Development 124, 4039-4051 (1997)
Printed in Great Britain © The Company of Biologists Limited 1997
DEV5144

The *Drosophila* gene *Bearded* encodes a novel small protein and shares 3′ UTR sequence motifs with multiple *Enhancer of split* Complex genes

Michael W. Leviten†, Eric C. Lai and James W. Posakony*

Department of Biology and Center for Molecular Genetics, University of California San Diego, 9500 Gilman Drive, La Jolla, CA 92093-0349, USA

†Present address: Deltagen Inc., 1031 Bing Street, San Carlos, CA 94070, USA
*Author for correspondence (e-mail: jposakony@ucsd.edu)

SUMMARY

Gain-of-function alleles of the *Drosophila* gene *Bearded* (*Brd*) cause sensory organ multiplication and loss phenotypes indistinguishable at the cellular level from those caused by loss-of-function mutations in the genes of the *Notch* pathway (Leviten, M. W. and Posakony, J. W. (1996). *Dev. Biol.* 176, 264-283). We have carried out a molecular analysis of the structure and expression of both wild-type and mutant *Brd* transcription units. We find that the *Brd* transcript is truncated and accumulates to substantially higher levels in the gain-of-function mutants, due to the insertion of a transposable element of the *blood* family in the *Brd* 3′ untranslated region (UTR). The wild-type *Brd* 3′ UTR includes three copies of a 9-nucleotide sequence (CAGCTTTAA) that we refer to as the 'Brd box'. Moreover, the 3′ UTRs of *Brd* and of the *m4* transcription unit of the *Enhancer of split* gene complex [E(spl)-C] exhibit an unusually high degree of sequence identity that includes not only Brd box sequences but also a second motif we refer to as the 'GY box' (GTCTTCC). We find that both the Brd box and the GY box are also present in the 3′ UTRs of several basic helix-loop-helix repressor-encoding genes of the E(spl)-C, often in multiple copies, suggesting that a novel mode of post-transcriptional regulation applies to *Brd* and many E(spl)-C genes. The fact that the more abundant *Brd* mutant mRNA lacks the GY box and two of the Brd boxes present in wild-type *Brd* mRNA suggests that either or both of these elements may confer instability on transcripts that contain them. Finally, we find that *Brd* encodes a novel small protein of only 81 amino acids that is predicted to include a basic amphipathic α-helix. The deduced Brd protein shows sequence similarity to the E(spl)m4 protein, which is likewise expected to include a basic amphipathic α-helix, suggesting that the two proteins have related biochemical functions.

Key words: neurogenesis, sensory organ development, *Enhancer of split* Complex, *Notch* signalling, basic amphipathic α-helix, transposable element, 3′ untranslated region, post-transcriptional regulation

INTRODUCTION

Distributed in a stereotyped pattern on the body surface of adult *Drosophila* is a diverse assortment of external sensory organs that together comprise most of the peripheral nervous system (PNS) of the fly. Each organ is composed of a specific set of differentiated cell types, which includes one or more sensory neurons and a small number of non-neuronal accessory cells that produce the characteristic stimulus-receiving structures of the organ. For example, the mechanosensory bristles that occur in large numbers over most of the cuticular surface each contain a single bipolar neuron and three other cells (the thecogen, trichogen and tormogen) that construct, respectively, a cap over the dendrite of the neuron, the bristle shaft, and the socket around the base of the shaft. The four cells that comprise each bristle derive, via a fixed lineage, from a single sensory organ precursor (SOP) cell (Hartenstein and Posakony, 1989).

The development of these sensory bristles commences during the late larval and early pupal stages within undifferentiated ectodermal sheets (the imaginal discs and histoblast nests) that will ultimately give rise to the cuticle of the adult fly. First, small groups of cells called proneural clusters express two proneural genes, *achaete* (*ac*) and *scute* (*sc*), which encode transcriptional activators of the basic helix-loop-helix (bHLH) class (Alonso and Cabrera, 1988; Cabrera and Alonso, 1991; Cubas et al., 1991; Skeath and Carroll, 1991; Van Doren et al., 1992). The activity of the ac and sc proteins confers on the cells of the proneural cluster the potential to adopt the SOP cell fate. However, inhibitory cell-cell interactions within each cluster, mediated by the neurogenic genes, then restrict the expression of the SOP fate to a single cell, leaving the inhibited cells to adopt the ordinary epidermal fate (Dietrich and Campos-Ortega, 1984; Hartenstein and Posakony, 1990; Parks and Muskavitch, 1993; Schweisguth, 1995; Schweisguth and Posakony, 1992, 1994; Tata and Hartley, 1995). Once the SOP cell has been stably specified, the four constituent cells of the bristle organ are generated by three successive asymmetric cell divisions (Hartenstein and Posakony, 1989). Cell-cell interac-

tions requiring neurogenic gene function are required for the asymmetry of daughter cell fates at each division (Posakony, 1994).

Several genes of the neurogenic group encode proteins that are known to function as part of the *Notch* cell-cell signalling and intracellular transduction pathway. Two transmembrane proteins, the products of *Notch* (*N*) and *Delta* (*Dl*), act as receptor and ligand, respectively, in this system (Fehon et al., 1990; Heitzler and Simpson, 1991; Struhl et al., 1993). Interaction of their extracellular domains results in the activation of Suppressor of Hairless [Su(H)], a sequence-specific DNA-binding protein, by means of a direct protein-protein interaction with the intracellular domain of the N receptor (Fortini and Artavanis-Tsakonas, 1994; Hsieh et al., 1996; Jarriault et al., 1995; Tamura et al., 1995). Su(H) then transduces the signal in part by functioning as a direct transcriptional activator of most of the genes of the *Enhancer of split* Complex [E(spl)-C] (Bailey and Posakony, 1995; Furukawa et al., 1995; Lecourtois and Schweisguth, 1995). This complex includes seven transcription units that encode bHLH transcriptional repressor proteins, as well as *E(spl)m4*, the product of which is a novel 152-amino acid (aa) protein of unknown function (Delidakis and Artavanis-Tsakonas, 1992; Klämbt et al., 1989; Knust et al., 1992).

A number of other known genes exhibit loss-of-function phenotypes similar to those of *N* pathway genes, but at present the specific roles of their protein products remain unknown. These include *mastermind* (*mam*) (Smoller et al., 1990), *neuralized* (*neu*) (Price et al., 1993) and *kuzbanian* (*kuz*) (Rooke et al., 1996). Recently we reported that gain-of-function mutations at a new locus, *Bearded* (*Brd*), cause phenotypic effects on adult sensory organ development that closely mimic those resulting from loss-of-function mutations in *N* pathway genes (Leviten and Posakony, 1996). These include the commitment of supernumerary proneural cluster cells to the SOP fate, and the appearance of additional sheath cell/neuron pairs at the expense of the socket and shaft cells in the sensory organ lineage. Moreover, we found that these dominant phenotypic effects of *Brd* gain-of-function alleles are sensitive to the dosage of two neurogenic genes, *N* and *neu*, and of *Hairless* (*H*), a negative regulator of *N* pathway activity (Bailey and Posakony, 1995; Bang et al., 1995; Bang and Posakony, 1992; Brou et al., 1994; Schweisguth and Posakony, 1994; Vässin et al., 1985). In an earlier study, we showed that both *Brd* and the genes of the E(spl)-C are directly activated in the proneural clusters of the wing imaginal disc by protein complexes that include ac and sc, and that the same genes are expressed in the vicinity of the morphogenetic furrow of the eye imaginal disc under *ac/sc*-independent control (Bailey and Posakony, 1995; Singson et al., 1994).

Here we present a molecular analysis of the structure and expression of both wild-type and mutant *Brd* genes. We have determined that *Brd* gain-of-function alleles are associated with an insertion of a retrotransposon of the *blood* family into the 3′ untranslated region (UTR) of the gene. We find that the wild-type *Brd* gene encodes a novel small protein of only 81 aa that is predicted to include a highly basic amphipathic α-helix. We show that the 3′ UTRs of *Brd* and *E(spl)m4* exhibit an unusually high degree of sequence identity, that the Brd and E(spl)m4 proteins are also related, and that the E(spl)m4 protein is likewise expected to form a basic amphipathic α-helix, suggesting that the two proteins have related biochemical activities. Finally, we report that *Brd* shares with several transcription units of the E(spl)-C, including *E(spl)m4*, two novel heptanucleotide sequence motifs in their 3′ UTRs, suggesting that these genes are subject to a previously unrecognized mode of post-transcriptional regulation. These observations strengthen the functional association between *Brd* and the *N* signalling pathway.

MATERIALS AND METHODS

Fly strains

Flies were grown on standard yeast-cornmeal-molasses-agar medium. Canton S flies were utilized as the wild-type stock for in situ hybridization experiments. *Brd* alleles and additional stocks are described in Leviten and Posakony (1996).

In situ hybridization

In situ hybridization to larval imaginal discs and pupal nota was performed using digoxygenin-UTP-labeled RNA probes essentially as described by Tautz and Pfeifle (1989), with the modifications of Jiang et al. (1991) and Schweisguth and Posakony (1992). Double-labeling experiments involving the enhancer trap marker A101 were performed as described previously (Romani et al., 1989).

Isolation of *Brd* cDNA clones

Brd cDNA clones were isolated from a 4-8 hour embryonic cDNA library carried in the pNB40 plasmid vector (Brown and Kafatos, 1988). Library DNA was transformed into Epicurean SCS1 supercompetent cells (Stratagene) as described by the manufacturer, and colony blotting and hybridization was performed as described (Maniatis et al., 1982). Five independent cDNA clones were isolated from approximately 1×10^4 colonies screened.

Purification of genomic DNA

Drosophila genomic DNA used in Southern blot analysis and in PCR reactions was isolated as described in Ellis et al. (1990).

Construction of a genomic DNA mini-library

To isolate a P element-tagged DNA fragment from the Brd^{P1} chromosome, genomic DNA from this strain was digested with *Eco*RI and electrophoresed on a 0.8% agarose gel. Fragments between 6.0 and 9.0 kb were gel purified using Geneclean (Bio101) and ligated to *Eco*RI-digested and dephosphorylated λZAP phagemid vector (Stratagene). Phage were plated and DNA transferred to nylon filters (Gene Screen) according to the Stratagene protocol. Filters were then probed with cloned P element sequences. Positive clones were subjected to plasmid rescue according to the λZAP protocol and characterized by restriction mapping and Southern blot analysis.

