hairy function as a DNA-binding helix-loop-helix repressor of *Drosophila* sensory organ formation

Shunji Ohsako,¹ Jeanette Hyer,¹ Grace Panganiban,² Ian Oliver,¹ and Michael Caudy^{1,3}

¹Department of Cell Biology and Anatomy, Cornell University Medical College, New York, New York 10021 USA; ²Howard Hughes Medical Institute and Laboratory of Molecular Biology, University of Wisconsin, Madison, Wisconsin 53706 USA

Sensory organ formation in *Drosophila* is activated by proneural genes that encode basic-helix-loop-helix (bHLH) transcription factors. These genes are antagonized by hairy and other proline-bHLH proteins. hairy has not been shown to bind to DNA and has been proposed to form inactive heterodimers with proneural activator proteins. Here, we show that hairy does bind to DNA and has novel DNA-binding activity: hairy prefers a noncanonical site, CACGCG, although it also binds to related sites. Mutation of a single CACGCG site in the *achaete* (*ac*) proneural gene blocks hairy-mediated repression of *ac* transcription in cultured *Drosophila* cells. Moreover, the same CACGCG mutation in an *ac* minigene transformed into *Drosophila* creates ectopic sensory hair organs like those seen in *hairy* mutants. Together these results indicate that hairy represses sensory organ formation by directly repressing transcription of the *ac* proneural gene.

[**Key Words:** Helix-loop-helix; DNA binding; transcription; repression; neurogenesis; *Drosophila*]

Received August 15, 1994; revised version accepted September 21, 1994.

Sensory organs in insects derive from sensory organ precursors (SOPs) that divide to produce both the sensory organ, such as a sensory hair, and the sensory neuron that innervates that organ (for review, see Chyzen and Dambly-Chaudiere 1989; Jan and Jan 1993a). Genetic analysis in *Drosophila* has shown that SOPs are formed in a precise pattern that is controlled by both positive and negative regulatory genes. The positive regulators include *daughterless* (*da*), the proneural genes of the *achaete-scute* complex [*AS-C*: *achaete* (*ac*), *scute* (*sc*), *lethal of scute* (*l'sc*), *asense* (*ase*)] and the proneural gene *atonal* (for review, see Campuzano and Modolell 1992; Jan and Jan 1993b). Recessive loss-of-function mutations in *da* lead to the loss of all SOPs and the sense organs (SOs) to which they give rise, whereas loss-of-function mutations in *atonal* and *AS-C* genes result in the loss of specific subsets of SOPs and SOs. In contrast, dominant gain-of-function mutations in the *AS-C* genes result in the formation of ectopic SOPs and SOs. Together, these dominant and recessive phenotypes indicate that the proneural genes act as genetic "switches" for activating SOP formation during development (Campuzano and Modolell 1992; Jan and Jan 1993b).

Negative regulation of proneural gene expression, and therefore of SOP and SO formation, occurs at two distinct levels. *AS-C* genes are initially expressed in proneural clusters in which many or all cells are competent to become an SOP (Ghyzen and Dambly-Chaudiere

1989; Romani et al. 1989). Genes such as *hairy* [gene; (hairy) protein] and *extramacrochaetae* (*emc*) control the pattern in which proneural clusters form (Cubas et al. 1991; Skeath and Carroll 1991). That pattern is subsequently refined by the restriction of proneural gene expression to the single cell that will become the SOP. This restriction depends on the neurogenic genes such as *Notch*, *Delta*, and *Enhancer of split* [*E(spl)*], which mediate lateral inhibition of SOPs by cell-cell interactions (Artavanis-Tsakonas and Simpson 1991; Campos-Ortega 1993). In neurogenic mutants proneural clusters form at the normal positions but multiple SOPs and SOs are formed within each cluster, whereas in *hairy* and *emc* mutants proneural clusters form at ectopic positions but only one or a few SOPs and SOs are formed within each cluster.

Molecular analysis has shown that the majority of the activator and repressor genes that regulate SOP formation encode helix-loop-helix (HLH) transcription factors (Jan and Jan 1993b). The HLH domain is a dimerization domain (Murre et al. 1989a), and many HLH proteins contain an adjacent basic region that mediates DNA binding. Heterodimer formation between basic-HLH (bHLH) proteins is an important mechanism for regulating DNA-binding activity of bHLH proteins (for review, see Jones 1990). For example, the *AS-C* proteins do not bind DNA as homodimers, but *da*/*AS-C* heterodimers bind with high affinity to specific "E-box" sequences present in the promoter of the *ac* proneural gene, and heterodimer binding to these sites activates *ac* transcrip-

³Corresponding author.

tion in vivo (Van Doren et al. 1991, 1992; Martinez et al. 1993). This is analogous to the activation of muscle-specific genes by MyoD/E12 and myogenin/E12 heterodimers in mammalian cells (for review, see Weintraub et al. 1991). Heterodimer formation also regulates DNA binding by the Myc and Max bHLH leucine zipper proteins (Blackwood and Eisenman 1991; Prendergast et al. 1991).

Heterodimer formation is also an important mechanism for mediating repression by HLH proteins. The *emc* repressor protein also contains an HLH domain, but it lacks a basic region and, therefore, cannot bind to DNA. *emc* protein forms non-DNA-binding heterodimers with the *da* and AS-C activator proteins, thereby repressing their DNA-binding activity in vitro and repressing *ac* transcription activation by *da*/AS-C heterodimers in vivo (Van Doren et al. 1991, 1992). In contrast with *emc*, the proline-bHLH repressor proteins all have basic regions, but unlike other bHLH proteins they each contain a proline in the basic region. The proline has been proposed to disrupt the DNA-binding activity of the basic region so that these proteins would function biochemically like *emc* and its mammalian homolog, Id (Benezra et al. 1990; Jones 1990). However, at least some proline-bHLH proteins can bind to DNA: The E(spl) *m8* protein and mammalian HES-1 protein have been shown to bind to noncanonical sequences (N boxes: CACGAG and CA-CAAG) that are present in the *m8* promoter region (Sasai et al. 1992; Tietze et al. 1992).

hairy has not been shown previously to bind to DNA, and several models of *hairy* repression invoke heterodimer formation between *hairy* and other bHLH proteins. Genetic analysis has shown that *hairy* specifically represses the activity of the *ac* gene in vivo (Moscato del Prado and Garcia-Bellido 1984). One model for this repression is that *hairy* forms inactive heterodimers with specific AS-C target proteins such as *ac*. These inactive heterodimers may be unable to bind to DNA, or they may bind DNA but be transcriptionally inactive (Jones 1990; Parkhurst et al. 1990). *hairy* also functions in segmentation as a repressor of *fushi tarazu* (*ftz*), and one model for this repressor function is that *hairy* forms DNA-binding heterodimers with a bHLH corepressor protein (Wainwright and Ish-Horowicz 1992).

In this study we combine biochemical and genetic approaches to directly determine whether *hairy* is a DNA-binding protein and the mechanism by which it acts to repress SO formation. Our analysis has been guided by previous structure/function analyses of bHLH proteins, which have shown that many bHLH proteins can be divided into two classes, depending on whether they bind to class A sites or class B sites (see Fig. 1; Dang et al. 1992). Although both class A and class B sites fit the canonical HLH-binding sequence, CANNTG, class A proteins do not bind to class B sites, and vice versa. These differences in DNA-binding specificity result from differences in the amino acid sequence of the basic region (Dang et al. 1992; Blackwell et al. 1993).

Here, we show that the *Drosophila* activator and proline-bHLH repressor proteins bind to class A and class B

sites, respectively. In addition, although *hairy* binds the class B canonical sites, it prefers a noncanonical site, CACGCG, and also binds other noncanonical sites. These noncanonical sites are closely related to class B sites and are defined as class C sites. We show that the same sites are recognized by other proline-bHLH proteins, and we define the structural features of sites recognized by that family of proteins. We then show that a CACGCG class C site present in the promoter of the *ac* proneural gene mediates repression of *ac* by *hairy* in vivo: Mutation of that site blocks repression of *ac* transcription by *hairy* in cultured cells. Moreover, in *Drosophila* the same *ac* promoter mutation creates ectopic sensory hairs like those seen in *hairy* mutants. These results indicate that *hairy* represses sensory organ formation by binding to a specific repressor site in *ac* and directly repressing transcription of the *ac* proneural gene.

