of the first E-box strongly increases β-galactosidase accumulation in some areas that do not correspond to proneural clusters, like the pc, wn, and an regions in panels B and C (see text for a description of these areas), and are only slightly stained with the unmodified *ac* promoter (A). The panels shown represent typical examples of patterns observed on discs of 8 (A), 15 (B), 9 (C), and 14 (D) independent transformant lines.

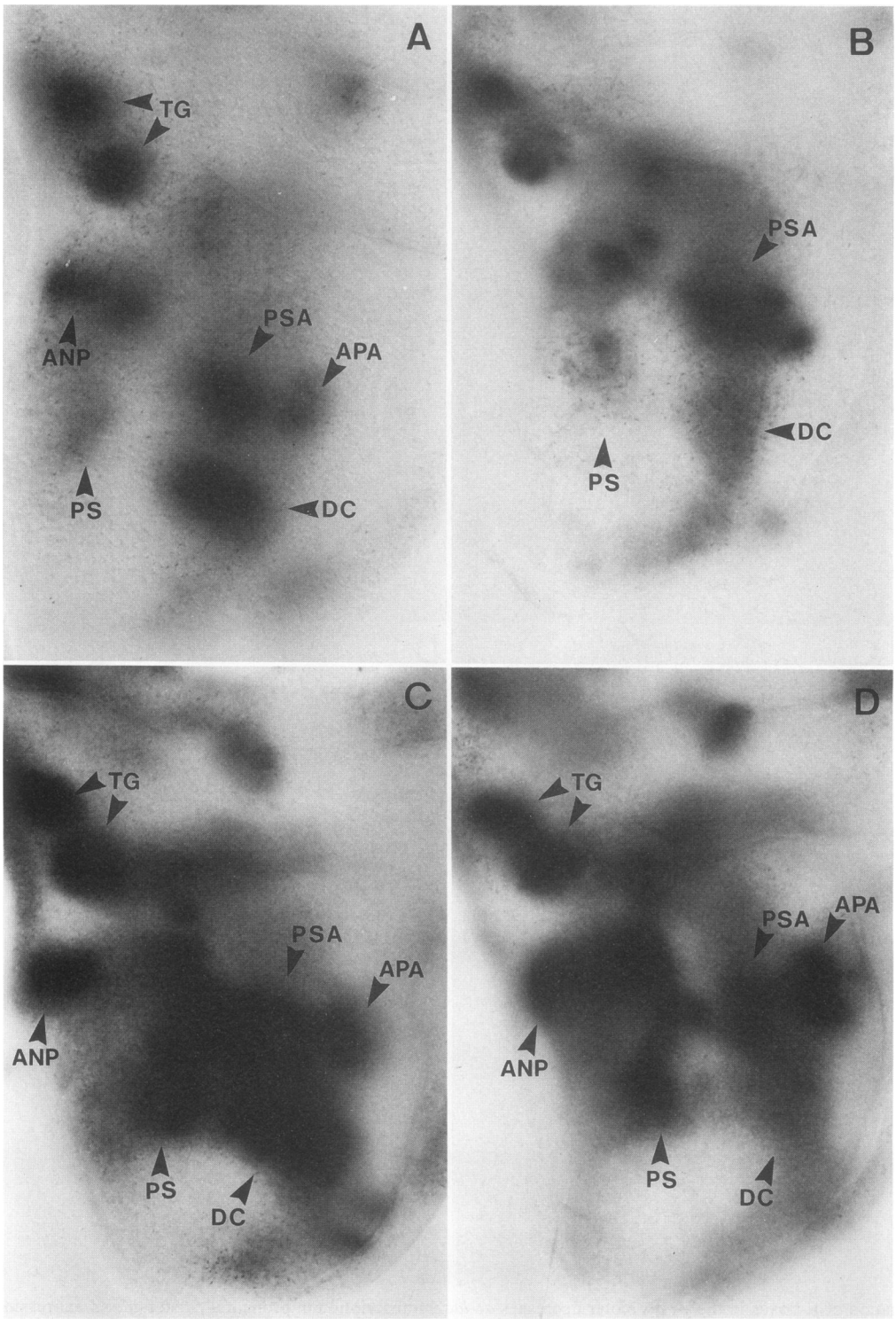

FIG. 3. *emc* negatively regulates the *ac* promoter by interfering with *ac* self-activation and stimulation by *sc*. *lacZ* expression was driven by a 3.8-kb *ac* promoter fragment in the following genetic backgrounds: A, wild type; B, *ac* mutant $Df(1)y^{3PL}sc^{8R}$; C, *emc* mutant $Df(3)emc^{E12}/emc^{pel}$; and D, *ac emc* mutant $Df(1)y^{3PL}sc^{8R}$; $Df(3)emc^{E12}/emc^{pel}$ imaginal discs. To obtain the discs shown in panel C, $Df(3)emc^{E12}$/TM6B females were mated with $Df(1)y^{3PL}sc^{8R}$; emc^{pel} males carrying the *ac-lacZ* insertion. In the case of the disc shown in panel D, $Df(1)y^{3PL}sc^{8R}$; emc^{pel} females carrying the *ac-lacZ* insertion were mated with $Df(3)emc^{E12}$/TM6B males. In both cases, larvae not TM6B were selected for the *ac* mutant genotype according to the phenotype of the mouth hooks. In the first case, larvae of the adequate genotype were extremely unusual. Pictures are centered on the presumptive notum region of imaginal wing discs from third instar larvae. The following presumptive regions are named: TG, tegula; ANP, anterior notopleural; PS, prescutal; DC, dorsocentral; PSA, posterior supraalar; APA, anterior postalar. To best show differences in staining levels, X-Gal staining of all discs shown was carried out simultaneously and under the same conditions.

the binding of proneural proteins to an *ac* E-box. The current model for the interference of Emc with proneural protein function proposes that Emc, by means of its HLH motif, dimerizes with these proteins and that the resulting complexes, with only one basic region, cannot bind to DNA. This model is largely based on biochemical evidence obtained with a system of similar proteins involved in controlling mammalian myogenesis (1). No experimental evidence is yet available on whether Emc interacts with the proneural proteins by means of its HLH domain or whether the absence of the basic region is essential for the inability of Emc-containing complexes to