Isolation of wild-type genomic DNA clones

Two phage clones carrying wild-type genomic DNA from the *Brd* locus were isolated from two separate libraries by standard screening techniques (Maniatis et al., 1982). λBrd1 was isolated from the EMBL3 'D' library obtained from M. Muskavitch. This phage contains approx. 15 kb of genomic DNA extending from just 5′ of the *Brd* transcription unit to nearly 14 kb downstream. λBrd2 contains approx. 13 kb of genomic DNA, including approx. 8 kb upstream of the *Brd* transcription unit. This phage was isolated from a λFixII library constructed by Stratagene.

Isolation of RNA

Poly(A)$^+$ RNA preparations from carefully staged embryos were the same as those described by Ellis et al. (1990) and Margolis et al.

(1994). RNA from larvae and pupae was prepared as described by Ellis et al. (1990). Poly(A)$^+$ RNA was isolated using the polyATract kit (Promega) according to the manufacturer's instructions.

PCR amplification of DNA from *Brd* point mutants

Two oligonucleotide primers were used to amplify by PCR a product corresponding to bp 1-400 of the *Brd* cDNA sequence from genomic DNA prepared from various *Brd* mutant stocks. The genomic DNA was purified from homozygous mutant adults when possible, and in other cases from flies transheterozygous for a particular mutant allele and the Brd^{Prv1} chromosome, which is completely deleted for the *Brd* gene. The 5′ primer corresponds to bp 1-18 and has an attached *Eco*RI site at its 5′ end; the 3′ primer extends from bp 383-400, with a synthetic *Hin*dIII site at its 5′ end. Amplified fragments were cloned into pBluescript KS (Stratagene), and sequenced using the Sequenase 2.0 kit (USB). For each mutant chromosome analyzed, at least two clones from independent amplification reactions were isolated and sequenced.

PCR amplification of wild-type and mutant RNA sequences

Total RNA was isolated from Brd^1 and w^{1118} late third-instar larvae as described above. First-strand cDNA was synthesized from 50-100 μg of total RNA using the oligonucleotide JK1 (C_{13}AAGCTTT$_{15}$; provided by J. Kavaler) and first-strand synthesis components and protocol from the Riboclone cDNA Synthesis System (Promega). A portion of this cDNA product was PCR amplified using a 5′ primer corresponding to *Brd* coding sequence (bp 149-171 of *Brd* cDNA; AAGGCCATGAAGAAGAAGATCTTCAA) and a 3′ primer JK2 (ATCAGGTGACGTCGAC$_{13}$; provided by J. Kavaler). A small fraction of this first amplification reaction was then reamplified using JK2 and a second *Brd* primer (bp 206-223; AGAAGATCTC-AGCAACTGAAGAACCAC) positioned downstream in the *Brd* cDNA sequence from the first *Brd* primer.

Germline transformation

A 550-bp *Sal*I restriction fragment, extending from the 5′ UTR of *Brd* to a site approximately 180 bp within the *blood* element inserted in the *Brd* 3′ UTR (see Fig. 1), was filled in using Klenow polymerase, and linkered with phosphorylated *Xba*I linkers. The resulting *Xba*I fragment was cloned into the unique *Xba*I site of the CaSpeR-*Hsp70* P element transformation vector (Bang and Posakony, 1992). DNA of the resulting P[w^+, *Hs-Brd*] transposon clone was injected into w^{1118} recipient embryos in combination with a Δ2-3 helper plasmid to provide transposase activity. Six independent transformant lines were established.

Heat-shock treatments

Crawling third-instar larvae and 0-, 6- and 16-hour APF pupae were placed in plastic Petri dishes containing a moistened paper towel and exposed to a 4-hour heat-shock regimen (2× 60 minutes in a 37°C air incubator, separated by 120 minutes at 25°C).

Adult cuticle preparations

Adult flies were anesthetized and dissected in PBS. Dissected cuticle was placed in 1:1 (v/v) lactic acid:Hoyer's medium mixture and incubated in a 55°C oven for 60 minutes to dissolve any associated tissue. The cuticle was then mounted on a slide, covered with a drop of Hoyer's, and incubated at 55°C to harden.

RESULTS

Molecular cloning of the *Brd* locus

The Brd^{P1} allele was recovered in a P element mutagenesis screen designed to obtain phenotypic revertants of the original dominant Brd^1 allele, and displays a partial suppression of the Brd^1 bristle multiplication phenotype (Leviten and Posakony, 1996). Southern blot analysis of genomic DNA and in situ hybridization to salivary gland polytene chromosomes (data not shown) indicated that the Brd^{P1} chromosome carries a single P element insertion at cytological position 71A1-2, the known site of the *Brd* locus (Leviten and Posakony, 1996). Exposure of the Brd^{P1} chromosome to P transposase activity (Robertson et al., 1988) yielded a high frequency of derivative chromosomes that generate various modified *Brd* phenotypes, indicating that the partial suppression of the Brd^1 phenotype exhibited by Brd^{P1} is most likely due to a P transposon insertion near the gene.

From a Brd^{P1} genomic DNA mini-library, we isolated an *Eco*RI fragment of 7.0 kilobases (kb) that was found to contain a 1.5-kb defective P element (Fig. 1). Southern blot analysis showed that the P element is flanked in this fragment by 1.5 kb and 0.5 kb of unique genomic sequences and that the remainder of the clone consists of additional, non-P-element repetitive DNA.

Subsequent Southern blot and DNA sequence analysis revealed the following features of the cloned genomic region

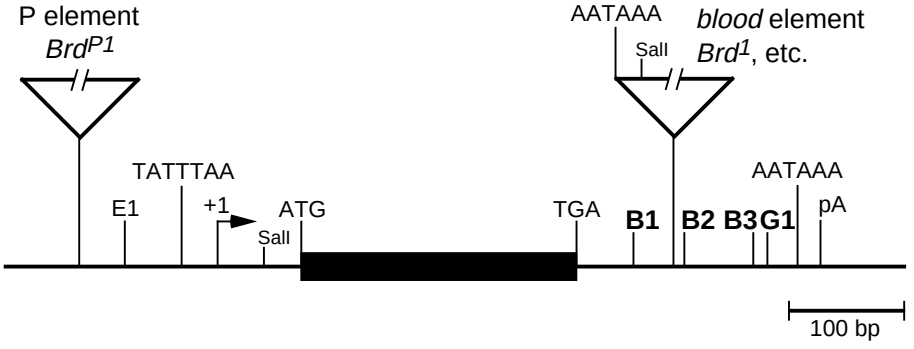

Fig. 1. Structure of wild-type and mutant *Brd* transcription units (see also Fig. 4). Filled box represents the *Brd* protein coding region, with start and stop codons indicated. The transcription start site (+1) is indicated by a horizontal arrow. Upstream elements shown are the E1 'E box' sequence that mediates transcriptional activation of *Brd* by proneural proteins (Singson et al., 1994), and a TATA box consensus sequence located at –30. Downstream elements include a polyadenylation signal (AATAAA) and the polyadenylation site (pA), as determined from cDNA sequences. There are no introns in the *Brd* gene. Bold letters mark the positions of 'Brd box' and 'GY box' motifs in the *Brd* 3′ UTR (see also Figs 4 and 7). Dominant gain-of-function alleles of *Brd* (Brd^1 and derivatives; see Leviten and Posakony, 1996) are associated with an insertion into the *Brd* 3′ UTR of a transposable element of the *blood* family; the proximal LTR of this transposon includes a polyadenylation signal, as shown. The partial revertant allele Brd^{P1} is associated with the additional insertion of a 1.5-kb P transposable element in the proximal upstream region of the gene. The two *Sal*I restriction sites shown mark the termini of the mutant genomic DNA fragment used in the construction of a P[*Hs-Brd*] transposon (see Materials and methods).

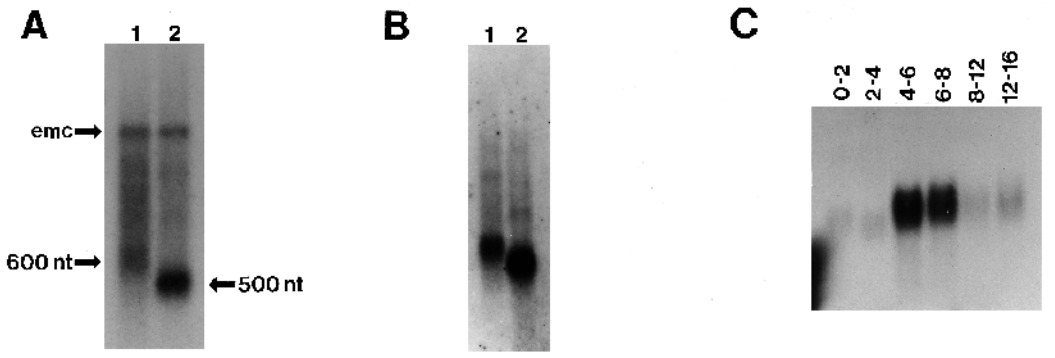

Fig. 2. Northern blot analysis of wild-type and *Brd* mutant transcripts. (A) Wild-type (lane 1) and Brd^3 mutant (lane 2) poly(A)$^+$ RNA isolated from 0- to 24-hour APF pupae was probed with genomic DNA from the *Brd* transcription unit, as well as with a cDNA clone representing the *extramacrochaetae* (*emc*) gene (Ellis et al., 1990) as a loading control; the *emc* transcript is approximately 2.3 kb in length. The *Brd* mutant transcript is approximately 100 bases shorter than the wild-type transcript and several-fold more abundant. (B) A longer exposure of a similar blot more clearly reveals the wild-type *Brd* transcript (lane 1); Brd^3 mutant RNA is in lane 2. (C) Temporal pattern of *Brd* transcript accumulation in poly(A)$^+$ RNA of wild-type embryos. Staging is indicated in hours after egg-laying.

in wild-type and *Brd* mutant DNAs (Fig. 1). Firstly, all of the chromosomes derived by mobilization of the 1.5-kb P element in Brd^{P1} display alterations in restriction fragment sizes in this region. Secondly, as expected, the parental Brd^1 chromosome from which Brd^{P1} was derived lacks the P element entirely. Finally, Brd^1 and most of its derivatives do contain the non-P-element repetitive DNA, but these sequences are not present in this region in wild-type chromosomes. As described in more detail below, we have found that this repetitive DNA corresponds to a transposable element of the *blood* family. We suggest that the insertion of this element is the direct cause of the original Brd^1 mutation.