Results

hairy has the key structural feature of a class B protein and binds to class B sites

Figure 1 shows an alignment of the basic regions of several class A and class B proteins with those of the *da* and AS-C activator proteins and the proline-bHLH repressor proteins. Previous analysis has shown that the Arg residue at position 13 (R13) present in all class B proteins is the key structural criterion for defining class B-binding specificity to the class B basic regions (see Fig. 1 legend; Dang et al. 1992; Van Antwerp et al. 1992; Blackwell et al. 1993). As shown in Figure 1, *hairy* and all of the other proline-bHLH proteins have R13 residues in their basic domains. This led us to predict that *hairy* would bind to class B sites but not to class A sites. Similarly, the lack of R13 in the *da* and AS-C proteins suggested that they would not bind to class B sites. Figure 2A shows that *da*/l'sc and *da*/sc heterodimers and *da*/*da* homodimers bind to the class A site CACCTG (Fig. 2A). However, those protein complexes do not bind to the class B CACGTG site. Furthermore, binding of those complexes to the labeled class A site is competed by an excess of cold (unlabeled) A site, but it is not competed by an equivalent 1000-fold excess of cold B site. Thus, the *da* and AS-C proneural proteins bind with high specificity to class A sites but not to class B sites, as predicted.

In contrast, Figure 2B shows that both full-length *hairy* protein and a truncated, bHLH version of *hairy* (containing only the bHLH domain) bind to the class B site, CACGTG. However, neither full-length nor bHLH *hairy* binds to the class A site. Moreover, binding by *hairy* to the class B site is not competed by a 1000-fold excess of class A sites, showing that *hairy* also is highly specific for the class B site. The class B protein Max shows the same general pattern of binding as *hairy*: Max binds to the class B site with high affinity, although it does cross-react somewhat with the class A site.

We next used polymerase chain reaction (PCR) site selection (Blackwell and Weintraub 1990; Sun and Baltimore 1991) to determine whether *hairy* would select

position:	Basic Region														
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
Class A	R	R	E	I	A	N	S	N	E	R	R	R	M	Q	S
	R	R	K	A	A	T	M	R	E	R	R	R	L	S	K
	R	R	Q	A	N	N	A	R	E	R	I	R	I	R	D
	S	V	A	R	R	N	A	R	E	R	N	R	V	K	Q
	S	V	Q	R	R	N	A	R	E	R	N	R	Y	K	Q
Class B	K	R	R	T	H	N	V	L	E	R	Q	R	R	N	E
	R	R	R	N	H	N	I	L	E	R	Q	R	R	N	D
	K	R	K	N	H	N	F	L	E	R	K	R	R	N	D
	K	R	A	H	H	N	A	L	E	R	K	R	R	D	H
	R	R	A	Q	H	N	E	V	E	R	R	R	R	D	K
proline-bHLH	K	K	D	N	H	N	L	I	E	R	R	R	R	F	N
	K	K	D	N	H	N	L	I	E	R	R	R	R	F	N
	R	R	S	N	K	P	I	M	E	K	R	R	R	A	R
	Q	K	V	K	K	P	M	L	E	R	Q	R	R	A	R
	L	K	V	K	K	P	L	L	E	R	Q	R	R	A	R
Canonical HLH site: C A N N T G															
Class A Sites: C A C C T G ; C A G C T G															
Class B Sites: C A C G T G ; C A T G T G															

Figure 1. Basic region similarities between class A, class B, and proline-bHLH proteins, and the class A and class B sites recognized by those proteins. AP-4 (Dang et al. 1992) and MyoD (Van Antwerp et al. 1992; Blackwell et al. 1993) are previously defined class A proteins, and Myc, Max, USF, TFE3, and TFEB are previously defined class B proteins (Dang et al. 1992). The classes are operationally defined in that class A proteins bind to class A sites but not to class B sites, etc. da/AS-C heterodimers can bind to sites that fit the class A consensus (Murre et al. 1989b; Cabrera and Alonso 1991; Van Doren et al. 1991), and their basic regions are similar to those of MyoD and AP-4. The proline-bHLH proteins have been aligned in a third group. The P (Pro) residues at position 6 in the proline-bHLH proteins, which is the defining feature of that class of bHLH proteins, are boxed. The R at position 13 (R13) in the class B and proline-bHLH proteins also is boxed. Basic amino acids R (Arg) and K (Lys) are shaded. Crystal structure analysis of protein-DNA cocrystals of the class B bHLH protein Max has shown that the R13 in Max makes direct contact with the G nucleotide at position 4 (G4) in the class B site, CACGTG (Ferre-D'Amare et al. 1993). The G4 nucleotide is the distinguishing feature of class B sites (Fig. 1, bottom), and by binding to the G4 nucleotide the R13 amino acid provides the key specificity determinant for class B bHLH proteins. The presence of an R13 residue in all of the proline-bHLH proteins suggests that they will bind to class B sites. The sequences are taken from AP-4 (Hu et al. 1990), MyoD (Davis et al. 1987), da (Caudy et al. 1988), ac, sc (Villares and Cabrera 1987), l'sc, ase (Alonso and Cabrera 1988; Gonzalez et al. 1989), c-, N-, and L-Myc (Batterly et al. 1983; Kohl et al. 1986; DePinho et al. 1987), Myn/Max (Blackwood and Eisenman 1991; Prendergast et al. 1991), USF (Gregor et al. 1990), TFE3 (Beckmann et al. 1990), TFEB (Carr and Sharp 1990), hairy (Rushlow et al. 1989), E(spl) m8, m5, m7, m3 (Klammt et al. 1989; Delidakis and Artavanis-Tsakonas 1992; Knust et al. 1992), HES-1 (Sasai et al. 1992), and HES-5 (Akazawa et al. 1992).

class B sites over class A sites when mixed with a pool of degenerate HLH canonical oligonucleotides; NCAN-

NTGN. hairy exclusively selected the CACGTG class B palindrome (Fig. 2C). In contrast, the class A protein da selected CACCTG in the same assays (data not shown). Thus, as predicted by the presence of the R13 residue in the basic region, hairy binds to, and also selects, class B sites in vitro.

hairy preferentially selects a noncanonical sequence

Although hairy clearly binds to class B sites, we were led to ask whether it could recognize additional, noncanonical sites. Analysis of other proline-bHLH proteins, the *Drosophila* E(spl) m8 protein (Tietze et al. 1992) and the mammalian HES-1 and HES-5 proteins (Akazawa et al. 1992; Sasai et al. 1992) has shown that those proteins can bind to a DNA sequence CCACGAGCCACAAGG present in the m8 promoter. This sequence contains two related hexamer sequences, CACGAG and CACAAG, that are called N-box sites (Sasai et al. 1992). The N boxes do not fit the HLH canonical sequence but are closely related to it. Gel retardation analysis has shown that hairy also can bind to the m8 double hexamer sequence (data not shown). Therefore, we wished to determine whether hairy would select the N-box hexamer sequences from C>NNNG oligonucleotides (Fig. 3A).

Surprisingly, hairy preferentially selected neither class B nor N-box sites (Fig. 3A). Rather, it preferentially selected a different noncanonical sequence, CACGCG, with the highest frequency (22/35). The class B CACGTG site was also selected with high frequency (9/35), and the CACGAG N-box site was selected with lower frequency (2/35). Given that hairy selects noncanonical sites, we asked whether the class B protein Max was also capable of selecting noncanonical sites. In contrast with hairy, Max exclusively selected the class B canonical sites, CACGTG and CATGTG (Fig. 3B). Like hairy, Max always selected sites with a G nucleotide at position 4, as predicted by the R13 residue present in both class B and proline-bHLH proteins. However, Max showed distinctly different selection than hairy at position 5 in the hexamer. Whereas hairy selected sequences with various nucleotides at position 5, Max exclusively selected sequences with T, which therefore are canonical class B sites.

The three classes of sites are closely related to one another, and in some cases, sites of distinct classes differ by only a single nucleotide within the core hexamer. The three classes are

Class A sites:	CACCTG	CAGCTG
Class B sites:	CACGTG	CATGTG
Class C sites:	CACGCG	CACGAG

where the CACGCG and CACGAG class C sites shown are the two noncanonical sites selected most frequently by hairy (Fig. 3A). These sites differ from the class B CACGTG site only by a single nucleotide change at position 5.