bind to DNA. To examine the first of these points, we prepared in a cell-free system a modified Emc protein that contained the single amino acid substitution (aspartic acid for valine) detected in the coding sequence of the *emc*^{1} loss-of-function mutant (15). This modification introduces a hydrophilic amino acid at a location, the carboxyl end of helix 1, where most HLH proteins have a hydrophobic residue. Although not proven, this substitution, which may disrupt the putative hydrophobic cavity created by the pair of amphipathic helices (21), has been proposed to interfere with the ability of the protein to form dimers (15). We compared, by using an electrophoretic mobility shift assay, the abilities of wild-type and Emc^1 protein to interfere with the binding of the proneural products to DNA. We used L'sc/E47S heterodimers (E47S is the murine homolog of the *Drosophila* Da protein [21]) because, in our conditions, these proteins yielded the most clear and well-separated retarded bands that did not overlap with background bands. In contrast to wild-type Emc, Emc^1 protein, even in a threefold molar excess, does not affect the binding by L'sc/E47S of a probe containing the ac-A E-box (Fig. 4A, compare lanes 5 and 6 with 7 and 8). This strongly suggests that complexes of Emc with proneural proteins are mediated by the Emc HLH domain and that the above modification is indeed responsible for the loss-of-function phenotype of the *emc*^{1} mutant.

To test whether the absence of the basic domain in Emc is essential for its function, we introduced the basic domain of L'sc into this protein (Fig. 1B). We checked whether this modification altered the properties of Emc—namely, whether it conferred to it the ability to bind to DNA or at least reduced its antagonistic behavior. We prepared two Emc variants, bas1Emc and bas2Emc (Fig. 1B), that differed in the spacing between the basic domain and the presumptive beginning of helix 1. bas2Emc has the same spacing as the L'sc protein, and bas1Emc has one residue less. Figure 4B shows that bas1Emc did not bind to DNA, either alone or in combination with E47S or L'sc (lanes 3, 5, and 7, respectively). Moreover, wild-type Emc (lane 9) and bas1Emc (lane 10) inhibited L'sc/E47S binding to the ac-A probe to a similar extent. The same results were obtained with bas2Emc (not shown). Thus, the mere absence of a basic domain in Emc does not seem sufficient to explain its antagonistic behavior.

