MOLECULAR AND CELLULAR BIOLOGY, Nov. 1997, p. 6727–6735 Vol. 17, No. 11
0270-7306/97/$04.00+0

Tissue-Specific Splicing and Functions of the *Drosophila* Transcription Factor Grainyhead

ANNE E. UV,† EMMA J. HARRISON, AND SARAH J. BRAY*

Department of Anatomy, University of Cambridge, Cambridge CB2 3DY, United Kingdom

Received 29 May 1997/Returned for modification 16 July 1997/Accepted 18 August 1997

Grainyhead belongs to a recently identified group of transcription factors which share a 250-amino-acid domain required for binding to DNA and a carboxy-terminal dimerization domain. The activities of Grainyhead and other members of the family appear to be modulated so that they can participate in different developmental processes. We have examined the structure and function of mRNAs from the *Drosophila grainyhead* gene and demonstrated that alternate splicing is responsible for generating a neuroblast-specific isoform of the protein. A mutation which abolishes this isoform results in pupal and adult lethality. Reporter genes containing different Grainyhead binding sites exhibit tissue-specific patterns of expression that correlate with the Grainyhead isoforms, suggesting that the alternate splicing serves to alter the repertoire of target genes controlled in the neuroblasts.

During development, relatively small numbers of transcription factors are available to coordinate the complex and dynamic patterns of gene activities within the organism. Thus, the ability of many of these factors to be employed in more than one developmental process is essential for providing the required complexity. There are several ways this could be accomplished. One is through the accessibility of target genes. Another is through modifications in the behaviors of the transcription factors themselves, either through changes in their structures or through their interactions with one another, as in the formation of heterodimers. It is thus important to determine how the activity of transcription factors is modulated to allow them to participate in the regulation of different genes during development.

The transcription factor encoded by the *Drosophila* gene *grainyhead* (*grh*) is present in several tissues, where it appears to participate in different developmental programs. Grainyhead (GRH) is essential for development; embryos which lack *grh* function do not hatch, and they have flimsy cuticles, grainy and discontinuous head skeletons, and patchy tracheal tubes (7). GRH has therefore been suggested to play a role in activating genes whose products are involved in cuticle production, and consistent with this, GRH is found in cuticle-secreting cells during embryogenesis (6). One putative target gene of GRH in these cells is *Ddc*, the gene encoding dopa decarboxylase; this gene is expressed in epidermal cells during the time of cuticle secretion and is essential for hardening and coloring of the cuticle (1, 21). Two GRH binding sites have been identified within the *Ddc* gene (reference 18 and unpublished data) from which GRH can activate transcription in vitro (11), and expression of *Ddc* in epidermal cells requires GRH (7). In addition to being expressed in cuticle-producing tissues, GRH is also present in a particular set of cells in the embryonic central nervous system (CNS), and it has also been called neuronal transcription factor 1 (NTF-1) (3, 4, 11). No particular function has yet been ascribed to GRH in these cells. However, since neuronal cells do not produce cuticle, it is likely that GRH plays a distinct role in the two tissue types. In support of this, GRH does not influence neuronal expression of *Ddc* (7). GRH is also present in the newly fertilized embryo, where it may function as a repressor, contributing to the mechanisms which restrict the expression of *tailless* and *decapentaplegic* to particular domains (16, 22).

GRH belongs to a recently identified group of transcription factors which includes human proteins that may modulate transcription of human immunodeficiency virus type 1 (HIV-1) (27, 34, 35). These proteins are related through a region of similarity which spans approximately 250 amino acids and is required for binding to DNA (3, 23, 28, 33, 34). In addition, both GRH and other members of the family (CP2/LSF) have been shown to homodimerize (3, 28, 33), and although the DNA-binding domain on its own is sufficient for high-affinity DNA binding, the adjacent dimerization domain in GRH greatly stabilizes the DNA-protein interaction (33). There is evidence that the activity of the GRH-CP2 family of transcription factors is modified to allow them to participate in distinct developmental processes or to confer positive versus negative regulation. For example, human CP2 is implicated in the stage-specific expression of globin genes (23). Since CP2 itself is expressed ubiquitously, it appears that this specificity resides in its ability to form heteromeric complexes with another transcription factor, and indeed a candidate protein has been identified biochemically (17).