Figure 2. Binding site specificity of the da/AS-C activator proteins and the hairy repressor protein for class A and class B sites. (A) Gel retardation assays showing specific binding of the *Drosophila* neuronal activator proteins da, l'sc, and sc to a class A (CACCTG) site. In each set of four lanes, the indicated proteins were mixed with class A probe (lanes 1,5,9) or class B probe (lanes 2,6,10), either with (+) or without competitor. The patterns of binding are similar for each of the da/l'sc, da/sc, and da/da combinations and show that binding of da/AS-C activator proteins is highly specific for the class A site, despite the fact that the core hexamers of the class A and class B hexamers differ by a single nucleotide. Similar observations have been made for the CAGCTG class A sites (data not shown). (But note that flanking sequences can affect the binding affinity of specific class A and class B sites.) (da) 500 ng of full-length da protein; (da/l'sc) 500 ng of full-length da + 125 ng full-length l'sc protein; (da/sc) 500 ng of full-length da + 125 ng full-length sc protein. (B) Binding of full-length hairy, bHLH hairy, and the class B protein Max, to class A and class B sites.

In each set of four lanes the proteins show specific binding to class B sites. Full-length hairy binds less stably than bHLH hairy, resulting in a broadened band of retarded DNA; however, both show the same pattern of binding specificity. Max, a previously defined class B protein (Blackwood and Eisenman 1991; Prendergast et al. 1991), also shows the same pattern of binding. Although it does bind to the class A site, it does so with about a fivefold lower affinity than to the class B site. (hairy) 300 ng of full-length hairy protein; (h BHLH) 50 ng of bHLH hairy protein (bHLH hairy is a truncated protein containing only the bHLH region); (Max) 20 ng of bHLH Max protein (Ferre-D'Amare et al. 1993). (C) PCR site selection of class B sites by hairy. Highly purified, full-length hairy protein was used in the PCR-assisted site selection, which was performed as described in Materials and methods. The actual sequences selected show the number of clones (of 20 total that were sequenced) that had a specific sequence that was bound and selected by hairy from the degenerate oligonucleotide pool, NCANNTGN. The frequency histogram shows the frequency with which a given nucleotide was found at a specific position within the overall set of 20 sequenced sites.

We then performed gel retardation assays to further examine the relative affinities of hairy and Max for the class B (CACGTG) and class C (CACGCG) sites. Figure 3C shows that Max binds very strongly to the class B site but only weakly to the class C site. In addition, a 1000-fold excess of class C site cannot compete the class B site away from Max. In contrast, full-length hairy and bHLH hairy bind to both of the class B and class C sites, and binding to the class B site can be competed by an excess of either class B or class C site.

Thus, PCR site selection and gel retardation assays both show that the class B protein Max does not have the same sequence specificity as the proline-bHLH protein hairy. Rather, hairy appears to be a class B protein with an expanded repertoire of binding sites, which include

the class C CACGCG and the CACGAG sequences, and possibly other related noncanonical sites.

Another proline-bHLH protein, E(spl) m5, also binds to class B and class C sites in vitro

We wished to determine whether other proline-bHLH proteins also recognize the class B and class C sites. As shown in Figure 1, the hairy and E(spl) m5 proteins represent relatively divergent members of the proline-bHLH family: Together, they span the range of sequence variations seen within the basic regions of proline-bHLH proteins. Figure 4 shows that m5 binds specifically to the class B and class C sites but not to the class A site, and binding to the class B and C sites cannot be

Having identified the general set of binding sites recognized by hairy in vitro, we asked whether hairy binds to those sites to repress the *ac* proneural gene in vivo. Previous analysis of the *ac* promoter region has shown that there is a cluster of three sites to which da and AS-C heterodimer complexes bind to activate *ac* transcription in vivo (Fig. 5A; Van Doren et al. 1991, 1992; Martinez et al. 1993). These are sites that we now recognize as class A sites. Having identified CACGCG as a high-affinity binding site for hairy in vitro, we observed that there is a CACGCG site ~50 bp upstream of the class A cluster in the *ac* promoter (Fig. 5A). Previous work using cultured *Drosophila* cells has shown that cotransfection of an *ac* reporter gene with da and AS-C expression con-

If the CACGCG site is mutated (Fig. 5C), *da* and *AS-C* still activate *ac* transcription to the same level, and addition of *emc* still can repress that activation. However, *hairy* no longer represses *ac* transcription. Thus, *hairy* requires the CACGCG site for repression. Given that this is the preferred binding site for *hairy* in vitro and that the presence of the binding site does not mediate

Figure 4. Binding by E(spl) m5 to class B site and class C sites. m5 protein binds to the class B and the class C sites but not to the class A site. m5 is competed away from the class B site and partially competed from the class C site by an excess of class C DNA but not by an excess of class A DNA. (m5) 300 ng of full-length m5 protein was used per reaction in this experiment.

repression of *ac* by endogenous proline-bHLH proteins in the absence of *hairy*, it appears that *hairy* directly represses *ac* transcription by binding to the CACGCG site in the *ac* promoter region.

hairy represses sensory organ formation in *Drosophila* through the CACGCG site in the *achaete* promoter

To test for *hairy* binding to the CACGCG site in *Drosophila*, we mutated the site in an *ac* minigene and introduced it into *Drosophila* by P-element transformation (Rubin and Spradling 1982). Previous work has shown that an *ac* minigene can rescue many of the sensory hairs that are missing in *achaete* null (*ac*⁻) mutants (G. Panigian and S. Carroll, unpubl.). Therefore, this provides a favorable in vivo system for directly testing the function of the CACGCG-binding site in the repression of sensory organ formation, not just in the repression of *ac* transcription.

Analysis of adult wings from transformants revealed that disruption of the CACGCG-binding site resulted in the presence of ectopic hairs (bristles) along the veins of the wing (Fig. 6). In wild-type flies there are no sensory

hairs along the L2 wing vein (Fig. 6, top) or other non-margin veins. In *hairy*¹ mutants there are many ectopic hairs on the L2 vein (Fig. 6, middle). Figure 6, bottom, shows that a similar phenotype is produced in transformants in which the single CACGCG site present in the *ac* promoter region has been mutated. In 10 of 13 homozygous transformant lines generated with a mutated CACGCG site there were large numbers of ectopic hairs on the wings and body and these hairs were in similar

Figure 5. *hairy*-mediated repression of *achaete* transcription via a CACGCG-binding site in the *ac* promoter region. (A) Schematic representation of the *ac* proximal promoter region (Villares and Cabrera 1987). A cluster of three class A sites is upstream of the TATA box in the *ac* promoter. A CACGCG (class C) hairy-binding site is located immediately upstream of the three class A sites. (B) Effect of *hairy* and *emc* on the transcriptional activation of the native *ac* promoter by da/AS-C activator proteins. Cotransfection of *Drosophila* Schneider cells with da and sc expression constructs results in activation of *ac* transcription. Additional cotransfection with either *hairy* or *emc* results in repression of *ac* transcription. The reporter construct contains 0.9 kb of the wild-type *ac* 5' promoter fused to a luciferase reporter gene (see Materials and methods). Values are the mean of at least four samples and error bars represent the standard error. (C) Effect of mutating the CACGCG site on repression of *ac* transcription by *hairy* and *emc*. Mutation of the CACGCG-binding site does not affect activation by da/sc proteins. In addition, *emc* still represses *ac* transcription by da/sc activator complexes in the absence of the CACGCG site. However, in contrast, *hairy* does not repress in the absence of the site. Thus, the CACGCG site is essential for *hairy*-mediated repression of *ac* transcription in cultured *Drosophila* cells.