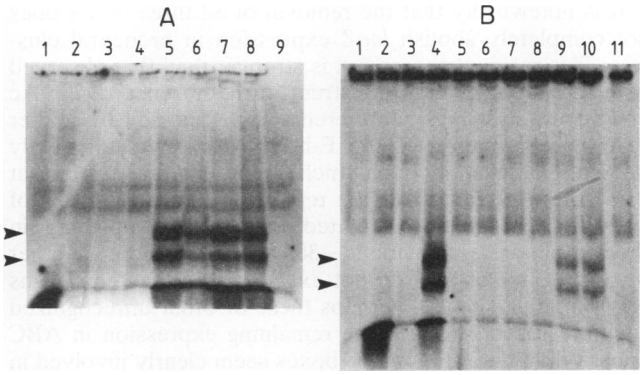

FIG. 4. Effect of different mutations in the Emc protein on its inhibitory effect on L'sc/E47S binding to DNA. (A) An Emc protein carrying a mutation in the HLH domain (Emc^1) does not inhibit L'sc/E47S binding to T5E1 oligonucleotide (28). The following proteins were synthesized in the reticulocyte lysates (by lane): 1, E47S; 2, L'sc; 3, Emc; 4, Emc^1; 5, L'sc and E47S; 6, L'sc, E47S, and Emc; 7, L'sc, E47S, and a threefold excess of Emc^1; 8, L'sc, E47S, and an approximately equimolar amount of Emc^1. Lane 9 shows results with no mRNA added. Retarded bands corresponding to L'sc/E47S heterodimers are shown by arrowheads. Two bands are present, due, most likely, to the translation in vitro of *l'sc* RNA, which yielded two proteins of different molecular weights (probably caused by different translation initiator sites). Other retarded bands appear in the nonprogrammed reticulocyte lysate (lane 9) and are probably caused by HLH proteins endogenous in the lysate. The bottom one has mobility equal to that obtained with reticulocyte lysate programmed with E47S mRNA (compare lanes 1 and 9). The apparent mismatch between the loading wells and the numbers on the lanes is due to breakage of the gel during drying, which displaced the upper left corner and the wells towards the left. The numbers match the middle and bottom portion of the gel, where the retarded probe is found. (B) bas1Emc protein does not bind to T5A oligonucleotide, and it inhibits L'sc/E47S binding to T5A. The following proteins were synthesized in the reticulocyte lysates (by lane): 1, L'sc; 2, E47S; 3, bas1Emc; 4, L'sc and E47S; 5, L'sc and bas1Emc; 6, E47S and Emc; 7, E47S and bas1Emc; 8, Emc and bas1Emc; 9, L'sc, E47S, and Emc; 10, L'sc, E47S, and bas1Emc; 11, E47S, Emc and bas1Emc. Arrowheads point to the same bands as in panel A. No additional bands appeared in the gels between the bottom band and the unretarded oligonucleotide in panel A or B.

DISCUSSION

E-boxes mediate control of *ac* expression. We have previously shown that, in the wing imaginal disc, most of the sequences necessary for the correct expression of the *ac* gene are contained within 0.8 kb upstream from the *ac* transcription start site (19). We now demonstrate that three E-boxes present in this stretch of DNA and relatively close to the transcriptional start (less than 262 bp) are important for the proper activation of the *ac* promoter. In transformed flies carrying *ac-lacZ* constructs in which these E-boxes have been eliminated, *lacZ* expression in proneural clusters is strongly reduced and the clusters become ill-defined. Ac, Sc, and L'sc have been shown to bind in vitro as heterodimers with Da to oligonucleotides containing these E-boxes (4, 28, and our unpublished observations). Thus, it is most probable that these E-boxes mediate two important regulatory actions on the *ac* gene: the *ac* self-activation and the activation by *sc* (19). These actions are most likely accomplished by direct binding with dimers containing Ac and/or Sc with Da and are most likely responsible for maintaining a high level of expression of both *ac* and *sc* genes in the proneural groups of cells. Given the relatively low specificity demonstrated in vitro by proneural and other bHLH proteins for the different *ac* and other E-box-containing sites (4, 22, 28), it is possible that other bHLH proteins may also recognize the *ac* sites and participate in the regulation of this gene. For example, Asense, another proneural protein encoded in the AS-C, activates *ac-lacZ* when it accumulates in the imaginal discs after a heat shock in transformant larvae harboring an *hsp70-asense* chimeric gene (8a).