Identification of wild-type and mutant *Brd* transcripts

The two regions of single-copy genomic DNA in the 7.0-kb Brd^{P1} clone were isolated and used separately as probes on northern blots containing poly(A)$^+$ mRNA from both embryos and pupae. The 0.5-kb fragment that lies between the P element and the non-P repetitive DNA (Fig. 1) detects a single transcript of about 600 nucleotides (nt) at both developmental stages in wild-type RNA (Fig. 2), while the transcript in the *Brd* mutants is both shorter (approximately 500 nt) and strikingly more abundant (Fig. 2A,B). This is consistent with the gain-of-function nature of the original *Brd* mutant alleles, and more specifically with the hypothesis that these mutations are hypermorphic (Leviten and Posakony, 1996).

Fig. 2C shows the temporal pattern of *Brd* transcript accumulation in poly(A)$^+$ RNA of wild-type embryos. We observe only a very low level of *Brd* RNA in preblastoderm embryos, including a small amount of presumably maternal transcript at 0-2 hours. In 4- to 6- and 6- to 8-hour embryos, the abundance of the transcript is sharply higher; interestingly, this is the period during which neuroblast segregation, a *N*-regulated process, takes place. By the next time interval (8-12 hours), *Brd* transcript levels have declined sharply, and remain low throughout the rest of embryogenesis. This finding indicates that, at least in the embryo, wild-type *Brd* RNA is likely to be short-lived. Overall, it is clear that embryonic levels of *Brd* transcript are subject to strong temporal modulation.

Brd is expressed in a proneural cluster pattern in imaginal discs

Several cDNA clones (see below) representing the wild-type *Brd* mRNA were isolated from a 4- to 8-hour embryo library, and used to analyze the spatial pattern of *Brd* transcript accumulation. Whole-mount in situ hybridization to imaginal discs from late third-instar larvae reveals that the *Brd* transcript accumulates specifically in a proneural cluster pattern in both wild-type and *Brd* mutant imaginal discs (Fig. 3A,B; see also Singson et al., 1994). Consistent with the northern blot data (see Fig. 2), *Brd* RNA is significantly more abundant in the mutant discs, as reflected by a much stronger in situ hybridization signal (Fig. 3B). These observations provide further support for the interpretation that the gain-of-function alleles of *Brd* are hypermorphs (Leviten and Posakony, 1996), since *Brd* is apparently overexpressed, and not spatially misexpressed, in these mutants. In addition, the finding that *Brd* is normally expressed in wild-type proneural clusters is consistent with the hypothesis that the gene has a wild-type function in adult sensory organ development (see Leviten and Posakony, 1996).

In the wing imaginal disc, *Brd* transcript is detectable at the positions of all developing adult external sensory organs, analogous to the expression of the proneural genes *ac* and *sc*. We have shown previously that *Brd* bristle phenotypes are dependent on *ac/sc* function (Leviten and Posakony, 1996), that *Brd* transcript accumulation within wing disc proneural clusters likewise requires *ac/sc* activity (Singson et al., 1994), and that the *Brd* promoter is directly activated in proneural clusters by bHLH protein complexes that include ac and sc (Singson et al., 1994). However, in contrast to *ac* and *sc*, there is generally no clear elevation of *Brd* transcript in presumptive SOPs.

To examine the spatial and temporal relationship of *Brd* expression to SOP specification, we performed double-labeling experiments with the SOP-specific enhancer trap marker A101. The results presented here are documented in *Brd* mutant discs, in which the in situ hybridization signal is stronger and more clusters are clearly represented than in wild-type discs.

Double-labeled third-instar wing discs demonstrate clearly that *Brd* transcript is localized to sites of SOP development

(Fig. 3C). These experiments also reveal that *Brd* expression, like *ac/sc* expression, precedes SOP specification as marked by A101. For example, in the anterior scutellar (aSC) macrochaete proneural cluster, uniform *Brd* expression can clearly be seen prior to detectable A101 staining (Fig. 3D). Subsequently, a single A101-positive cell is observed among the *Brd*-positive cells (Fig. 3E). In many clusters there is clearly a period in which the multiplied SOPs express both *Brd* and A101. This is most evident in regions in which a cross-sectional view of the disc epithelium can be seen, and the A101-positive nucleus of an SOP cell is clearly surrounded by cytoplasmic *Brd* transcript (Fig. 3F). Finally, *Brd* expression in A101-positive cells appears to become diminished at some positions in late third-instar and early pupal discs. For example, in the A101-positive SOP cells at the posterior scutellar (pSC) macrochaete position, *Brd* expression is much weaker than in the aSC region at the same time (Fig. 3E). Other positions, including the anterior postalar (aPA) macrochaete site, also exhibit undetectable levels of *Brd* transcript while expressing A101 strongly (data not shown). The pSC and aPA are two of the first macrochaete SOPs to be determined (Huang et al., 1991), suggesting that the decreased *Brd* RNA levels observed there represent the turnover of transcript and not simply a relatively lower level of *Brd* expression throughout the development of these particular clusters.

In summary, these results suggest a general pattern for the progression of *Brd* expression within imaginal disc proneural clusters. Initially, *Brd* transcript is present at roughly equal levels throughout the cluster, and this expression precedes that of the early SOP marker A101. Subsequently, *Brd* and A101 are coexpressed within the SOP cell (with *Brd* transcript remaining in the non-SOP cells as well), and then *Brd* transcript levels become diminished in both the SOPs and the surrounding cluster cells.

Characterization of the *Brd* transcription unit

We began the detailed characterization of wild-type and mutant *Brd* transcription units by isolating five independent cDNA clones from a plasmid library representing poly(A)$^+$ RNA of wild-type 4- to 8-hour embryos. All contained inserts of approximately 550 bp, consistent with the size of the normal *Brd* transcript detected on northern blots (see Fig. 2). In addition to their identical sizes, two other criteria suggest that these cDNAs are full-length: two of the clones were found to start at the same nucleotide, and analysis of upstream genomic DNA sequence reveals a consensus TATA box (TATTTAA) 30 bp upstream of the first base of the cDNAs (Fig. 4) (Singson et al., 1994). In addition, the 3′ ends of the two clones analyzed are identical in sequence, each containing a single consensus polyadenylation signal (AATAAA) 18 bp 5′ of the poly(A) tail, suggesting that the complete 3′ end of the *Brd* transcript is also represented within these cDNAs (Fig. 4).

One cDNA clone was sequenced in its entirety, revealing the presence of a single major open reading frame extending from the first ATG triplet at position +71 to a TGA stop codon at +314, flanked by 70 bp of 5′ untranslated region (UTR) and 233 bp of 3′ UTR (Fig. 4). The sequence of the entire *Brd* transcription unit from a wild-type genomic DNA clone (data not shown) showed that the genomic and cDNA sequences are completely colinear, indicating that there are no introns in the gene.

The exact structure of the *Brd* mutant transcription unit was similarly determined by DNA sequence analysis of the Brd^{P1} genomic DNA clone (see above and Fig. 1). The P element insertion in the Brd^{P1} allele lies 125 bp upstream of the putative transcription start site (Fig. 4), suggesting that it could partially suppress the *Brd* phenotype by interfering with the transcriptional regulation of the gene. The non-P repetitive element associated with *Brd* gain-of-function alleles (see above) was found to lie at position +399, in the 3′ UTR of the gene. Upstream of the point of insertion of this element, the mutant

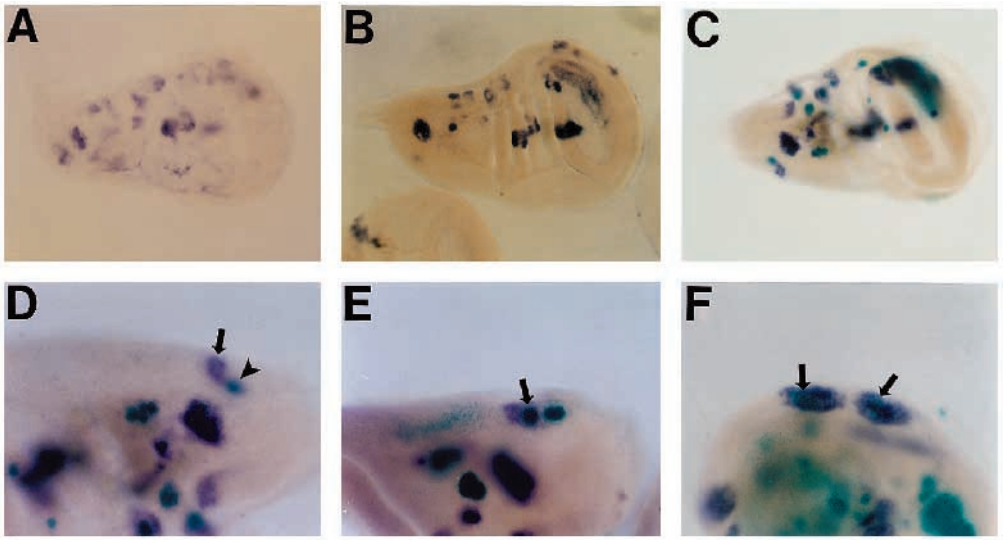

Fig. 3. Spatial and temporal pattern of *Brd* transcript accumulation in wild-type and *Brd* mutant larval imaginal discs. Imaginal discs were hybridized using a digoxygenin-labeled *Brd* RNA probe; expression of the *lacZ* reporter gene in the A101 enhancer trap insertion was detected by β-galactosidase activity staining. (A) Wild-type and (B) Brd^{1} homozygous late third-instar wing imaginal discs exhibit similar proneural cluster patterns of *Brd* expression. (C) Brd^{3} *A101/+* wing disc double-labeled for A101 expression in sensory organ precursors (blue) and *Brd* RNA accumulation (purple). *Brd* transcript clearly surrounds sites of A101 expression. (D,E) Accumulation of *Brd* transcript precedes the appearance of A101-positive SOPs, as illustrated in the region from which the anterior scutellar (aSC) macrochaete arises (arrows). (D) An A101-positive cell is present at the posterior scutellar (pSC) macrochaete position (arrowhead), while at the aSC position (arrow) a cluster of *Brd*-expressing cells lacking A101 expression is observed. (E) An older third-instar disc shows A101-positive cells at both the pSC and aSC (arrow) positions. Note that *Brd* transcript is still present at the aSC position, but no longer appears at the earlier-determined pSC position. (F) Leg imaginal disc showing A101-positive nuclei surrounded by cytoplasm containing *Brd* transcript (arrows).