Figure 6. Derepression of sensory hair formation in *Drosophila* by mutation of the CACGCG site in an *ac* minigene. (Top) Wild-type wings have sensory hairs (bristles) only along the wing margins; the L2 veins and other nonmargin veins (not shown) are normally completely devoid of hairs. (Middle) In *hairy*¹ mutants loss of *hairy* repressor function results in the formation of ectopic sensory hairs along the L2 veins and other veins. Previous work has shown that *ac* expression along L2 is derepressed in *hairy*¹ mutants (Skeath and Carroll 1991) and that *hairy* functions genetically as a specific repressor of the *ac* proneural gene (Moscoso del Prado and Garcia-Bellido 1984). (Bottom) Disruption of the CACGCG site in an *ac* minigene results in the formation of ectopic sensory hairs along the L2 vein. This line (M4) has an intermediate phenotype relative to the strongest transformant line (M1, see Table 1) but it is comparable to the phenotype of *hairy*¹ mutants. In contrast, none of the control lines had ectopic hairs on L2, and the controls at most had two hairs on L5 (Table 1). The wing photographs are 175× magnifications of the anterior wing margin and the L2 vein.

positions as in *hairy*¹ mutants (Table 1A). In contrast, the control lines were essentially wild type, with no

more than two hairs on any wing (one or less, on average). Similar ectopic hair phenotypes were seen in both the presence and absence of the endogenous, wild-type *ac* gene (Table 1B), indicating that the ectopic hairs result from derepression of the *ac* minigene, not from an increase in overall expression of *ac* product in the presence of additional copies. Similar phenotypes were also seen in *hairy*¹ mutants (data not shown), indicating that *hairy* represses *ac* primarily by acting through the CACGCG site. Both mutant and control transformants rescued most of the hairs on the notum that are missing in *ac*[−] mutant flies; thus, mutation of the CACGCG site does not affect activation of the *ac* minigene (data not shown).

The phenotype observed in the CACGCG[−] mutants further mimics the *hairy*¹ phenotype in that ectopic hairs also are present at other locations like those seen in *hairy*¹ mutants, for example, on the mesopleurae and head (data not shown). This phenotype is distinct from that seen in clones deficient for neurogenic genes such as *E(spl)*, in which clusters of hairs are seen at the normal positions (P. Heitzler and P. Simpson, pers. comm.) rather than single hairs at many ectopic positions, as in *hairy*¹. Thus, it appears that the ectopic sensory hairs created in the *ac* binding site mutants results from the lack of hairy protein binding to that site rather than the lack of *E(spl)* protein binding to the *ac* promoter.

Discussion

Mechanism of action of hairy as a proline-bHLH repressor of sensory organ formation

We have combined biochemical and genetic analyses to determine the mechanism by which the proline-bHLH protein hairy specifically represses the *ac* proneural gene. We have shown that despite the presence of the proline residue in the basic region, hairy is a sequence-specific DNA-binding protein with novel DNA-binding specificity. hairy binds as a homomer to class B canonical HLH sites, as predicted by the presence of the key amino acid residue for class B binding in its basic region (Fig. 1). However, unlike the class B protein Max, hairy binds preferentially to a site, CACGCG, that does not fit the HLH canonical sequence (Fig. 3). Determination of hairy-binding sites in vitro has allowed us to identify a single CACGCG site present in the *ac* promoter region (Fig. 5A). In a cell culture assay, hairy represses transcriptional activation of the native *ac* promoter by da/AS-C proteins. However, mutation of the CACGCG promoter site results in blocking repression of *ac* transcription by hairy, although repression by *emc* is not affected. Thus, the CACGCG site is essential for repression of *ac* by hairy. Moreover, mutating that site in a *ac* minigene transformed into *Drosophila* creates ectopic sensory organs like those seen in *hairy*¹ mutants (Fig. 6).

The above results provide strong evidence that hairy represses sensory organ formation by directly repressing transcription of the *ac* proneural gene, and similar results have been found by another group (van Doren et al.

Table 1. Ectopic hair formation in *achaete* transformants with hairy-binding site mutations

A		L2 L5 TOT					L2 L5 TOT		
		1	3	7	1	7	9	3	
	<i>hairy</i>	1	3	7	1	7	9	3	
	M1	4	8	2	7	1	3	5	
	M2	3	1	2	7	9	6		
	M3	3	6	2	3	9	0		
	M4	2	8	2	3	8	6		
	M5	3	0	2	0	6	8		
	M6	2	5	1	8	6	5		
	M7	2	5	2	0	6	5		
	M8	2	4	1	7	6	2		
	M9	1	6	1	3	6	0		
	M10	1	9	1	7	5	5		
	M11	0	0	4	6	5			
	M12	0	0	0	0	0			
	M13	0	0	0	0	0			

B		ac- (Df[ac])						ac+					
		females			males			females			males		
		L2	L5	TOT	L2	L5	TOT	L2	L5	TOT	L2	L5	TOT
	M1	55	30	140	48	35	151	59	32	145	48	27	135
	M4	26	25	88	26	23	86	28	21	75	28	23	87
	M3	43	31	93	29	23	78	11	14	34	36	23	90
	C1	0	1	0.7	0	0	0.4	0	0	0.3	0	1	1
	C5	0	0	0	0	1	0.8	0	1	0.5	0	1	0.7
	C4	0	0	0.2	0	0	0.2	0	1	0.5	0	1	0.7

M1, M2, etc., are transformant lines with the CACGCG site mutated; C1, C2, etc., are control lines with identical *ac* minigenes, except that the CACGCG site has not been mutated. The numbers given under the L2, L5, and TOT headings are the average number of hairs seen along the L2 vein, L5 vein, and total wing, respectively. In all cases, the averages are of 10 wings, each from different flies. The mutant and control lines have been numbered according to the strength of their phenotypes, with the strongest lines first. Both the mutant and control lines rescued most of the hairs missing from the notum and other regions in *ac* flies. (A) Thirteen mutant lines were viable as homozygotes, and 10 of these 13 had clear hairy-like phenotypes. One line had substantially more hairs (135 total per wing) than are seen in *hairy*¹ mutants. Three of the lines had ectopic hairs in numbers (80–100) comparable to those seen in *hairy*¹ mutants (93). Six lines had 50–60 hairs per wing. These ectopic hairs were present in patterns similar to those seen *hairy*¹ mutants (Ingham et al. 1985; Lindsley and Zimm 1992). Three lines had few or no hairs, apparently because of position effects resulting from their P elements having inserted in transcriptionally inactive regions of the genome (Spradling and Rubin 1983). Overall, the lines with the mutated CACGCG site exhibited a clear hairy-like phenotype. (B) Ectopic hair formation in homozygous transformants in the presence or absence of the endogenous *ac* gene. Similar ectopic hair phenotypes were seen in both the presence and absence of the endogenous, wild-type *ac* gene, and in both males and females. In contrast, the control lines were essentially wild type, with no more than two hairs on any wing (one or less, on average); these hairs were invariably on the mid- to proximal region of the L5 vein and never on the L2 vein. This region appears to be a "hot spot" for expression of the truncated *ac* minigene present in the transposon, and/or a region very sensitive to low levels of ectopic *ac* protein expression.

1994). These results are further supported by previous genetic evidence that *hairy* represses SOP formation by specifically repressing *ac* activity (Moscoso del Prado and Garcia-Bellido 1984). During wild-type wing development, *hairy* is expressed along the wing veins, and *ac* is not (Skeath and Carroll 1991), but in *hairy* mutants, *ac* protein is ectopically expressed along the wing veins, indicating that *hairy* represses *ac* expression, not the activity of *ac* protein. We cannot exclude the possibility that *hairy* activates the expression or activity of another proline-bHLH (or other) protein that then binds to the CACGCG site to mediate *ac* repression. However, given that CACGCG is the preferred site for *hairy* binding in vitro (Fig. 3), that *hairy* represses *ac* through that site in Schneider cells (Fig. 5), and that the resulting phenotype exactly mimics the *hairy* mutant phenotype rather than that of *E(spl)* or some other proline-bHLH mutant, we conclude that *hairy* directly represses *ac* transcription by binding to that site in vivo.

How does *hairy* binding upstream of the class A activator sites repress *ac* expression? Previous work (Van Doren et al. 1992; Martinez et al. 1993) has shown that *da*/AS-C activator proteins bind to the class A binding sites that are immediately upstream of the TATA box in the *ac* gene and thereby activate *ac* transcription in vivo. As we have shown, *hairy* does not bind to the class A activator-binding sites in vitro so that *hairy* does not appear to function as a "passive repressor" by competing with activator proteins for those sites (Jaynes and O'Farrell 1991; Cowell 1994). Rather, *hairy* binds to the distinct, distally located CACGCG repressor-binding site, and mutation of that site results in derepression of *ac* both in Schneider cells and in *Drosophila*. This indicates that *hairy* functions as an "active repressor" (Fig. 7), that interferes with *ac* transcription presumably through an

Figure 7. Model for active repression of the *achaete* proneural gene by hairy. hairy protein binds to the CACGCG class C site that lies distal to the three class A sites that are bound by *da*/AS-C heterodimers. Binding of the activator proteins to the class A sites facilitates (arrows) the activation of transcription of the *ac* gene. hairy does not bind to the class A sites (Fig. 3) and therefore does not passively repress by competing with activator proteins for those sites. Rather, hairy bound to the distally located CACGCG site actively represses *ac* transcription by interfering with the activation of *ac* transcription, apparently by interacting with either the bound activator HLH proteins or with more proximally bound components of the general transcriptional machinery.

intrinsic repression domain (Jaynes and O'Farrell 1991; Cowell 1994). Although an intrinsic repression domain has not been identified in hairy, it does have a polyglutamine/polyalanine domain similar to the repression domains observed in other active repressor proteins (Cowell 1994). hairy also has a carboxy-terminal WRPW amino acid domain that is present in all of the proline-bHLH repressor proteins and may function as a repression domain.