We set out to investigate mechanisms which could account for the participation of GRH in differentiation of the epidermis and the CNS. Analysis of *grh* cDNAs indicated the existence of mRNAs encoding different protein isoforms which could have distinct functions (6). Here we present evidence that one specific isoform is generated in the CNS and appears to be involved in equipping GRH with a tissue-specific activity. A *grh* mutant which abolishes this isoform lacks GRH protein in the CNS and results in lethality during postembryonic development, consistent with nervous system-specific effects. Since the CNS-specific isoform does not exhibit any novel characteristics in vitro, we propose that its effects are mediated through interactions with other proteins.

* Corresponding author. Mailing address: Department of Anatomy, University of Cambridge, Downing Street, Cambridge CB2 3DY, United Kingdom. Phone: 44-1223-333792. Fax: 44-1223-333786. E-mail: sjb32@mole.bio.cam.ac.uk.

† Present address: Umeå Center for Molecular Pathogenesis, Umeå University, S-90187 Umeå, Sweden.

MATERIALS AND METHODS

Isolation and mapping of genomic phage. To characterise the *grh* genomic structure, we isolated overlapping lambda genomic clones spanning the entire *grh* gene from a *Drosophila melanogaster* library (in EMBL3; a gift of Ron Blackman) by standard procedures. Initially a probe corresponding to the entire *grh* cDNA was used, and subsequently further phages were isolated by "walking," using fragments from the existing lambda phage. Similar techniques were used to characterize part of the *grh* gene in *Drosophila virilis*, and three independent lambda clones were isolated from a genomic library (a gift of Phil Ingham) by low-stringency hybridization, using a 177-bp fragment of the *grh* cDNA (corresponding to bp 3251 to 3428) as a probe (hybridization was performed at 37°C in 4× SSC [1× SSC is 0.15 M NaCl plus 0.015 M sodium citrate), and the washes were performed at 22 and 45°C in 2× SSC).

After constructing a restriction map of the region covered by both the *D. melanogaster* and *D. virilis* lambda clones, we used Southern blotting to determine the restriction fragments which hybridized to the *grh* cDNA. The DNA was transferred onto Zeta-Probe membranes (Bio-Rad) by the alkaline method described by the manufacturer, and these membranes were subsequently incubated with radioactive fragments prepared from the *grh* cDNA under the conditions described by Church and Gilbert (9) and modified in the protocol provided with the Zeta-Probe membranes. Fragments that hybridized to the cDNA fragments were subsequently subcloned into Bluescript KS (Stratagene). In each case, a series of overlapping deletions was then constructed by using exonuclease III according to the method of Henikoff (14). Selection of clones with deletions near the beginning and end of each exon was based on hybridization to specific *grh* fragments as detected by Southern blotting of the series of overlapping clones. Selected clones were sequenced by using Sequenase (U.S. Biochemicals) in accordance with the manufacturer's procedures and searched for matches to the cDNA sequence by using Macvector (International Biotechnology Inc.) or Wisconsin (Genetics Computer Group) sequence analysis programs.

Antibody staining of embryos and larvae. Larval tissues were dissected in BSS (10 mM Tricine base, 55 mM NaCl, 40 mM KCl, 7 mM $MgCl_2$, 5 mM $CaCl_2$, 20 mM glucose, 50 mM sucrose [pH 6.95]). For anti-GRH staining, the tissues were fixed for 10 min on ice in 4% paraformaldehyde (Polysciences) and EM buffer [100 mM KCl, 40 mM NaCl, 5 mM disodium EGTA, 1 mM spermidine HCl, 0.4 mM spermidine HCl, 30 mM piperazine-*N*,*N*′-bis(2-ethanesulfonic acid) (PIPES; pH 7.4), 0.2% β-mercaptoethanol]. The tissues were blocked for 1 h in PBT (phosphate-buffered saline [PBS; 130 mM NaCl, 7 mM Na_2HPO_4, 3 mM NaH_2PO_4) containing 0.3% Triton X-100 and 0.5% bovine serum albumin), and all subsequent washes employed this buffer. Incubation with anti-GRH monoclonal antibody (1:5) (6) in PBT was at 4°C overnight. After incubation with the primary antibody, the tissues were washed and subsequently incubated for 3 to 5 h with fluorescein isothiocyanate-conjugated anti-mouse antibody (Jackson) (1:200). After being washed, the preparations were mounted in CITIFLUOR (City University of London, London, United Kingdom) and viewed with a standard fluorescence microscope.