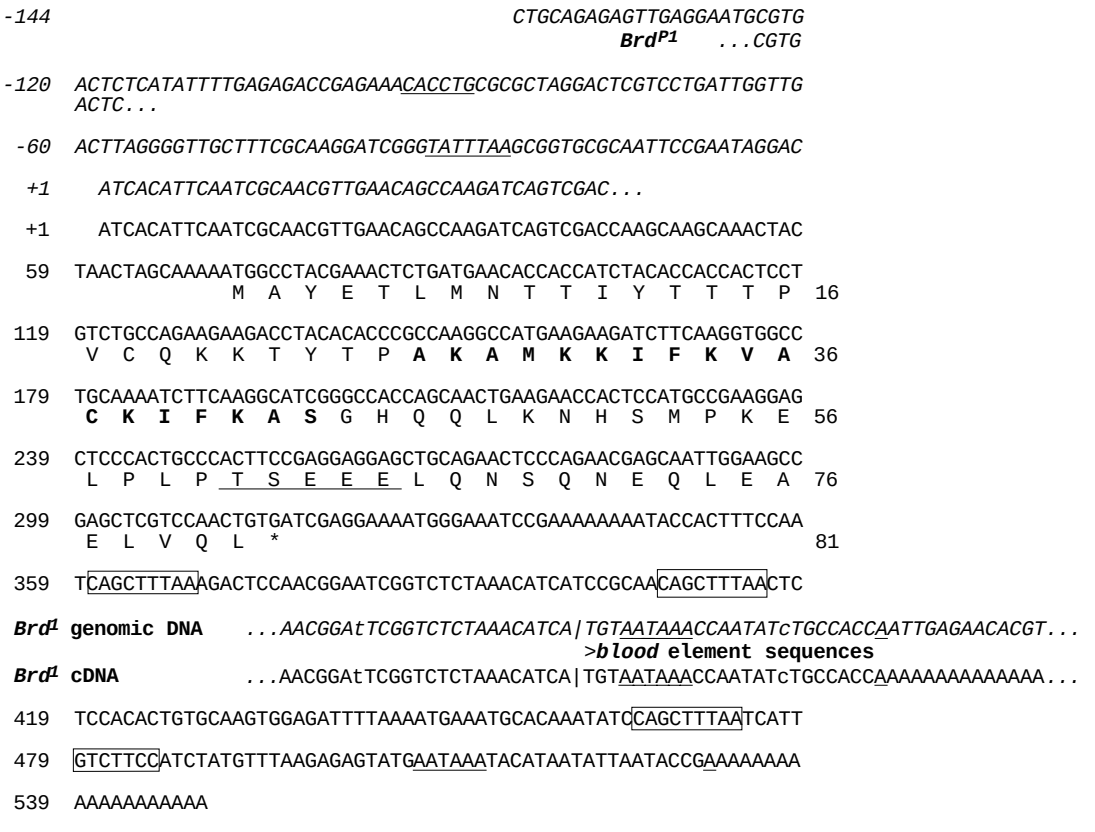

Fig. 4. Sequence of wild-type and mutant *Brd* genomic DNAs and cDNAs. Genomic DNA sequences are in italics, cDNA sequences in Roman type. Nucleotide numbers are shown at left (transcription start site at +1), amino acid numbers at right. The sequence of a full-length, wild-type *Brd* cDNA clone from a 4- to 8- hour embryonic library is shown (1 to 549), along with overlapping 5′ UTR and flanking sequences of wild-type genomic DNA. The duplicated eight bases flanking the upstream P element insertion in Brd^{P1}, as well as 3′ UTR and flanking sequences of Brd^{1} genomic DNA (obtained from Brd^{P1}) and cDNA (from Brd^{1} RNA), are aligned under the corresponding wild-type sequence. LTR sequences of the *blood* transposon insertion in Brd^{1} and other gain-of-function alleles are indicated. Underlined nucleotide sequences are, from top, an 'E box' sequence (CACCTG) that mediates transcriptional activation of *Brd* by proneural proteins (Singson et al., 1994); the TATA box consensus sequence (TATTTAA); consensus polyadenylation signals (AATAAA); and polyadenylation sites (single underlined As). The three 'Brd box' motifs and single 'GY box' element in the *Brd* 3′ UTR are boxed. Lower-case letters represent nucleotide sequence polymorphisms. The *Brd* cDNA sequence includes an open reading frame capable of encoding an 81-aa protein; the conceptual translation is presented in the one-letter code. The *Brd* start codon shown is preceded by an in-frame stop codon (TAA, nt 59-61), and the sequence preceding the start codon (AAAAATG) conforms to the *Drosophila* consensus translation initiation sequence (Cavener, 1987). Two features of the predicted Brd protein are indicated: A putative basic amphipathic α-helix (aa 26-43, bold lettering), and a consensus casein kinase II phosphorylation signal (aa 61-65; underlined) (Kemp and Pearson, 1990). The stop codon is indicated by an asterisk. The GenBank accession number for the *Brd* cDNA sequence is AF016542.

genomic DNA is identical to the wild-type cDNA sequence, with the exception of a 1-bp silent polymorphism in the coding region and another 1-bp polymorphism in the 3′ UTR (Fig. 4). The sequence of the 5′ terminus of this repetitive element (Fig. 4) identified it as a member of the *blood* family of retrotransposons, first recognized in association with the *blood* allele of the *white* locus (w^{bl}) (Bingham and Chapman, 1986). The *blood* element in *Brd* mutants is oriented such that a polyadenylation signal is located just 3 bp inside the LTR proximal to the *Brd* transcription start (Fig. 4). Efficient use of this *blood* polyadenylation signal could account for the truncation of the transcript in the *Brd* mutants (see Fig. 2A,B). To confirm this interpretation, we cloned the 3′ ends of the *Brd* transcript from both wild-type and Brd^{1} larvae using RT-PCR (see Materials and methods), and determined their sequence. We found that the Brd^{1} transcript does include *blood* element LTR sequences at its 3′ terminus, and is polyadenylated 15 nt downstream of the *blood* polyadenylation signal (Fig. 4). This generates a mutant transcript that is approximately 100 nt shorter than the wild-type larval transcript, which itself terminates at exactly the same position as indicated by the embryonic cDNA clones (Fig. 4). These data are entirely consistent with the relative sizes of the wild-type and mutant RNAs as observed on northern blots (Fig. 2).

Examination of the sequence of the wild-type *Brd* gene revealed that a 9-bp sequence motif, CAGCTTTAA, is directly repeated three times in the 3′ UTR (Fig. 4). We refer to this motif hereafter as the 'Brd box'. The *blood* transposon present in the gain-of-function alleles is inserted between the first and second of these elements, and causes the two most 3′ Brd boxes to be excluded from the mutant transcript (Figs 1, 4; see Discussion).

Brd encodes an unusually small novel protein

The *Brd* open reading frame encodes a predicted protein of only 81 aa (9.2 kDa) (Fig. 4). While the protein overall is slightly basic (predicted isoelectric point 8.65), it has a distinctly basic N-terminal half and a relatively acidic C-terminal half. The N-terminal region contains a high proportion of residues potentially subject to phosphorylation, including eight threonine and three tyrosine residues in the first 24 aa. In addition, there is a clear casein kinase II consensus phosphorylation site (Kemp and Pearson, 1990) near the *Brd* C terminus (Fig. 4).

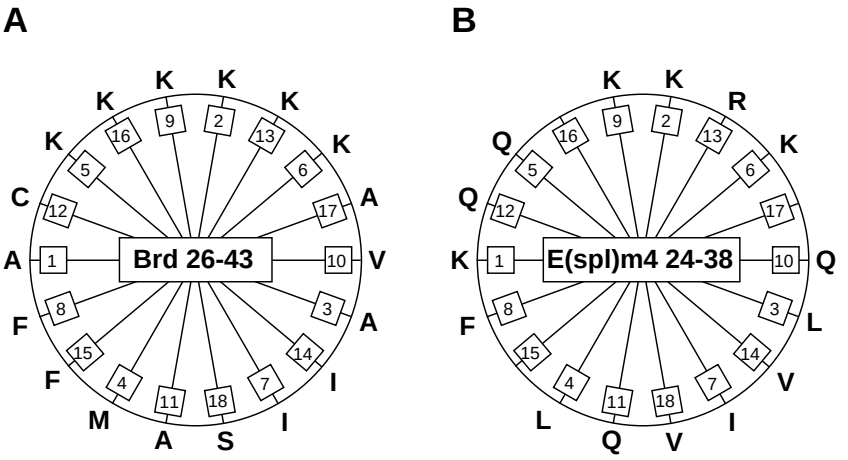

Fig. 5. Helical wheel plots illustrating predicted basic amphipathic α-helices in the deduced Brd (A) and E(spl)m4 (B) proteins. Numbers in the center of the wheel indicate the beginning and ending amino acid positions of the predicted helical region in each protein. Numbers on the spokes of the wheel indicate the linear order of the amino acids within the helical region.

The primary distinguishing structural feature of the *Brd* protein is an 18-aa region (residues 26-43; Fig. 4) that is strongly predicted to form a basic amphipathic α-helix. A helical wheel plot of this region (Fig. 5A) reveals the alignment of six positively charged lysine residues on one face of the helix, with largely hydrophobic residues on the opposing face.