Our results do not exclude the possibility that hairy, or other proline-bHLH proteins, may form non-DNA-binding heterodimers with *da* or *AS-C* proteins in other contexts. There is evidence that the mammalian HES-1 proline-bHLH protein can repress transcriptional activation of an artificial target gene in cultured mammalian cells both by binding to N-box sequences and by directly interfering with activator protein function (Sasai et al. 1992), and similar repression events may occur in *Drosophila*. In addition, although our results clearly show that hairy can bind to DNA as a homodimer in vitro, we cannot exclude the possibility that hairy binds as a heterodimer with other bHLH proteins in Schneider cells or in *Drosophila*.

The proline-bHLH proteins form a third class of bHLH proteins

Our results extend the class A versus class B scheme proposed by Dang and co-workers (Dang et al. 1992) to include a third class, class C, which are the proline-bHLH proteins. These proteins have distinct basic region sequences and other distinct domains not found in the class A and class B proteins. None of the class A or class B proteins have the WRPW amino acid domains that are present in all proline-bHLH proteins. In addition, whereas all class B proteins have a leucine zipper dimerization domain adjacent to their bHLH domain, none of the proline-bHLH proteins have a leucine zipper. Thus, the class C proteins are structurally distinct overall from the class A and class B proteins.

The class C basic regions have DNA-binding properties that are clearly distinct from the class A basic regions. The class C repressor proteins bind to both class B and class C sites as we have shown for hairy and the E(spl) m5 protein. The binding to class B sites is presumably attributable to the presence of the R13 residue in all class C proteins, as R13 residues in class B proteins directly contact the G4 nucleotide that distinguishes class B sites from class A sites (Ferre-D'Amare et al. 1993). All of the class C sites also contain a G4 nucleotide, but they vary from the class B sites at position 5 in the core hexamer. Class C sites include the CACGAG N-box site that is recognized by the E(spl) m8 protein (Tietze et al. 1992) and by mammalian HES homologs of the hairy and E(spl) proteins (Akazawa et al. 1992; Sasai et al. 1992). Together, the results for the hairy, m5, m8, and HES proteins indicate that many or all proline-bHLH proteins will bind to class B and class C sites.

The class C basic regions are clearly distinct from the class B regions, and this may reflect different DNA-bind-

ing properties. Another group has found that Myc-Max heterodimers occasionally (2/59) select CACGCG sites in PCR site selection assays (Blackwell et al. 1993). However, in our studies, Max exclusively selected class B sites under the same conditions in which hairy selected both class B and class C sites. Fusion of the hairy basic region onto the Max HLH-zipper domain produces a chimeric protein that binds as well as a homodimer to class C sites (T. White and M. Caudy, unpubl.), confirming that some feature of the hairy basic region confers high-affinity binding to class C sites.

Structure/function relationships within bHLH protein classes

Our results suggest that precise structure/function relationships may hold for the various classes of bHLH proteins. For the *Drosophila* proline-bHLH proteins and class A proteins, there are distinct biological functions associated with the different protein classes. The hairy and the E(spl) proline-bHLH proteins all function as repressors of sensory precursor formation during nervous system development. In contrast, the class A *da* and *AS-C* proteins clearly function genetically as activators of neuronal (sensory) precursor formation.

In general, class A proteins tend to function as activators of cell determination pathways in both vertebrates and invertebrates. In both *Drosophila* and mammals, neuronal cell determination pathways are mediated by the proneural *da* and *AS-C* bHLH genes or their homologs in other systems (Campos-Ortega 1993; Guillemot et al. 1993; Jan and Jan 1993b). Muscle cell determination is regulated by another set of class A bHLH proteins, the myogenic bHLH factors that include MyoD, myogenin, MRF4, and myf-5 (Weintraub et al. 1991). Like the *AS-C* proteins and homologs, the myogenic class A proteins bind to DNA as heterodimers with the multifunctional E12 and E47 class A proteins, which are homologs of *da* (Brennan and Olson 1990; Lassar et al. 1991).

The class B proteins are the least understood in terms of biological function. Although it appears that the Myc family is involved in apoptosis, cell proliferation control and/or repression of cell differentiation (Marcu et al. 1992), it seems unlikely that all of the class B proteins have similar functions during normal development. What does seem clear, though, from the biochemical analysis of the class B proteins is that they are likely to be competing for binding to the same DNA sites in target genes, and thus interacting at some level in vivo.

The classification of bHLH proteins as class A, class B, and class C proteins provides a useful framework for analyzing and understanding how those proteins function in gene regulation. Although this may be an oversimplification and some exceptions will likely be found, the utility of this framework lies in its predictive power, such as the ability to successfully predict that the proline-bHLH proteins would bind to class B sites because they have R13 residues. As illustrated here for hairy, the detailed knowledge of the binding sites recognized by

these various classes of proteins is useful in identifying candidate *in vivo* binding sites in genetically defined targets. By combining the biochemical analysis of hairy *in vitro* with molecular genetic analysis *in vivo*, we have been able to identify a site present in the *ac* proneural gene through which hairy acts as a specific repressor of *ac* transcription and thereby represses sensory organ formation in *Drosophila*.

Materials and methods

Drosophila stocks

Flies were raised on the standard yeast/cornmeal/molasses/agar medium. The w^- , $y^{3PLsc^{8R}}$ strain is described in Lindsley and Zimm (1992) and is used here as an *ac*⁻ background for P element-mediated transformation.

Plasmid constructions

Nomenclature for plasmids is as follows; T3=l'sc, T4=sc, and T5=ac. da, T3, T4, h, hbHLH, and m5 bacterial expression vectors were constructed from pET14b (Novagen) that has a polyhistidine leader sequence and thrombin cleavage site added to a conventional pET3 expression construct. pET14b-da: The 2.5-kb *Bgl*III fragment of a *da* genomic DNA clone was cloned into Bluescript (BS, Strategene) and subsequently cloned into the *Bam*HI site of the pET3a, generating pET3a-da (generously provided by H. Vaessin, Ohio State University, Columbus). This construct encodes a fusion protein that contains the pET gene 10 leader sequence, 12 amino acids encoded by the BS polylinker, and 4 amino acids encoded by the 5' noncoding region of *da* fused to the full 710-amino-acid native *da* protein. The pET3a-da *Nde*I-*Eco*RI fragment was then transferred into pET14b, generating pET14b-da. pET14b-T3: A *Bgl*III linker was added to a *Pvu*II site (at nucleotide 43) of a T3 cDNA clone, and the resulting *Bgl*II-*Bam*HI (at nucleotide 631) fragment was cloned into the *Bam*HI site of the pET3b. The *Bam*HI-*Eco*RI fragment of the 3' portion of the T3 cDNA clone was cloned into the *Bam*HI-*Eco*RI site of the above pET3b, generating pET3b-T3 (H. Vaessin). This contained the region corresponding to the T3 16–259 amino acids and was transferred into pET14b, generating pET14b-T3. T4, pET14b-T4: The PCR fragment (nucleotides 39–754 in T4 cDNA) was cloned into the *Sma*I site in BS into which a *Bam*HI linker was added to *Eco*RV site and the resultant *Bam*HI fragment was cloned into the *Bam*HI site of the pET3c, generating pET3c-T4 (H. Vaessin). This construct encodes a protein containing the pET gene 10 leader sequence, an additional 7 amino acids from the BS polylinker, and amino acids 14–345 of T4. The *Nde*I-*Hind*III fragment from pET3c-T4 was shuttled into pET14b to make pET14b-T4. pET14b-h: The hairy cDNA clone (BSD2E6, generously provided by C. Rushlow, Columbia University, New York) was digested with *Nde*I filled in with Klenow DNA polymerase to kill this site, and then mutated at a *Bst*EII site at the amino terminus to generate an *Nde*I site at the initiation codon of the open reading frame. This construct was then shuttled into pET14b by *Nde*I-*Eco*RI-directed cloning, generating an expression construct for full-length hairy protein fused to the polyhistidine leader encoded by pET14b. pET14b-h bHLH: the region encoding the bHLH domain of hairy was PCR amplified with primers that added *Nde*I and *Bam*HI sites at the 5' and 3' ends, respectively, which allowed directed cloning into *Nde*I-*Bam*HI-digested pET14b to yield pET 14b-h bHLH that expressed a truncated hairy protein consisting of amino acids 29–96 in which methi-