It is noteworthy that the removal of all three boxes does not completely abolish *lacZ* expression in proneural clusters. This residual expression is stronger than that observed with the intact *ac-lacZ* construct in the absence of the Ac and Sc proteins (Fig. 2 in reference 19). This indicates either that the elimination of the E-boxes does not completely prevent Ac-Sc interaction (which seems unlikely [28]) or that these proteins can affect the regulation of *ac* by means of sites different from the mutated ones. Further upstream in the *ac* promoter, at positions −382 and −816, there are other putative E-boxes that do not exactly match the consensus HLH binding site (2). Perhaps these or other unrecognized sites are responsible for the remaining expression in ABC lines. While the A, B, and C boxes seem clearly involved in the *ac* self-stimulation and the activation by *sc*, the residual expression in clusters further suggests that these HLH binding sites probably do not participate in the initial activation of *ac* that occurs when *ac*-dependent proneural clusters (dorsocentral, postsupraalar) are founded (most other clusters are initiated by the expression of *sc*) (19). At present, the factors and factor-binding sites responsible for the initial activation of *ac* or *sc* in clusters of cells are unknown.

Several regions of the wing disc that hardly express *lacZ* in *ac-lacZ* larvae do so in larvae that carry *ac-lacZ* constructs lacking one, two, or three E-boxes (Fig. 2). The expression is strongest in the triple mutant lines. Overstained discs carrying the unmodified *ac-lacZ* construct also show expression in these areas, albeit at very reduced levels (not shown). Thus, it seems unlikely that replacement of the E-boxes by $(T:A)_6$ sequences creates new sites for binding of activators. A more likely explanation is that elimination of the E-boxes prevents the binding of unknown factors, probably HLH proteins, that under normal conditions bind to the E-sites and inhibit *ac* transcription promoted by activator(s) localized to these areas of the disc. Whatever the mechanism, the opposite effect of E-box removal in proneural clusters (decreased transcription) and other disc areas (transcriptional activation) reflects the complex regulation of *ac* and the heterogenous distribution of the factors involved. The gradually weaker or stronger *lacZ* expression (depending on which regions of the disc are considered) with increasing number of removed E-boxes suggests that these act cooperatively in both effects.

***emc* negatively regulates *ac* transcription.** We and another group have previously proposed that *emc* negatively regulates *ac* and *sc* proneural functions by sequestering their products (9, 15, 28). The titration by Emc of the proneural products, together with the cross-regulation between *ac* and *sc* and the *ac* self-stimulation, predicts an effect of *emc*, mediated by the Ac and Sc proteins, on *ac* transcription (15, 19). This indeed seems to be the case: the strong hypomorphic *emc* combination $Df(3)emc^{E12}/emc^{pel}$ promotes overexpression of *lacZ* in the imaginal discs of *ac-lacZ* transformed flies. (The same genetic combination also promotes *sc* overexpression [6].) Within the wing disc, this *lacZ* overexpression is largely reversed in the central region of the presumptive notum (Fig. 3D), when $Df(3)emc^{E12}/emc^{pel}$ larvae are also deficient for *ac*. The effect is localized to this region of the disc because the expression of *sc* in this area depends mostly on *ac* and, consequently, the *ac* deficiency causes the absence of Ac and most of the Sc protein (19, 27). The *lacZ* overexpression is maintained in other areas of the disc. Thus, as previously shown in an emc^+ background (19), Sc alone can maintain a high level of activity of the *ac* promoter. Moreover, the Sc protein activates the *ac* promoter ubiquitously when it is homogeneously distributed in *ac sc* mutant; *ac-lacZ*; *hsp70-scute* larvae (22a). Taken together, these results indicate that, under wild-type conditions, *emc* partially titrates Ac, Sc, and other HLH products needed for *ac* (and *sc*) activation, maintaining the activity of these genes at levels compatible with the development of the wild-type complement of SOs. The release of the negative effect of Emc enhances these stimulations and leads to an increased level of *ac* and *sc* transcription and to an increased number of cells committed into sensory mother cells (7). This condition may actually prevail in some primitive drosophilids, which display patterns of SOs similar to those of *emc* mutants in *D. melanogaster* (12).

In the central region of the presumptive notum, *ac emc* mutant discs show a level of *lacZ* expression slightly higher than *ac* mutant discs (compare Fig. 3C and D). This may be due to factors that activate the *ac* promoter in that region of the disc (19) and are not completely titrated by the residual amount of Emc. The fact that some bristles develop infrequently in the DC-PSA region of *emc ac* mutant flies (10, and our unpublished results) suggests that one of these factors is Sc. As previously indicated, in the absence of *ac*, *sc* is not expressed in this region at detectable levels, but it is possible that under conditions of Emc insufficiency, a basal, low level of *sc* expression may be sufficient to promote the occasional development of SOs and to activate weakly the *ac* promoter of our constructs.