Cloning and sequence analysis of the *Brd* protein coding region from three *Brd* point mutants (Leviten and Posakony, 1996) confirmed our identification of the *Brd* gene. Two of these mutants, Brd^{1rv1} and Brd^{1rv2}, exhibit identical 1-bp changes (relative to an isogenic parental chromosome) that generate stop codons in the middle of the protein just C-terminal to the *Brd* helical region (Q47 to stop, CAA to TAA). These mutations completely revert the dominant *Brd* phenotypes to wild type. The third mutant, Brd^{1rv3}, which is also a complete revertant, exhibits three nucleotide changes, creating two non-conservative amino acid changes in the N-terminal region of the protein (I11 to T, ATC to ACC; Q19 to E, CAG to GAA). Although the impact of an additional threonine residue in the already threonine-rich N terminus is uncertain, the substitution of an acidic glutamate residue adjacent to the basic helical region might very well affect charge-dependent interactions involving this region of the protein.

The *Brd* and *E(spl)m4* genes and proteins are related

Although the predicted Brd protein (Fig. 4) is not strongly related to any known protein in the GenBank database, we were interested to find that Brd shares weak sequence similarity with the product of the *m4* transcription unit of the E(spl)-C (Klämbt et al., 1989). Alignment of the Brd and E(spl)m4 sequences reveals two regions of limited similarity, one in the N-terminal portion of both proteins, and a second within acidic portions of both proteins (Fig. 6A). Since the region of highest identity overlaps the Brd amphipathic helix, we investigated whether the corresponding region of E(spl)m4 would likewise be expected to form a such a structure. Indeed, as shown in Fig. 5B, residues 24-38 of m4 are predicted to adopt an α-helical conformation with clear basic amphipathic character, suggesting that the Brd and E(spl)m4 proteins may have related biochemical functions.

Even more striking evidence for the re-latedness of the *Brd* and *E(spl)m4* genes was provided by the finding that they exhibit an unusual degree of nucleotide sequence identity in their 3′ UTRs. Fig. 6B shows an alignment of a continuous 40-nt region in which 31 nt are shared between *Brd* and *m4*. This region includes a 14/14 identity containing the third (most 3′) Brd box in *Brd*, as well as an adjacent 7/7 identity that we refer to as the ‘GY box’ (see Fig. 4 and below). A search of the

A

```
                1                                                   50
Brd             MAYETLMNTTIYTTTPVCQKKTYTPAKAMKKIFKVACKIFKASGHQQLKN
                     +NT    T      K +Y+  K + KIFK   ++
E(spl)m4         MCQNKINTNNTMTIKSNKKLSYSVKKLLQKIFKQQQRVEEEQNLKNALK
                 1                                                  49

                51                              81
Brd             HSMPKELPLPTSEEELQNSQNEQLEAELVQL
                          S E + NS N  LE   +
E(spl)m4               ANSLESLESMENSRNADLESASICASLESCENEANERLSQSCEI
                       50                                          93

E(spl)m4        EDYDFEQLPTVPVHFVRTAHGTFFWTAVSDLPADNDLVEPLYCSTSNAIAIPQDRWVQA
                94                                                                152
```

B

```
                460   Brd Box       GY Box            499
Brd             ATATCCAGCTTTAATCATTGTCTTCCATCTATGTTTAAGA
                ** ***************  ******* *  ** *** * *
E(spl)m4        ATCTCCAGCTTTAATCAACGTCTTCCGTTAATTTTTGAAA
                613                                 652
```

Fig. 6. Amino acid and nucleotide sequence alignments illustrating relatedness of the *Brd* and *E(spl)m4* genes and their predicted protein products. (A) Amino acid sequence similarity between Brd and E(spl)m4. The complete protein sequences (numbered in aa) are presented; amino acid identities are indicated by letter code, while conservative changes are identified by a + symbol. The predicted basic amphipathic α-helical regions of each protein are shown in bold. (B) Nucleotide sequence identity in a 40-nucleotide stretch of the *Brd* and *E(spl)m4* 3′ UTRs. Numbers are in nt from the transcription start site (this paper and Klämbt et al., 1989). The 31 identical bases in this region are indicated by asterisks. Shared Brd box and GY box sequence motifs are marked in bold.

GenBank database with the 23-nt sequence from *Brd* that contains both of these identities (Fig. 6B) reveals that *E(spl)m4* is unique among known genes in its degree of relatedness to *Brd* in this region.

Two shared sequence motifs in the 3′ UTRs of *Brd* and E(spl)-C genes

The observation that the 3′ UTRs of *Brd* and *E(spl)m4* share two extended sequence identities, the Brd box and the GY box, prompted us to search more systematically for these motifs in genes of the E(spl)-C. The results are presented in Fig. 7. We find that both sequence elements are present in the 3′ UTRs of most genes of the Complex, sometimes in multiple copies.

The Brd box is most often represented by the 7-nt core sequence AGCTTTA (Fig. 7A), although in two cases [E(spl)m4-B1 and E(spl)m5-B3], the full 9-nt element CAGCTTTAA (repeated three times in *Brd*) is found. In both of these cases, the Brd box motif is part of a longer sequence identity with *Brd* [the 14/14 identity noted above that includes Brd-B3 and E(spl)m4-B1, and a 12/12 identity that includes Brd-B2 and E(spl)m5-B3]. Overall, of the nine transcription units in the E(spl)-C for which appropriate sequence data are available [seven bHLH repressor-encoding genes plus *E(spl)m4* and *groucho* (*gro*)], the Brd box occurs in the 3′ UTRs of six of them (Fig. 7A). The exceptions are *E(spl)mβ*, *E(spl)m8* and *gro*.

The 7-nt GY box (GTCTTCC) is also present in the 3′ UTRs of multiple E(spl)-C genes (Fig. 7A). As with the Brd box, GY boxes are also frequently part of longer sequence identities between genes, such as the 11/11 block that includes E(spl)m3-G1 and E(spl)m5-G3, or the 10/10 identity that contains Brd-G1 and E(spl)m3-G1. No gene was identified that contains a GY box without an accompanying Brd box.

Fig. 7B illustrates the relative positions of Brd and GY boxes in the 3′ UTRs of *Brd* and the E(spl)-C genes. Overall, no consistent positional relationship can be discerned between the Brd and GY boxes in the different genes, or between these elements and the stop codon, polyadenylation signal or polyadenylation site.

Overexpression of *Brd* under control of a heat-shock promoter generates weak *Brd* gain-of-function phenotypes

To investigate the effects of general and/or overexpression of *Brd*, we constructed a fusion gene in which the *Brd* protein coding

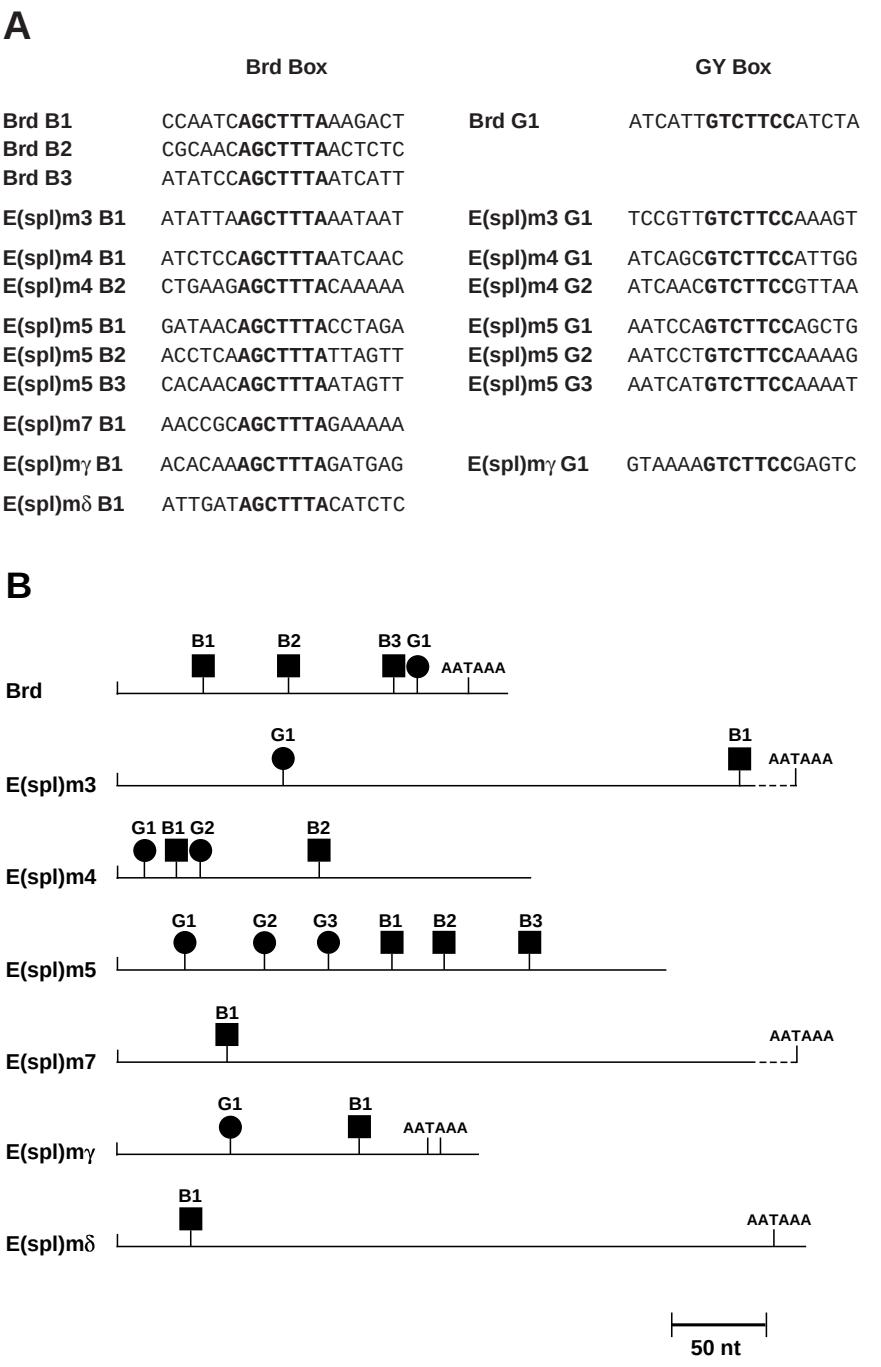

Fig. 7. Brd box and GY box sequence motifs found in the 3′ UTRs of *Brd* and several genes of the E(spl)-C. (A) Alignment of Brd box and GY box sequences (bold letters). Each motif (indicated by B or G) in a given gene is numbered in order of its 5′-to-3′ position in the 3′ UTR. Brd boxes are aligned by the core 7-nucleotide sequence AGCTTTA; GY boxes by the core GTCTTCC. Both motifs are frequently contained in longer sequence identities between pairs of sequences, as can be seen by inspection of the flanking nucleotides shown. The Brd box in *E(spl)mγ* was identified after correcting an apparent error in the published sequence (Knust et al., 1992). The GenBank accession number for our partial *E(spl)mγ* 3′ UTR sequence (from a full-length cDNA clone) is AF016543. (B) Diagram showing the locations of Brd boxes and GY boxes in the 3′ UTRs of *Brd* and E(spl)-C genes. Tick mark at the left end of each horizontal line represents the position of the stop codon in that gene. Length of each horizontal line indicates the extent of available sequence data. Where present, positions of consensus polyadenylation signals (AATAAA) are shown. Filled squares denote Brd boxes; filled circles, GY boxes. Numbering of motifs (B1, G1, etc.) corresponds to that in (A).