onine was substituted for serine at residue 29. The PCR amplification was performed using the primers 5'-TGGGAATTC-CATATGAGTGACCGTCGGTCAAC-3' and 5'-AGGGATC-CAGATCTACTGCATGGCTGCCTGCTG-3'. m5, pET14b-m5: The DNA fragment containing the m5-coding region was made by PCR using primers 5'-TGGGAATTCATATGGCAC-CACAGAGCAACAAC-3' and 5'-TGGGATCCACATTTAC-CAAGGGCGCCACATGGT-3' from an m5 cDNA clone (cm5h, generously provided by C. Delidakis and S. Artavanis-Tsakonas, Yale University, New Haven, CT). The PCR product was digested with *Nde*I and *Bam*HI and cloned into the *Nde*I-*Bam*HI site of the pET14b vector, generating pET14b-m5.

Reporter plasmids, which consisted of the native *ac* promoter and upstream region fused to the luciferase gene, were constructed as follows: BS SK—containing the 1.6-kb *Eco*RI fragment of the *ac* genomic clone (λ sc-112, generously provided by J. Modolell, Centro Biologica Molecular, University of Madrid, Spain) was digested with *Eco*47III and *Eco*RV to remove the coding region and then self-ligated. This plasmid containing 0.9 kb of *ac* upstream sequence was digested with *Sac*I and *Hind*III, and the resultant DNA fragment was cloned into the *Sac*I-*Hind*III site of pGL2-Basic vector (Promega) containing the luciferase reporter gene, generating pT5-0.9 wt/luc. To examine the functional significance of the hairy-binding site, CACGCG, the site was mutated to a *Sma*I site, CCCGGG, using an oligonucleotide-directed *in vitro* mutagenesis kit (Amersham), generating pT5-0.9 mut/luc. Expression vectors controlled under the *Drosophila* actin 5C promoter, pAct5CPPA, and pAcda, pAct4, and pAcemc, were generously provided by J. Colgan and J. Manley, and M. Van Doren and J. Posakony, respectively (Han et al. 1989; Van Doren et al. 1992). The expression vector pACh was constructed as follows: The *Not*I-*Eco*RI fragment of the 3' portion of the hairy cDNA clone (BSD2E6) was cloned into the *Not*I-*Eco*RI site in BS into which a *Bgl*III linker was added to an *Eco*RV site. The *Sac*I-*Bgl*III fragment of this DNA was cloned into the *Sac*I-*Bgl*III site of pAct5CPPA. The *Kpn*I-*Not*I fragment of the 5' portion of the hairy cDNA clone (BSD2E6, 10#3) was cloned into the *Kpn*I-*Not*I site of the above DNA, generating pACh.

P-element vectors containing the wild-type (p[T5-C]) and the hairy-binding site mutant ([pT5-M]) of the *ac* minigene were constructed as follows: p[T5-C] and p[T5-M] were generated by sequentially cloning two adjacent genomic fragments of *ac* DNA into a P-element vector. First, a 3.2-kb *Bam*HI fragment (coordinates 59–56 of the AS-C; Campuzano et al. 1985) was cloned into the *Bam*HI and *Bgl*III sites of Casper 2 (Thummel et al. 1988) with the 3' *Bam*HI site of the fragment ligated to the *Bgl*III site of Casper to generate pG1 (G. Panganiban, J. Skeath, and S. Carroll, pers. comm.). Then the 0.8-kb *Hind*III-*Bam*HI fragment of BS SK containing the 1.6-kb *Eco*RI fragment of the *ac* genomic clone (λ sc-112) or the mutated one that was described above, was cloned into the *Bam*HI site and the polylinker *Hind*III site of pG1 giving p[T5-C] or p[T5-M].

Production and purification of recombinant proteins

Proteins were expressed in BL21(DE3)lysS by induction with 0.4 mM IPTG for 1.5–2 hr at room temperature. da, l'sc, sc, hairy, and m5 proteins were purified from the inclusion bodies using His-Bind Resin (Novagen) according to the manufacturer's protocol. hairy bHLH protein was purified from the supernatant using His-Bind Resin after polyethyleneimine precipitation to remove bacterial DNA and ammonium sulfate precipitation (final 50% saturation). The purified proteins from inclusion bodies were renatured by dialyzing sequentially against a thrombin cleavage buffer (20 mM Tris-HCl at pH 8.4, 150 mM NaCl, 2.5

mm CaCl₂, 0.1% Triton X-100, 10% glycerol) containing 4 M urea, then 2 M urea, and then straight buffer. Next, they were cleaved by human thrombin to remove the polyhistidine leader and were dialyzed against 50 mM Tris-HCl (pH 7.5), 150 mM NaCl, 0.1% Triton X-100, 10% glycerol, 1 mM DTT, and 1 mM EDTA. The Max bHLH leucine zipper protein was kindly provided by A. Ferre-D'Amare and S. Burley (Ferre-D'Amare et al. 1993). This Max protein is identical to the protein used for X-ray crystallography analysis. Proteins were quantitated with bovine serum albumin (BSA) as the standard, and the quantitations were also confirmed by Coomassie blue staining of the proteins fractionated by SDS-PAGE with protein standards of known concentrations. In general, the proteins purified using the histidine leader were >90% pure, as estimated by Coomassie gel analysis.

Gel retardation assays

In all of the DNA-binding assays shown, highly purified proteins were used. In all cases, the class A site used was the CACCTG hexamer and flanking sequences from the *ac* promoter region (see Fig. 5A); the class B site used was the CACGTG hexamer and flanking sequences from the upstream stimulatory factor-binding site in the adenovirus major late promoter (Gregor et al. 1990); and the class C site used was the CACGCG site from the *ac* promoter region (Fig. 5A). The exact sequences of the DNA probes used are CACCTG (CAGGTG on the opposite strand), 5'-GATCGTCACGCAGGTGGGATC-CCTA-3' and 5'-GATCGGATCCACCTGCGTGACTAC-3' for the opposite strand for the class A site; and (CACGTG), 5' primer, 5'-GATCCAGGTACCGTGGCCTG-3' and 5'-AATTCAGGCCACGTGACCTG-3' on the opposite strand for the class B site; (CACGCG) 5'-GATCGCCGGCACGCGACAGG-3' and 5'-AATTCCTGTCCGCTGCCGCGC-3' on the opposite strand for the class C site. Gel retardation assays were performed as described (Benezra et al. 1990) except that the DNA-binding reaction mixture contained 10 mM DTT instead of 1 mM DTT and 0.3 mg/ml of BSA.

PCR-assisted DNA-binding site selection

PCR-assisted DNA-binding site selection was performed essentially as described previously (Blackwell and Weintraub 1990; Sun and Baltimore 1991). Briefly, 38-bp degenerate oligonucleotides (5'-CTAGTGGATCCTGGTNCANNTGNGCAAGAATTCGAGGG-3') or 36-bb degenerate oligonucleotides (5'-CTAGTGGATCCTGGTNCNNNGGCAAGAATTCGAGGG-3') were converted to double-stranded DNA by Klenow DNA polymerase after priming with the 15-bp 3' primer (5'-CCCTCGAAT-TCTTG-3'). This product was purified on a 10% polyacrylamide gel, eluted at 37°C overnight in 10 mM Tris-HCl (pH 8.0), 1 mM EDTA, and 100 mM NaCl, and recovered by ethanol precipitation. About 20 ng of this DNA and 5 × 10⁴ cpm of labeled class B probe was incubated separately with 300 ng of the full-length hairy protein or 20 ng of the Max bHLH leucine zipper protein in the binding buffer for 15 min, and the complexes were separated on a 4% native acrylamide gel in 1 × TBE buffer (50 mM Tris, 50 mM boric acid, 1 mM EDTA). The region next to the radioactive band revealed by class B probe was excised from the dried gel and eluted at 65°C for 2 hr in 10 mM Tris-HCl (pH 8.0), 1 mM EDTA, 50 mM NaCl, and 0.2% SDS. The eluate was extracted with phenol and chloroform, and precipitated with ethanol. A fraction of the sample (1/25) was amplified for 35 cycles in a 100-μl PCR reaction (Perkin-Elmer Cetus) using a 5'

primer (5'-CTAGTGGATCCTGGT-3') and the 3' primer. The products were purified on a 10% acrylamide gel as described above. After three rounds of selection, oligonucleotides were cloned into pT7Blue T-Vector (Novagen). DNA sequences were determined using the AmpliTaq Cycle Sequencing kit (Perkin-Elmer Cetus).