emc^1 is a homozygous-lethal mutation that, in heterozygosis with weak *emc* alleles, causes a strong phenotype of ectopic SOs, which is only slightly weaker than that of the combination of $Df(3)emc^{E12}$ with these weak alleles (10). Thus, emc^1 is a very strong hypomorphic allele. Because, at the molecular level, the only modification detected was a change of a valine to an aspartic acid at the carboxyl end of helix 1 of the Emc protein, it was proposed that this replacement would interfere with the ability of the mutant protein to form heterodimers (15). Indeed, we have now shown that, in contrast to wild-type Emc protein, Emc^1 protein does not abolish L'sc/E47S binding to DNA (Fig. 4A). This result supports the proposal that the interaction between proneural products and Emc would be achieved by heterodimerization through the HLH domain (21).

The introduction of a basic domain in the Emc protein, that of the L'sc protein, does not seem sufficient to confer to it the ability to bind DNA as a partner with proneural proteins. Moreover, we find no alteration in its ability to interfere with the binding of bHLH heterodimers to DNA. Although fragments of mammalian bHLH proteins conserving only the basic region and the HLH domain have been shown to be able to heterodimerize and specifically bind to DNA (8), Emc may have some other structural constraints that prevent, even in the presence of a basic domain, its interaction with DNA. A relatively trivial explanation might be that the conformation of the introduced basic domain in a molecule that has evolved as an antagonist of DNA interaction may not be appropriate for such an interaction. It is worth pointing out the contrast of our results with those showing that the removal of just the basic residues in MyoD eliminates its DNA binding ability and converts it into an Emc-like antagonist (8). Together, these data suggest that the basic domain is necessary, but not sufficient, to give the ability to bind to DNA. It remains to be tested whether this is a particular property of our system or is a general characteristic of HLH proteins.

ACKNOWLEDGMENTS

We thank our laboratory colleagues for constructive criticism of the manuscript and H. Vaessin for an E47S-containing plasmid. C.M. was a Rich Foundation fellow, and J.G. was a postdoctoral fellow of the Consejo Superior de Investigaciones Científicas.

This work was supported by grants from the Dirección General de Investigación Científica y Técnica (PB90-0085) and the Comunidad Autónoma of Madrid and an institutional grant from the Fundación Ramón Areces to the Centro de Biología Molecular.

ADDENDUM

Recently, Van Doren et al. (27a) reported results consistent with those described here concerning the mutagenesis of the *ac* promoter E-boxes and the overexpression of the *ac-lacZ* transgene in *emc* hypomorphic mutants.

REFERENCES

1. **Benezra, R., R. L. Davis, D. Lockshon, D. L. Turner, and H. Weintraub.** 1990. The protein Id: a negative regulator of helix-loop-helix DNA binding proteins. Cell **61:**49–59.
2. **Blackwell, K., and H. Weintraub.** 1990. Differences and similarities in DNA-binding preferences of MyoD and E2A protein complexes revealed by binding site selection. Science **250:**1104–1110.
3. **Botas, J., J. Moscoso del Prado, and A. García-Bellido.** 1982. Gene-dose titration analysis in the search of trans-regulatory genes in *Drosophila*. EMBO J. **1:**307–310.
4. **Cabrera, C. V., and M. C. Alonso.** 1991. Transcriptional activation by heterodimers of the *achaete-scute* and *daughterless* gene products of *Drosophila*. EMBO J. **10:**2965–2973.
5. **Campuzano, S., and J. Modolell.** 1992. Patterning of the *Drosophila* nervous system: the *achaete-scute* gene complex. Trends Genet. **8:**202–207.
6. **Cubas, P., J.-F. de Celis, S. Campuzano, and J. Modolell.** 1991. Proneural clusters of *achaete-scute* expression and the generation of sensory organs in the *Drosophila* imaginal wing disc. Genes Dev. **5:**996–1008.
7. **Cubas, P., and J. Modolell.** 1992. The *extramacrochaetae* gene provides information for sensory organ patterning. EMBO J. **11:**3385–3393.
8. **Davis, R. L., P. F. Cheng, A. B. Lassar, and H. Weintraub.** 1990. The MyoD DNA binding domain contains a recognition code for muscle-specific gene activation. Cell **60:**733–746.

8a. **Domínguez, M., and C. Martínez.** Unpublished observations.