sequences are expressed under the control of the *Hsp70* heat-shock promoter. This chimeric gene is carried in the P element vector Casper-*Hsp70* (Bang and Posakony, 1992). The design of the construct was influenced by the possibility that the high levels of transcript responsible for the *Brd* mutant effects require the presence of *blood* element sequences (e.g. for LTR-dependent truncation of the *Brd* transcript). Accordingly, a 550-bp *Sal*I genomic DNA fragment from the Brd^{P1} allele, extending from the *Brd* 5′ UTR to a position approximately 180 bp inside the *blood* element (see Fig. 1), was used to create the *Hs-Brd* transgene. Five independent transformant lines carrying this construct were analyzed.

Heat-shock pulses applied during the late larval and early pupal stages (see Materials and methods) generated bristle multiplication and bristle loss phenotypes resembling those of *Brd* gain-of-function mutations. The bristle multiplication effects were not as severe as those observed for the Brd^{1} and Brd^{3} mutations, even in fly lines carrying four copies of the *Hs-Brd* construct. As expected, pulses of late third-instar larvae affected the earlier-developing macrochaetes, causing bristle duplications at some macrochaete positions, and a more reliable and general macrochaete loss phenotype (data not shown). Heat pulses applied during early pupal development (0, 6 and 16 hours APF) resulted in increases in microchaete number and density on the head and notum, comparable to those characteristic of weak *Brd* mutant genotypes (data not shown). In addition, significant microchaete loss was observed, and in some positions sockets without shafts were found, along with microchaetes with severely reduced shafts. These phenotypes have all been observed previously in *Brd* gain-of-function mutants. Finally, no general effects on epidermal cell or wing vein development were observed.

The P[*Hs-Brd*] transformant lines were exposed to P transposase activity in order to mobilize the transgene construct and potentially reposition it adjacent to strong enhancers in the genome, which would confer more stable, high levels of expression on the *Brd* gene. This approach has generated two lines of flies that exhibit strong bristle defects. One of these, Brd^{HSJ1}, displays strong dominant macrochaete and microchaete multiplication phenotypes similar to Brd^{3} heterozygotes (Fig. 8). Heterozygotes of this line exhibit microchaete tufting primarily in the anterior half of the notum, with a nearly wild-type microchaete pattern posteriorly. Homozygotes exhibit stronger and more uniform notum microchaete multiplication (Fig. 8), strong head and notum macrochaete multiplication (Fig. 8), and severely roughened eyes. The dominant bristle phenotypes of Brd^{HSJ1} are almost certainly due to overexpression of the *Brd* gene. The recessive eye roughening is potentially caused by a mutational effect of the transposon insertion on an endogenous gene, although gain-of-function alleles of *Brd* are capable of producing similar eye roughening defects (M. W. Leviten and J. W. Posakony, unpublished).

For comparison, we also extensively investigated the ability of four copies of a *Hs-Brd* construct containing a wild-type *Brd* 3′ UTR to phenocopy *Brd* gain-of-function phenotypes. We found that this construct is incapable of conferring any dominant mutant phenotypes under a variety of heat-shock regimens. Moreover, we failed to detect any dominant *Brd*-like effects in over 5,000 progeny flies in which this construct was mobilized [using the Δ2-3 system (Robertson et al., 1988)] in an effort to position it near strong enhancers in the genome.

Fig. 8. Adult PNS phenotype associated with overexpression of a mutant *Brd* gene under the control of the *Hsp70* promoter (see text for details). Shown is a dorsal view of the notum of a Brd^{HSJ1} homozygote.

These results are consistent with the hypothesis that the *blood* transposon-mediated truncation of the mutant *Brd* transcript (and the consequent loss of two Brd boxes and one GY box) is important to its ability to interfere with cell fate decisions in the adult PNS.

DISCUSSION

Identification of the *Brd* transcription unit

Several criteria indicate conclusively that we have identified the previously unknown transcription unit that corresponds to the *Brd* locus. Firstly, particular mutant alleles of *Brd* are associated specifically with insertions of transposable elements in or near this transcription unit; namely, a *blood* transposon in the 3′ UTR in the case of the dominant gain-of-function alleles, and a defective P element upstream of the transcription start site in the case of the partial revertant allele Brd^{P1}. Secondly, by comparison to wild-type or gain-of-function alleles, EMS-induced mutations that completely revert the *Brd* dominant phenotype consist of specific base substitutions in the open reading frame of the gene we have identified as *Brd*. Finally, overexpression of only the identified *Brd* gene under the control of a heat-shock promoter in transgenic flies partially phenocopies the dominant effects of the gain-of-function mutations.

Molecular basis of *Brd* gain-of-function phenotypes

We have investigated in detail the molecular lesions associated with the dominant gain-of-function mutations of *Brd* described in our earlier report (Leviten and Posakony, 1996). These alleles cause, with varying degrees of severity, sensory organ multiplication and loss phenotypes very similar to those resulting from loss-of-function mutations in the genes of the *N* signalling pathway (for a review see Posakony, 1994). Our analysis indicates that the likely origin of these *Brd* dominant phenotypes is the insertion of a retrotransposon of the *blood* family into the 3′ UTR of the gene. Due to the efficient usage of a polyadenylation signal present in the LTR of the transposon, transcripts from the mutant *Brd* gene are shorter than wild-type and lack the most 3′ 132 bases of the normal 3′ UTR sequence. In contrast, the predicted protein product of the gain-of-function alleles is unchanged from the wild type.

We have also found that steady-state levels of *Brd* transcript are several-fold higher in the gain-of-function mutants. Perhaps the simplest interpretation of this effect is that the truncated *Brd* transcript in these mutants has a significantly greater half-life, possibly due to the loss of instability-conferring elements in the distal 3′ UTR of the normal mRNA (see below). However, it is also possible that enhanced transcriptional activity of *Brd* caused by the LTRs or other sequences of the *blood* transposon contributes to the higher RNA levels we observe in the mutants.

Our previous genetic analysis (Leviten and Posakony, 1996) indicated that the *Brd* gain-of-function mutants almost certainly represent either neomorphic or strong hypermorphic alleles of the gene. Since, as described here, the predicted Brd protein in the mutants is the same as the wild type, and the spatial pattern of transcript accumulation in late third-instar imaginal discs (i.e. in proneural clusters) is also apparently unchanged, we favor the latter interpretation; namely, that the *Brd* dominant effects we have described result from hyperactivity of the normal Brd protein product. Nevertheless, it is important to note that we cannot presently rule out the possibility that a higher steady-state level of *Brd* RNA (particularly one of greater stability) at one stage might result in its persistence into a stage (or a cell type) from which it might normally be absent, thus yielding what is formally a neomorphic misexpression effect.

Spatial and temporal dynamics of *Brd* expression in imaginal disc proneural clusters

Consistent with the specific effects of *Brd* gain-of-function mutations on adult sensory organ development, and the lack of wing vein or wing blade defects common to many other bristle mutants, *Brd* is expressed specifically in the proneural clusters of the wing and other imaginal discs as a direct target of transcriptional activation by ac and sc (Singson et al., 1994). We have shown here that *Brd* transcript accumulation precedes the appearance of the early SOP marker A101, and that following the onset of A101 expression, *Brd* expression is diminished in at least some SOPs, as has been reported for ac/sc expression (Cubas et al., 1991; Skeath and Carroll, 1991). However, in contrast to ac, sc and the ac/sc target gene *scabrous* (*sca*), we have not observed consistent elevation of *Brd* transcript levels within presumptive SOPs in either wild-type or *Brd* mutant wing discs. Similarly, in situ hybridization fails to reveal in SOPs elevated levels of transcript from several of the genes of the E(spl)-C that are also direct targets of ac/sc regulation (Singson et al., 1994). This difference in SOP expression of proneural proteins on the one hand, and of certain of their direct regulatory targets [*Brd* and E(spl)-C genes] on the other, suggests that some form of transcriptional or post-transcriptional repression acts in the latter case. Some insight is afforded by the observation that when a *lacZ* reporter gene is expressed under the control of the *Brd* promoter, higher levels of β-galactosidase activity are often detected in SOPs (Singson et al., 1994). Similar results have been obtained for reporter genes driven by certain E(spl)-C promoter fragments (Bailey and Posakony, 1995; Kramatschek and Campos-Ortega, 1994; Singson et al., 1994). Thus, in assays that specifically reflect transcription-level regulation, these genes' promoters are found to be responsive as expected to the higher levels of proneural protein activity in the SOP.