DNA transfection and transient expression assay

DNA transfection into *Drosophila* Schneider L2 cells (generously provided by J. Colgan and J. Manley, Columbia University, NY) was performed as described previously (Han et al. 1989). Ten micrograms of DNA was used per 60-mm culture dish, and 2 μg of pCopia LTR-lacZ (J. Colgan and J. Manley) was included as an internal control, 0.5 μg of either the pT5-0.9 wt/luc or pT5-0.9 mut/luc reporter plasmids, 3 μg total of expression vector DNAs (with or without cDNA insert), and 4.5 μg of BS as a carrier DNA. Following 2 days of incubation with the CaPO₄ precipitate, cells were washed with PBS, harvested by scraping with a rubber cell scraper, and lysed in 250 μl of the cell lysis buffer (25 mM Tris-PO₄ at pH 7.8, 2 mM EDTA, 20 mM DTT, 10% glycerol, 1% Triton X-100). Luciferase activity was determined using a Luciferase assay kit (Promega). β-Galactosidase activity was determined as described previously (Han et al. 1989). Luciferase activity was normalized to β-galactosidase activity to control for variations in transfection efficiency.

Genetic analysis of *ac* transgenic flies

Transgenic flies were generated using standard procedures of microinjection (germ-line transformation) and P element-mediated transposon mobilization (Rubin and Spradling 1982; Laski et al. 1986). A *w*⁻, *y*^{3PL_{sc}8R} (*ac* deficiency) line was used as the recipient strain. Flies heterozygous for the *ac* minigene-mutant construct (p[T5-M]) show very few ectopic hairs, whereas homozygous flies have a large number of ectopic hairs on the wing and on the mesopleura. Flies heterozygous for the *ac* minigene-control construct (p[T5-C]) never show any hairs; the homozygous flies show one or two ectopic hairs at a very low frequency (see Results). Results from genetic crosses indicate that all lines are single insertions. For counts of wing hairs, wings were removed from flies and mounted in 80% glycerol in PBS.

Acknowledgments

We thank Paul Castella for excellent technical assistance, and Adrian Ferre-D'Amare and Stephen Burley for the generous gift of highly purified Max bHLH-Zip protein. Colleagues who provided plasmid constructs are acknowledged throughout the text. We also thank Al Fisher, Stephen Burley, Tom Kadesch, Peter Gergen, Steve DiNardo, Claude Desplan, Harold Weintraub, Susan Parkhurst, and Sean Carroll for stimulating discussions and critical comments on an early version of the paper. Finally, we thank Mark Van Doren and Jim Posakony for exchanging information regarding hairy function prior to publication. This work was funded in part by the Uehara Memorial Foundation (S.O.), and by the Howard Hughes Medical Institute (G.P.). M.C. was funded by grants from the National Institutes of Health (R01 NS28652), the Alfred P. Sloan Foundation, the Pew Scholars in Biomedical Sciences Program, the Mather Foundation, the Cornell Scholars program, and also by an institutional grant from the Markey Foundation.

The publication costs of this article were defrayed in part by payment of page charges. This article must therefore be hereby marked "advertisement" in accordance with 18 USC section 1734 solely to indicate this fact.

References

- Akazawa, C., Y. Sasai, S. Nakanishi, and R. Kageyama. 1992. Molecular characterization of a rat negative regulator with a basic helix-loop-helix structure predominantly expressed in the developing nervous system. *J. Biol. Chem.* **267**: 21879–21885.
- Alonso, M.C. and C.V. Cabrera. 1988. The achaete-scute gene complex of *Drosophila melanogaster* comprises four homologous genes. *EMBO J.* **7**: 2585–2591.
- Artavanis-Tsakonas, S. and P. Simpson. 1991. Choosing a cell fate: A view from the Notch locus. *Trends Genet.* **7**: 403–408.
- Battery, J., C. Moulding, R. Taub, W. Murphy, T. Stewart, H. Potter, G. Lenoir, and P. Leder. 1983. The human c-myc oncogene: Structural consequences of translocation into the IgH locus in Burkitt lymphoma. *Cell* **34**: 779–787.
- Beckmann, H., L.K. Su, and T. Kadesch. 1990. TFE3: A helix-loop-helix protein that activates transcription through the immunoglobulin enhancer muE3 motif. *Genes & Dev.* **4**: 167–179.
- 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.
- Blackwell, T.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.
- Blackwell, T.K., J. Huang, A. Ma, L. Kretzner, F.W. Alt, R.N. Eisenman, and H. Weintraub. 1993. Binding of Myc proteins to canonical and noncanonical DNA sequences. *Mol. Cell. Biol.* **13**: 5216–5224.
- Blackwood, E.M. and R.N. Eisenman. 1991. Max: A helix-loop-helix zipper protein that forms a sequence-specific DNA-binding complex with Myc. *Science* **251**: 1211–1217.
- Brennan, T.J. and E.N. Olson. 1990. Myogenin resides in the nucleus and acquires high affinity for a conserved enhancer element on heterodimerization. *Genes & Dev.* **4**: 582–595.
- 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.
- Campos-Ortega, J.A. 1993. Early neurogenesis in *Drosophila melanogaster*. In *The development of Drosophila melanogaster* (ed. C.M. Bate and A. Martinez-Arias), pp. 109–1129. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York.
- Campuzano, S. and J. Modolell. 1992. Patterning of the *Drosophila* nervous system: The achaete-scute gene complex. *Trends Genet.* **8**: 202–208.
- Campuzano, S., L. Carramolino, C.V. Cabrera, M. Ruiz-Gomez, R. Villares, A. Boronat, and J. Modolell. 1985. Molecular genetics of the achaete-scute gene complex of *D. melanogaster*. *Cell* **40**: 327–338.
- Carr, C.S. and P.A. Sharp. 1990. A helix-loop-helix protein related to the immunoglobulin E box-binding proteins. *Mol. Cell Biol.* **10**: 4384–4388.
- Caudy, M., H. Vassin, M. Brand, R. Tuma, L.Y. Jan, and Y.N. Jan. 1988. daughterless, a *Drosophila* gene essential for both neurogenesis and sex determination, has sequence similarities to myc and the achaete-scute complex. *Cell* **55**: 1061–1067.
- Cowell, I.G. 1994. Repression versus activation in the control of gene transcription. *Trends Biochem. Sci.* **19**: 38–42.
- 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.
- Dang, C.V., C. Dolde, M.L. Gillison, and G.J. Kato. 1992. Discrimination between related DNA sites by a single amino acid residue of Myc-related basic-helix-loop-helix proteins. *Proc. Natl. Acad. Sci.* **89**: 599–602.
- Davis, R.L., H. Weintraub, and A.B. Lassar. 1987. Expression of a single transfected cDNA converts fibroblasts to myoblasts. *Cell* **51**: 987–1000.
- Delidakis, C. and T.S. Artavanis-Tsakonas. 1992. The Enhancer of split [E(spl)] locus of *Drosophila* encodes seven independent helix-loop-helix proteins. *Proc. Natl. Acad. Sci.* **89**: 8731–8735.
- DePinho, M.L., K.S. Hatton, A. Tesfaye, G.D. Yancopoulos, and F.W. Alt. 1987. The human myc gene family: Structure and activity of L-myc and an L-myc pseudogene. *Genes & Dev.* **1**: 1311–1326.
- Ferre-D'Amare, A.R., G.C. Prendergast, E.B. Ziff, and S.K. Burley. 1993. Recognition by Max of its cognate DNA through a dimeric b/HLH/Z domain. *Nature* **363**: 38–45.
- Ghysen, A. and C. Dambly-Chaudiere. 1989. Genesis of the *Drosophila* peripheral nervous system. *Trends Genet.* **5**: 251–255.
- Gonzalez, F., S. Romani, P. Cubas, J. Modolell, and S. Campuzano. 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.
- Gregor, P.D., M. Sawadogo, and R.G. Roeder. 1990. The adenovirus major late transcription factor USF is a member of the helix-loop-helix group of regulatory proteins and binds to DNA as a dimer. *Genes & Dev.* **4**: 1730–1740.
- Guillemot, F., L.C. Lo, J.E. Johnson, A. Auerbach, D.J. Anderson, and A.L. Joyner. 1993. Mammalian achaete-scute homolog 1 is required for the early development of olfactory and autonomic neurons. *Cell* **75**: 463–476.
- Han, K., M.S. Levine, and J.L. Manley. 1989. Synergistic activation and repression of transcription by *Drosophila* homeobox proteins. *Cell* **56**: 573–583.
- Hu, Y.F., B. Luscher, A. Admon, N. Mermod, and R. Tjian. 1990. Transcription factor AP-4 contains multiple dimerization domains that regulate dimer specificity. *Genes & Dev.* **4**: 1741–1752.
- Ingham, P.W., S.M. Pinchin, K.R. Howard, and D. Ish-Horowicz. 1985. Genetics analysis of the hairy locus in *Drosophila melanogaster*. *Genetics* **111**: 463–486.
- Jan, Y.N. and L.Y. Jan. 1993a. The peripheral nervous system. In *The development of Drosophila melanogaster* (ed. C.M. Bate and A. Martinez-Arias), pp. 1207–1244. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York.
- . 1993b. HLH proteins, fly neurogenesis, and vertebrate myogenesis. *Cell* **75**: 827–830.
- Jaynes, J.B. and P.H. O'Farrell. 1991. Active repression of transcription by the Engrailed homeodomain protein. *EMBO J.* **10**: 1427–1433.
- Jones, N. 1990. Transcriptional regulation by dimerization: Two sides to an incestuous relationship. *Cell* **61**: 9–11.
- Klamt, C., E. Knust, K. Tietze, and J.A. Campos-Ortega. 1989. Closely related transcripts encoded by the neurogenic gene complex enhancer of split of *Drosophila melanogaster*. *EMBO J.* **8**: 203–210.
- Knust, E., H. Schrons, F. Grawe, and J.A. Campos-Ortega. 1992. Seven genes of the Enhancer of split complex of *Drosophila melanogaster* encode helix-loop-helix proteins. *Genetics* **132**: 505–518.
- Kohl, N.E., E. Legouy, R.A. DePinho, P.D. Nisen, R.K. Smith, C.E. Gee, and F.W. Alt. 1986. Human N-myc is closely re-