9. **Ellis, H. M., D. R. Spann, and J. W. Posakony.** 1990. *extramacrochaetae*, a negative regulator of sensory organ development in *Drosophila*, defines a new class of helix-loop-helix proteins. Cell **61:**27–38.
10. **García-Alonso, L., and A. García-Bellido.** 1988. *Extramacrochaetae*, a *trans*-acting gene of the *achaete-scute* complex of *Drosophila* involved in cell comunication. Roux's Arch. Dev. Biol. **197:**328–338.
11. **García-Bellido, A.** 1979. Genetic analysis of the *achaete-scute* system of *Drosophila melanogaster*. Genetics **91:**491–520.
12. **García-Bellido, A.** 1981. From the gene to the pattern: chaetae differentiation, p. 281–304. *In* C. W. Lloyd and D. A. Rees (ed.), Cellular controls in differentiation. Academic Press, London.
13. **García-Bellido, A., and P. Santamaría.** 1978. Developmental analysis of the *achaete-scute* system of *Drosophila melanogaster*. Genetics **88:**469–486.
14. **Garrell, J., and S. Campuzano.** 1991. The helix-loop-helix domain: a common motif for bristles, muscles and sex. Bioessays **13:**493–498.
15. **Garrell, J., and J. Modolell.** 1990. The Drosophila *extramacrochaetae* locus, an antagonist of proneural genes that, like these genes, encodes a helix-loop-helix protein. Cell **61:**39–48.
16. **Ghysen, A., and C. Dambly-Chaudière.** 1989. Genesis of the *Drosophila* peripheral nervous system. Trends Genet. **5:**251–255.
17. **Ho, S., H. Hunt, R. Horton, J. Pullen, and L. Pease.** 1989. Site-directed mutagenesis by overlap extension using the polymerase chain reaction. Gene **77:**51–59.
18. **Marchuk, D., M. Drumm, A. Saulino, and F. Collins.** 1990. Construction of T-vectors, a rapid and general system for direct cloning of unmodified PCR products. Nucleic Acids Res. **19:**1154.
19. **Martínez, C., and J. Modolell.** 1991. Cross-regulatory interactions between the proneural *achaete* and *scute* genes of *Drosophila*. Science **251:**1485–1487.
20. **Moscoso del Prado, J., and A. García-Bellido.** 1984. Genetic regulation of the *achaete-scute* complex of *Drosophila melanogaster*. Roux's Arch. Dev. Biol. **193:**242–245.
21. **Murre, C., P. S. McCaw, and D. Baltimore.** 1989. A new DNA binding and dimerization motif in immunoglobin enhancer binding, daughterless, MyoD and myc proteins. Cell **56:**777–783.
22. **Murre, C., P. S. McCaw, H. Vaessin, M. Caudy, L. Y. Jan, Y. N. Jan, C. V. Cabrera, J. N. Buskin, S. D. Hauschka, A. B. Lassar, H. Weintraub, and D. Baltimore.** 1989. Interactions between heterologous helix-loop-helix proteins generate complexes that bind specifically to a common DNA sequence. Cell **58:**537–544.

22a. **Rodríguez, I., and C. Martínez.** Unpublished observations.

23. **Romani, S., S. Campuzano, E. Macagno, and J. Modolell.** 1989. Expression of *achaete* and *scute* genes in *Drosophila* imaginal discs and their function in sensory organ development. Genes Dev. **3:**997–1007.
24. **Rubin, G. M., and A. C. Spradling.** 1982. Genetic transformation of *Drosophila* with transposable element vectors. Science **218:**348–353.
25. **Sambrook, J., E. F. Fritsch, and T. Maniatis.** 1989. Molecular cloning: a laboratory manual. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y.
26. **Schröder, C., D. Tautz, E. Seifert, and H. Jäckle.** 1988. Differential regulation of the two transcripts from the *Drosophila* gap segmentation gene *hunchback*. EMBO J. **7:**2881–2887.
27. **Skeath, J. B., and S. B. Carroll.** 1991. Regulation of *achaete-scute* gene expression and sensory organ pattern formation in the *Drosophila* wing. Genes Dev. **5:**984–995.

27a. **Van Doren, M., P. A. Powell, D. Pasternak, A. Singson, and J. W. Posakony.** 1992. Spatial regulation of proneural gene activity: auto- and cross-activation of *achaete* is antagonized by *extramacrochaetae*. Genes Dev. **6:**2592–2605.

28. **Van Doren, M., H. M. Ellis, and J. W. Posakony.** 1991. The *Drosophila extramacrochaetae* protein antagonizes sequence-specific DNA binding by *daughterless/achaete-scute* protein complexes. Development **113:**245–255.
29. **Villares, R., and C. V. Cabrera.** 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.