We suggest, then, that regulatory mechanisms exist to reduce transcript accumulation from certain ac/sc activation targets in the SOP. While it is presumably advantageous to upregulate SOP expression of ac/sc target genes that promote the SOP fate, or are required for sending an inhibitory signal (e.g. *sca*; Baker et al., 1990; Mlodzik et al., 1990), the genes of the E(spl)-C act oppositely by antagonizing proneural activity, so high levels of these proteins in the presumptive SOP are undesirable. Indeed, their inappropriate expression in SOPs extinguishes this fate (Bang et al., 1995; Nakao and Campos-Ortega, 1996; Tata and Hartley, 1995). It seems reasonable that negative post-transcriptional modes of regulation, in particular, could play an essential role in limiting E(spl)-C transcript and protein accumulation in the SOP. It is not yet clear whether a similar rationale applies to *Brd*, because the mode of action of the *Brd* dominant mutants in generating supernumerary SOPs is unknown.

Amino acid, structural and nucleotide sequence similarities suggest related functions for *Brd* and *E(spl)m4*

A key result of our molecular analysis of the *Brd* gene is the finding that it is related to the *m4* transcription unit of the E(spl)-C. Though at present the function of neither gene is understood (see below), we have identified clear similarities in both the primary and secondary structures of their predicted protein products, and even more extensive (and unique) nucleotide sequence identities in their 3′ UTRs. Besides allowing us to recognize the shared GY box motif in the 3′ UTRs of *Brd* and the genes of the E(spl)-C, this discovery provides strong support for the hypothesis that *Brd* and *E(spl)m4* have related functions during development.

The presence of a predicted basic amphipathic α-helix in the deduced protein products of both *Brd* and *E(spl)m4* does not immediately suggest the nature of their biochemical activities. Helical domains of this type are associated with a variety of functions, including protein-protein interaction (e.g. binding of calmodulin) and interaction with lipid membranes (Segrest et al., 1990). It is certainly plausible that small proteins like Brd and E(spl)m4, lacking other apparent functional domains, might well exert their effects primarily via interaction with other proteins. We are currently pursuing this possibility in systematic screens for interacting protein partners, as well as by investigating the possible *in vivo* significance of the predicted interaction with calmodulin.

The available genetic data are unfortunately insufficient to give a clear indication of whether *Brd* and *E(spl)m4* exert similar regulatory effects *in vivo*. Neither gain- nor loss-of-function point mutations in the *E(spl)m4* gene have been reported, and accordingly its specific function within the proneural cluster (if any) has yet to be determined. Since the other genes of the E(spl)-C function to antagonize proneural activity, *E(spl)m4* might be expected to act similarly. In contrast, the simplest deduction from the *Brd* gain-of-function phenotypes (Leviten and Posakony, 1996) is that *Brd* normally acts to promote commitment to neural cell fates. Thus, even if the Brd and E(spl)m4 proteins rely on similar biochemical mechanisms of action (as suggested by their structural similarities), it is a distinct possibility that they have antagonistic, rather than cognate, regulatory effects. For example, both proteins might interact with a common protein or proteins, but have opposite effects on the activity of these factors.

In light of the structural evidence that their protein products may have related (though not necessarily cognate) biochemical functions, it is of course noteworthy that *Brd* and *E(spl)m4* are co-expressed in imaginal disc proneural clusters (Bailey and Posakony, 1995; Singson et al., 1994). Both genes are subject to direct transcriptional activation by the proneural proteins (Bailey and Posakony, 1995; Singson et al., 1994), while *E(spl)m4* (Bailey and Posakony, 1995), but apparently not *Brd* (A. Bailey, A. Singson and J. W. Posakony, unpublished observations), is also directly activated by Su(H) in response to N receptor activity. Thus, as N-mediated signalling in the proneural cluster proceeds, the expression of *E(spl)m4* would be expected to increase relative to that of *Brd*. As we learn more about the *in vivo* functions of the two genes, this distinction in their transcriptional regulation is likely to take on added significance.

Wild-type function of *Brd*

Our analysis of the molecular structure and expression of both wild-type and mutant *Brd* genes provides, we believe, further support for the hypothesis that *Brd* normally plays a role in adult sensory organ development in *Drosophila*, and more specifically in the cell-cell signalling events mediated by the *N* pathway (Leviten and Posakony, 1996). Previously, we had established that *Brd* is expressed in the proneural clusters of the wing imaginal disc under the direct control of the proneural transcriptional activators ac and sc, and that it is also expressed near the morphogenetic furrow of the eye disc (another site of *N*-mediated cell-cell interaction) under distinct, *ac/sc*-independent, control (Singson et al., 1994). The detailed similarity at the cellular level between *Brd* gain-of-function phenotypes and those caused by loss-of-function mutations in the genes of the *N* pathway, combined with the sensitivity of *Brd* dominant effects to the dosage of both *N* and *H*, constitute the strongest circumstantial evidence for a functional connection between *Brd* and the *N* pathway (Leviten and Posakony, 1996). Three findings described in the present paper appear to strengthen this association. First, we have presented molecular evidence that the gain-of-function alleles of *Brd* are hypermorphic, indicating that their phenotypic effects are likely to rely on the hyperactivity of a normal Brd protein expressed in its normal setting. Second, the temporal dynamics of *Brd* transcript accumulation in wing disc proneural clusters, clearly preceding the specification of the SOP, are consistent with a role in the cell-cell signalling events that lead to this cell fate choice. Finally, the clear structural relatedness of *Brd* and a *N* pathway-regulated gene, *E(spl)m4*, as well as the sharing of the Brd box and GY box motifs with *m4* and several other *N*-regulated transcription units in the E(spl)-C (see below), are especially compelling.

Sharing of novel 3′ UTR sequence motifs suggests a common mode of post-transcriptional regulation for *Brd* and genes of the E(spl)-C

We believe that one of the most important results of the present study is the identification of two novel heptanucleotide sequence motifs, the Brd box and the GY box, in the 3′ UTRs of *Brd* and the genes of the E(spl)-C. These elements, notable for the length of the strictly conserved sequence, are sometimes present in multiple copies in a single gene, and are often part of even more extensive sequence identities (up to 14/14) between particular pairs of genes. The existence of these motifs strongly suggests that *Brd* and most of the E(spl)-C genes are subject to a shared, and previously unrecognized, mode of regulation. Given their location and their preferential orientation, we currently favor the working model that the Brd and GY elements function as part of RNA transcripts to mediate some type of post-transcriptional control. However, it is also possible that they function instead in DNA, in an unusual mode of transcriptional regulation that relies on oriented sequence elements located in the 3′ UTR.

If the Brd and GY motifs do signify a common mode of post-transcriptional control for *Brd* and the E(spl)-C, there are several possibilities for the nature of this regulation, including RNA stability modulation, control of translation and RNA localization. We have described here our observation that the transposon-mediated elimination of two Brd boxes and a GY box from the *Brd* transcript is associated with a substantial increase in its steady-state level. This finding is consistent with the interpretation that either or both of these sequence elements normally function to confer instability on transcripts that include them. Indeed, we have recently found using reporter transgenes that specific mutation of the three Brd boxes in the *Brd* 3′ UTR results in a substantial increase in steady-state levels of both RNA and protein products *in vivo* (E. C. Lai and J. W. Posakony, unpublished observations).

At present it is perhaps simplest to suggest that the Brd and GY boxes represent sites of interaction of regulatory proteins that bind single-stranded RNA. If so, it is worth noting that the unusual conserved length of these motifs (7 nt) implies an unusual degree of sequence specificity for such RNA-binding activities. An alternative possibility, suggested by the precedent of the *lin-14* and *lin-4* genes of *C. elegans* (Ha et al., 1996; Lee et al., 1993), is that the Brd and GY elements bind complementary RNA molecules (encoded elsewhere), which are capable of forming short RNA duplexes that function in post-transcriptional regulation. Such RNA duplexes could themselves interact with regulatory proteins, either sequence-specifically or by sequence-non-specific recognition of the double-stranded structure.

Clearly, much remains to be learned about the functions of the Brd box and GY box sequence motifs. Nevertheless, it seems likely that understanding in detail the shared mode(s) of *Brd* and E(spl)-C regulation implied by the existence of these elements will be an important step in assembling an overall picture of the complex control of adult sensory organ development in *Drosophila*.

We would like to express our thanks to the following investigators for providing stocks and materials: Hugo Bellen, Nick Brown, Hilary Ellis, Josh Kavaler and Marc Muskavitch. We are very grateful to Dennis Spann for preparing staged embryo RNAs, to Mark Borowsky for his assistance with embryo injections, and to Jenee Wagner for her careful and patient work in preparing the figures. Adina Bailey, Ruth Bodner, Rick Firtel, Deborah Lycan and Dave Nellesen provided very helpful critical comments on the manuscript. This work was supported by an NIH grant (GM46993) to J. W. P. and by a graduate fellowship from the Lucille P. Markey Charitable Trust to E. C. L.

REFERENCES

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

Bailey, A. M. and Posakony, J. W. (1995). Suppressor of Hairless directly activates transcription of *Enhancer of split* Complex genes in response to Notch receptor activity. *Genes Dev.* **9**, 2609-2622.

Baker, N. F., Mlodzik, M. and Rubin, G. M. (1990). Spacing differentiation in the developing *Drosophila* eye: a fibrinogen-related lateral inhibitor encoded by *scabrous*. *Science* **250**, 1370-1377.

Bang, A. G., Bailey, A. M. and Posakony, J. W. (1995). *Hairless* promotes stable commitment to the sensory organ precursor cell fate by negatively regulating the activity of the *Notch* signaling pathway. *Dev. Biol.* **172**, 479-494.

Bang, A. G. and Posakony, J. W. (1992). The *Drosophila* gene *Hairless* encodes a novel basic protein that controls alternative cell fates in adult sensory organ development. *Genes Dev.* **6**, 1752-1769.

Bingham, P. and Chapman, C. (1986). Evidence that *white-blood* is a novel type of temperature-sensitive mutation resulting from temperature-dependent effects of a transposon insertion on formation of *white* transcripts. *EMBO J.* **5**, 3343-3351.

Brou, C., Logeat, F., Lecourtois, M., Vandekerckhove, J., Kourilsky, P., Schweisguth, F. and Israel, A. (1994). Inhibition of the DNA-binding activity of *Drosophila* Suppressor of Hairless and of its human homolog, KBF2/RBP-Jκ, by direct protein-protein interaction with *Drosophila* Hairless. *Genes Dev.* **8**, 2491-2503.