- lated in organization and nucleotide sequence to c-myc. *Nature* **319**: 73–77.
- Laski, F.A., D.C. Rio, and G.M. Rubin. 1986. Tissue specificity of *Drosophila* P element transposition is regulated at the level of mRNA splicing. *Cell* **44**: 7–19.
- Lassar, A.B., R.L. Davis, W.E. Wright, T. Kadesch, C. Murre, A. Voronova, D. Baltimore, and H. Weintraub. 1991. Functional activity of myogenic HLH proteins requires heterooligomerization with E12/E47-like proteins in vivo. *Cell* **66**: 305–315.
- Lindsley, D.L. and G.G. Zimm. 1992. *The genome of Drosophila melanogaster*. Academic Press, San Diego, CA.
- Marcu, K.B., S.A. Bossone, and A.J. Patel. 1992. myc function and regulation. *Annu. Rev. Biochem.* **61**: 809–860.
- Martinez, C., J. Modolell, and J. Garrell. 1993. Regulation of the proneural gene *achaete* by helix-loop-helix proteins. *Mol. Cell Biol.* **13**: 3514–3521.
- Moscoso del Prado, J. and A. Garcia-Bellido. 1984. Genetic regulation of the *achaete-scute* complex of *Drosophila melanogaster*. *Wilhelm Roux's Arch. Dev. Biol.* **193**: 242–245.
- Murre, C., P.S. McCaw, and D. Baltimore. 1989a. A new DNA binding and dimerization motif in immunoglobulin enhancer binding, daughterless, MyoD, and myc proteins. *Cell* **56**: 777–783.
- 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. 1989b. Interactions between heterologous helix-loop-helix proteins generate complexes that bind specifically to a common DNA sequence. *Cell* **58**: 537–544.
- Parkhurst, S.M., D. Bopp, and D. Ish-Horowicz. 1990. X:A ratio, the primary sex-determining signal in *Drosophila*, is transduced by helix-loop-helix proteins [erratum appears in *Cell* 1991, **64**(5): following p. 1046]. *Cell* **63**: 1179–1191.
- Prendergast, G.C., D. Lawe, and E.B. Ziff. 1991. Association of Myn, the murine homolog of max, with c-Myc stimulates methylation-sensitive DNA binding and ras cotransformation. *Cell* **65**: 395–407.
- Romani, S., S. Campuzano, E.R. 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.
- Rubin, G.M. and A.C. Spradling. 1982. Genetic transformation of *Drosophila* with transposable element vectors. *Science* **218**: 348–353.
- Rushlow, C.A., A. Hogan, S.M. Pinchin, K.M. Howe, M. Lardelli, and D. Ish-Horowicz. 1989. The *Drosophila* hairy protein acts in both segmentation and bristle patterning and shows homology to N-myc. *EMBO J.* **8**: 3095–3103.
- Sasai, Y., R. Kageyama, Y. Tagawa, R. Shigemoto, and S. Nakanishi. 1992. Two mammalian helix-loop-helix factors structurally related to *Drosophila* hairy and *Enhancer of split*. *Genes & Dev.* **6**: 2620–2634.
- 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.
- Spradling, A.C. and G.M. Rubin. 1983. The effect of chromosomal position on the expression of the *Drosophila* xanthine dehydrogenase gene. *Cell* **34**: 47–57.
- Sun, X.H. and D. Baltimore. 1991. An inhibitory domain of E12 transcription factor prevents DNA binding in E12 homodimers but not in E12 heterodimers. *Cell* **64**: 459–470.
- Thummel, C.S., A.M. Boulet, and H.D. Lipshitz. 1988. Vectors for *Drosophila* P-element-mediated transformation and tissue culture transfection. *Gene* **74**: 445–456.
- Tietze, K., N. Oellers, and E. Knust. 1992. Enhancer of splitD, a dominant mutation of *Drosophila*, and its use in the study of functional domains of a helix-loop-helix protein. *Proc. Natl. Acad. Sci.* **89**: 6152–6156.
- Van Antwerp, M., D.G. Chen, C. Chang, and E.V. Prochownik. 1992. A point mutation in the MyoD basic domain imparts c-Myc-like properties. *Proc. Natl. Acad. Sci.* **89**: 9010–9014.
- 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.
- 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.
- Van Doren, M., A.M. Bailey, J. Esnayra, K. Ede, and J.W. Posakony. 1994. Negative regulation of proneural gene activity: hairy is a direct transcriptional repressor of *achaete*. *Genes & Dev.* (this issue).
- 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.
- Wainwright, S.M. and D. Ish-Horowicz. 1992. Point mutations in the *Drosophila* hairy gene demonstrate in vivo requirements for basic, helix-loop-helix, and WRPW domains. *Mol. Cell Biol.* **12**: 2475–2483.
- Weintraub, H., R. Davis, S. Tapscott, M. Thayer, M. Krause, R. Benezra, T.K. Blackwell, D. Turner, R. Rupp, S. Hollenberg, Y. Zhuang, and A. Lassar. 1991. The myoD gene family: Nodal point during specification of the muscle cell lineage. *Science* **251**: 761–766.

Hairy function as a DNA-binding helix-loop-helix repressor of *Drosophila* sensory organ formation.

S Ohsako, J Hyer, G Panganiban, et al.

Genes Dev. 1994, **8**:

Access the most recent version at doi:[10.1101/gad.8.22.2743](https://doi.org/10.1101/gad.8.22.2743)

References

This article cites 59 articles, 26 of which can be accessed free at:
<http://genesdev.cshlp.org/content/8/22/2743.full.html#ref-list-1>

License

Email Alerting Service

Receive free email alerts when new articles cite this article - sign up in the box at the top right corner of the article or [click here](#).