Anti-dopa decarboxylase (DDC) staining of the larval CNS was performed essentially as described previously (19): larval CNSs were fixed in 10% formalin and PBS for 1 to 2 h at room temperature, left for 15 min in methanol, washed four times for 5 min each in PBT, and preincubated for 1 h in PBT. Primary antibody (anti-DDC rabbit antiserum; a gift from Larry Marsh) was added at 1:250, and the mixture was incubated for 1 h at 37°C. Subsequently the specimens were washed in PBT and incubated with secondary antibody (rhodamine-conjugated anti-rabbit) at 1:200 for 1 h at 37°C. After being washed, the tissues were mounted as described above.

In situ hybridization to larval tissues. In situ hybridization with digoxigenin-labelled DNA probes was carried out essentially as described previously (30). Fragments of *grh* genomic DNA (spanning exons 4 and 5) and cDNA (corresponding to the DNA-binding domain) were labelled by random priming with a DNA labelling and detection kit (Boehringer). The larval tissues were dissected in ice-cold PBS and fixed for 20 min at room temperature in 4% paraformaldehyde in PBS. After fixation, the tissues were washed for 10 min in PBS. The subsequent steps of pretreatment, hybridization, and staining were performed as described previously (30). The tissues were mounted in 50% glycerol.

5-Bromo-4-chloro-3-indolyl-β-D-galactopyranoside (X-gal) staining of larval tissues. Larval tissues were dissected in PBS, fixed for 5 min in 2% glutaraldehyde in PBS, stained by a standard procedure (2, 13), and then mounted in 10% glycerol.

Preparation and sequencing of genomic DNA from larvae. Five larvae each of the grh^{370} homozygote (*b pr cn* wx^+ grh^{370} *bw*) and the parent (*b pr cn* wx^{wxt} *bw*) were homogenized in 400 μl of H buffer (160 mM sucrose, 100 mM Tris-Cl [pH 8.0], 80 mM EDTA) and treated with proteinase by adding 25 μl of a 10-mg/ml proteinase K solution (Boehringer) and 50 μl of 10% sodium dodecyl sulfate and incubating at 68°C for 1 h. After phenol-chloroform extraction, 14 μl of 8 M potassium acetate was added per 100 μl of aqueous solution and the solution was kept on ice for 30 min. The precipitate was removed by microcentrifugation, and the DNA was precipitated from the supernatant by addition of 2 volumes of ethanol. The resulting pellet was dissolved at 68°C in 25 μl of Tris-EDTA containing 10 mg of RNase per ml.

The following sense and antisense oligonucleotides corresponding to the *grh* cDNA sequence were synthesized (Oswel) and used in PCR to generate fragments of exons 4 and 5 from the above genomic DNA: 5′ TGA GAA TTC ATT TCC TTA GCC ACC G 3′ and 5′ TCT GGA TCC TAA TGG ACA CGC AAC GC 3′. Two 2.5-μl aliquots from each DNA preparation were used in four independent PCR reactions using a combination of *Taq* (Bioline) and *Pfu* (Stratagene) polymerases in a total volume of 50 μl. The fragments obtained were cleaved with *Eco*RI and *Bam*HI and cloned into Bluescript KS for subsequent dideoxynucleotide sequencing (Sequenase; U.S. Biochemicals). Two clones from each PCR reaction were sequenced.

***Drosophila* strains.** For a description of marker and balancer chromosomes, see reference 23a. $Df(2R)Pcl^{7B}/CyO$ was described previously (10). grh^{B37} contains a stop codon in exon 14 (7), and the grh^{370} allele (*b pr cn* wx^+ grh^{370} *bw*/*CyO Roi*) was obtained from R. Tearle and was generated by ethyl methanesulfonate mutagenesis. The grh^{370} studies were carried out with grh^{370} homozygotes or $grh^{370}/Df(2R)Pcl^{7B}$ transheterozygotes. The *GRHbe1-lacZ* and *GRHbe2-lacZ* transformant lines were generated by inserting trimers of the be1 (5′-ctagaCTC GCATAACCGGTTTCCAAGt or be2 (5′-ctagaGCGATTGAACCGGTCCTG CGGt) oligonucleotides (lowercase letters denote the bases for the restriction site; *grh* binding sites are underlined) into the *Xba*I site of the HZ50PL reporter vector (15) and injecting the resulting constructs into *cinnabar rosy* host flies to generate $rosy^+$ transformants. Multiple lines of each construct were analyzed.