Brown, N. H. and Kafatos, F. C. (1988). Functional cDNA libraries from *Drosophila* embryos. *J. Mol. Biol.* **203**, 425-437.

Cabrera, C. V. and Alonso, M. C. (1991). Transcriptional activation by heterodimers of the *achaete-scute* and *daughterless* gene products of *Drosophila*. *EMBO J.* **10**, 2965-2973.

Cavener, D. R. (1987). Comparison of the consensus sequence flanking translational start sites in *Drosophila* and vertebrates. *Nucleic Acids Res.* **15**, 1353-1361.

Cubas, P., de Celis, J.-F., Campuzano, S. and Modolell, J. (1991). Proneural clusters of *achaete-scute* expression and the generation of sensory organs in the *Drosophila* imaginal wing disc. *Genes Dev.* **5**, 996-1008.

Delidakis, C. and Artavanis-Tsakonas, S. (1992). The *Enhancer of split* [*E(spl)*] locus of *Drosophila* encodes seven independent helix-loop-helix proteins. *Proc. Nat. Acad. Sci. USA* **89**, 8731-8735.

Dietrich, U. and Campos-Ortega, J. A. (1984). The expression of neurogenic loci in the imaginal epidermal cells of *Drosophila melanogaster*. *J. Neurogenetics* **1**, 315-332.

Ellis, H. M., Spann, D. R. and Posakony, J. W. (1990). *extramacrochaetae*, a negative regulator of sensory organ development in Drosophila, defines a new class of helix-loop-helix proteins. *Cell* **61**, 27-38.

Fehon, R. G., Kooh, P. J., Rebay, I., Regan, C. L., Xu, T., Muskavitch, M. A. and Artavanis-Tsakonas, S. (1990). Molecular interactions between the protein products of the neurogenic loci *Notch* and *Delta*, two EGF-homologous genes in Drosophila. *Cell* **61**, 523-534.

Fortini, M. E. and Artavanis-Tsakonas, S. (1994). The Suppressor of Hairless protein participates in Notch receptor signaling. *Cell* **79**, 273-282.

Furukawa, T., Kobayakawa, Y., Tamura, K., Kimura, K., Kawaichi, M., Tanimura, T. and Honjo, T. (1995). Suppressor of Hairless, the *Drosophila* homologue of RBP-Jκ, transactivates the neurogenic gene *E(spl)m8*. *Jpn. J. Genet.* **70**, 505-524.

Ha, I., Wightman, B. and Ruvkun, G. (1996). A bulged *lin-4/lin-14* RNA duplex is sufficient for *Caenorhabditis elegans lin-14* temporal gradient formation. *Genes Dev.* **10**, 3041-3050.

Hartenstein, V. and Posakony, J. W. (1989). Development of adult sensilla on the wing and notum of *Drosophila melanogaster*. *Development* **107**, 389-405.

Hartenstein, V. and Posakony, J. W. (1990). A dual function of the *Notch* gene in *Drosophila* sensillum development. *Dev. Biol.* **142**, 13-30.

Heitzler, P. and Simpson, P. (1991). The choice of cell fate in the epidermis of Drosophila. *Cell* **64**, 1083-1092.

Hsieh, J. J., Henkel, T., Salmon, P., Robey, E., Peterson, M. G. and Hayward, S. D. (1996). Truncated mammalian Notch1 activates CBF1/RBPJk-repressed genes by a mechanism resembling that of Epstein-Barr virus EBNA2. *Mol. Cell. Biol.* **16**, 952-959.

Huang, F., Dambly-Chaudière, C. and Ghysen, A. (1991). The emergence of sense organs in the wing disc of *Drosophila*. *Development* **111**, 1087-1095.

Jarriault, S., Brou, C., Logeat, F., Schroeter, E. H., Kopan, R. and Israel, A. (1995). Signalling downstream of activated mammalian Notch. *Nature* **377**, 355-358.

Jiang, J., Kosman, D., Ip, Y. T. and Levine, M. (1991). The *dorsal* morphogen gradient regulates the mesoderm determinant *twist* in early *Drosophila* embryos. *Genes Dev.* **5**, 1881-1891.

Kemp, B. E. and Pearson, R. B. (1990). Protein kinase recognition sequence motifs. *Trends Biochem. Sci.* **15**, 342-346.

Klämbt, C., Knust, E., Tietze, K. and Campos-Ortega, J. (1989). Closely related transcripts encoded by the neurogenic gene complex *Enhancer of split* of *Drosophila melanogaster*. *EMBO J.* **8**, 203-210.

Knust, E., Schrons, H., Grawe, F. and Campos-Ortega, J. A. (1992). Seven genes of the *Enhancer of split* complex of *Drosophila melanogaster* encode helix-loop-helix proteins. *Genetics* **132**, 505-518.

Kramatschek, B. and Campos-Ortega, J. A. (1994). Neuroectodermal transcription of the *Drosophila* neurogenic genes *E(spl)* and *HLH-m5* is regulated by proneural genes. *Development* **120**, 815-826.

Lecourtois, M. and Schweisguth, F. (1995). The neurogenic Suppressor of Hairless DNA-binding protein mediates the transcriptional activation of the *Enhancer of split* Complex genes triggered by Notch signaling. *Genes Dev.* **9**, 2598-2608.

Lee, R. C., Feinbaum, R. L. and Ambros, V. (1993). The C. elegans heterochronic gene *lin-4* encodes small RNAs with antisense complementarity to *lin-14*. *Cell* **75**, 843-854.

Leviten, M. W. and Posakony, J. W. (1996). Gain-of-function alleles of *Bearded* interfere with alternative cell fate decisions in *Drosophila* adult sensory organ development. *Dev. Biol.* **176**, 264-283.

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

Margolis, J. S., Borowsky, M., Shim, C. W. and Posakony, J. W. (1994). A small region surrounding the distal promoter of the *hunchback* gene directs maternal expression. *Dev. Biol.* **163**, 381-388.

Mlodzik, M., Baker, N. E. and Rubin, G. M. (1990). Isolation and expression of *scabrous*, a gene regulating neurogenesis in *Drosophila*. *Genes Dev.* **4**, 1848-1861.

Nakao, K. and Campos-Ortega, J. A. (1996). Persistent expression of genes of the *Enhancer of split* complex suppresses neural development in *Drosophila*. *Neuron* **16**, 275-286.

Parks, A. L. and Muskavitch, M. A. (1993). *Delta* function is required for bristle organ determination and morphogenesis in *Drosophila*. *Dev. Biol.* **157**, 484-496.

Posakony, J. W. (1994). Nature versus nurture: asymmetric cell divisions in Drosophila bristle development. *Cell* **76**, 415-418.

Price, B. D., Chang, Z., Smith, R., Bockheim, S. and Laughon, A. (1993). The *Drosophila neuralized* gene encodes a C3HC4 zinc finger. *EMBO J.* **12**, 2411-2418.

Robertson, H. M., Preston, G. R., Phillis, R. W., Johnson-Schlitz, D. M., Benz, W. K. and Engels, W. R. (1988). A stable source of P-element transposase in *Drosophila melanogaster*. *Genetics* **118**, 461-470.

Romani, S., Campuzano, S., Macagno, E. R. and Modolell, J. (1989). Expression of *achaete* and *scute* genes in *Drosophila* imaginal discs and their function in sensory organ development. *Genes Dev.* **3**, 997-1007.

Rooke, J., Pan, D., Xu, T. and Rubin, G. M. (1996). KUZ, a conserved metalloprotease-disintegrin protein with two roles in *Drosophila* neurogenesis. *Science* **273**, 1227-1231.

Schweisguth, F. (1995). *Suppressor of Hairless* is required for signal reception during lateral inhibition in the *Drosophila* pupal notum. *Development* **121**, 1875-1884.

Schweisguth, F. and Posakony, J. W. (1992). *Suppressor of Hairless*, the Drosophila homolog of the mouse recombination signal-binding protein gene, controls sensory organ cell fates. *Cell* **69**, 1199-1212.

Schweisguth, F. and Posakony, J. W. (1994). Antagonistic activities of *Suppressor of Hairless* and *Hairless* control alternative cell fates in the *Drosophila* adult epidermis. *Development* **120**, 1433-1441.

Segrest, J., De Loof, H., Dohlman, J., Brouillette, C. and Anantharamaiah, G. (1990). Amphipathic helix motif: Classes and properties. *Proteins* **8**, 103-117.

Singson, A., Leviten, M. W., Bang, A. G., Hua, X. H. and Posakony, J. W. (1994). Direct downstream targets of proneural activators in the imaginal disc include genes involved in lateral inhibitory signaling. *Genes Dev.* **8**, 2058-2071.

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

Smoller, D., Friedel, C., Schmid, A., Bettler, D., Lam, L. and Yedvobnick, B. (1990). The *Drosophila* neurogenic locus *mastermind* encodes a nuclear protein unusually rich in amino acid homopolymers. *Genes Dev.* **4**, 1688-1700.

Struhl, G., Fitzgerald, K. and Greenwald, I. (1993). Intrinsic activity of the Lin-12 and Notch intracellular domains in vivo. *Cell* **74**, 331-345.

Tamura, K., Taniguchi, Y., Minoguchi, S., Sakai, T., Tun, T., Furukawa, T. and Honjo, T. (1995). Physical interaction between a novel domain of the receptor Notch and the transcription factor RBP-Jκ/Su(H). *Curr. Biol.* **5**, 1416-1423.

Tata, F. and Hartley, D. A. (1995). Inhibition of cell fate in *Drosophila* by *Enhancer of split* genes. *Mech. Dev.* **51**, 305-315.

Tautz, D. and Pfeifle, C. (1989). A non-radioactive *in situ* hybridization method for the localization of specific RNAs in *Drosophila* embryos reveals translational control of the segmentation gene *hunchback*. *Chromosoma* **98**, 81-85.

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

Vässin, H., Vielmetter, J. and Campos-Ortega, J. A. (1985). Genetic interactions in early neurogenesis of *Drosophila melanogaster*. *J. Neurogenetics* **2**, 291-308.

(Accepted 6 August 1997)