RESULTS

Genomic organization of *grh*. Previous studies of *grh* cDNAs have suggested that several different *grh* transcripts are produced. Therefore, we embarked on an analysis of the *grh* gene to determine the full extent of potential variation. Genomic clones spanning the sequenced cDNA (designated N) (see Fig. 1 and reference 6) were isolated, and the initial mapping showed that the *grh* gene spans at least 37 kb. Subsequent characterization of the individual exons revealed that there are a total of 16 of them (Fig. 1A), with the exons making up the carboxyl 511 amino acids being much smaller and more closely spaced than those forming the amino part of the protein. This is particularly the case for the DNA-binding domain, which encompasses 263 amino acids and is contained within seven exons, many of which are very small (for example, exons 9 and 10 contain 94 and 56 bp, respectively). Given both the large size of the *grh* gene and its subdivision into many exons, there is considerable potential for alternative splicing to produce different isoforms.

We first characterized the four different classes of cDNAs isolated previously (N, N′, O, and O′ [Fig. 1C]) and found two sites of variation. One is a small intron within exon 12 which remains unspliced in some cDNA variants (e.g., N) and codes for an extra 30 amino acids. The other consists of exons 4 and 5, which are only present in the rarer large cDNAs (O and O′) and lead to an insertion of 810 bp. We also examined the genomic structure of the carboxyl part of the *grh* gene from the distantly related *Drosophila* species *D. virilis* and found that its organization is identical, apart from the precise lengths of some introns. This includes exons 4 and 5, which are significantly conserved (60 and 80% amino acid identity, respectively), although the degree of conservation is lower than that for the exons comprising the DNA-binding and dimerization domains (98% amino acid identity). However, the intron within exon 12, which in some cDNAs from *D. melanogaster* remains unspliced (N and O; labelled 12b in Fig. 1), has no obvious sequence similarity in *D. virilis* and indeed contains a stop codon, indicating that its coding potential is not conserved.

We subsequently analyzed another 22 *D. melanogaster grh* cDNAs isolated during different stages of embryonic development and from imaginal discs and found no evidence of other splicing variants. In addition, it appears that exons 4 and 5 are always inserted together, as we have not detected any evidence of the inclusion of one of these exons without the other in any of the cDNAs analyzed. Thus, despite the complex organization of the *grh* gene, we have evidence for only two splicing variants, one of which is conserved in *D. virilis*. This change, present in the O isoform, leads to an insertion of 270 amino

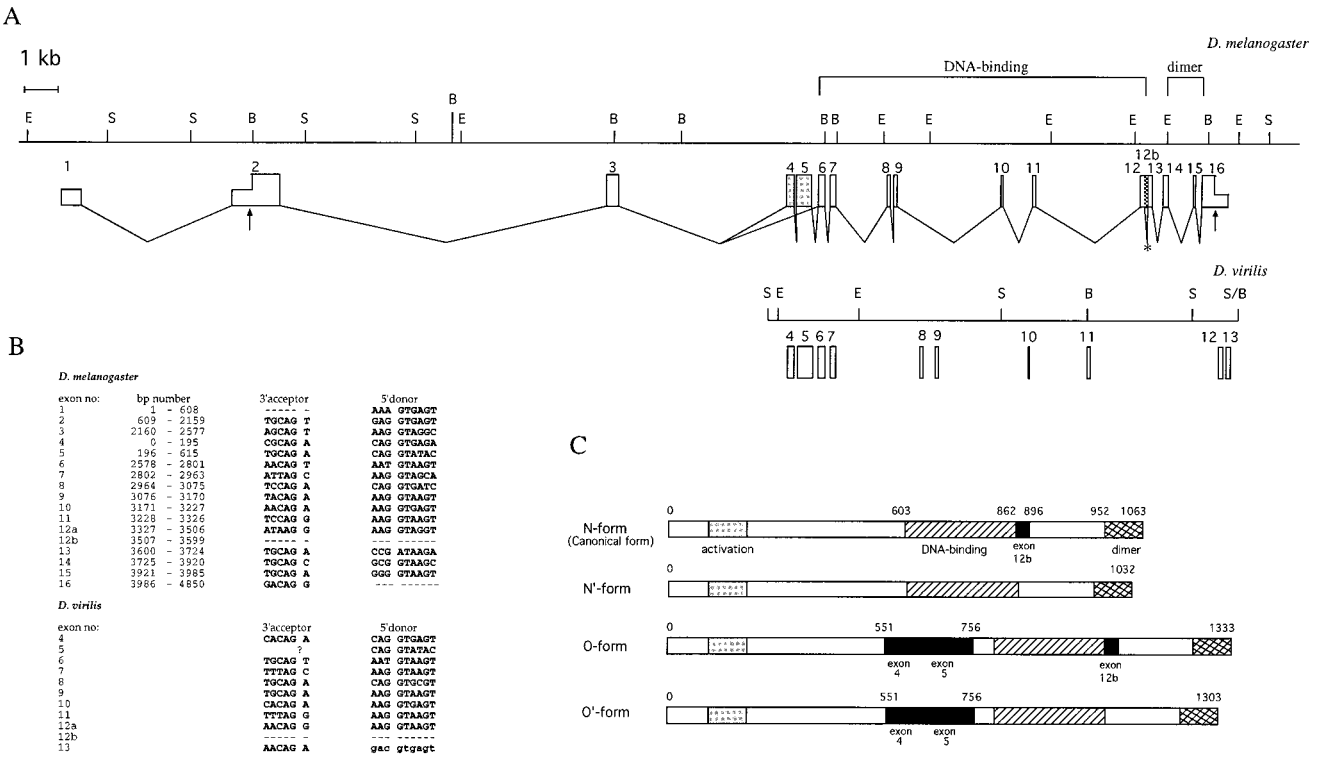

FIG. 1. Intron-exon structure of the *grh* gene and the composition of *grh* mRNAs. (A) Organization of the *grh* genes in *D. melanogaster* and *D. virilis*. The two horizontal lines represent the genomic DNAs of the *grh* genes in *D. melanogaster* (upper) and *D. virilis* (lower). Restriction sites for *Eco*RI, *Bam*HI, and *Sal*I are indicated above the line as E, B, and S, respectively. Boxes denote the positions of exons, and the numbers above them are the exon numbers; vertical arrows indicate the start and termination of the translated region. The transcription start site and polyadenylation site have not been identified; the first and last exons represent the extent of the longest cDNA obtained. Shaded boxes indicate exons which are alternately used; the asterisk marks alternate exon 12b. The exons encoding the DNA-binding and dimerization domains are denoted by brackets above the genomic DNA. (A scale bar is indicated at the top left corner.) (B) The DNA sequences for *D. melanogaster* and *D. virilis* corresponding to the splice donor and splice acceptor sites of each exon are shown. (C) The four identified GRH isoforms are presented. Each mRNA form is illustrated with a horizontal box, and their names are denoted to the left. The numbers above the boxes represent the amino acids at various positions: the N-terminal and C-terminal ends of each protein, the sites of alternate splicing, and, for the canonical N form, each end of the DNA-binding and dimerization domains. Lightly stippled boxes indicate the positions of the activation domains, hatched boxes indicate the DNA-binding domains, cross-hatched boxes indicate dimerization domains, and heavily stippled (dark) boxes indicate the alternate exons (3, 33). The N form is the mRNA that has been entirely sequenced, and thus we designate it the canonical form; and the protein that it encodes is called the GRH protein throughout the text. The O form contains an additional 270 amino acids (encoded by exons 4 and 5 [dark shading]) compared to the N form. The N′ form differs from the N form in that it lacks 30 amino acids downstream of the DNA-binding domain (exon 12b [dark shaking]), and the O′ form differs from the O form in lacking the same 30 amino acids.

acids next to the DNA-binding domain. We tested whether inclusion of this exon alters the interaction of GRH with its DNA target sites in vitro but were unable to detect any differences in binding affinity or in GRH dimerization (data not shown).

***grh* is spliced in the CNS in a tissue-specific manner.** The conservation of the alternately spliced exons 4 and 5 in *D. virilis* suggests that they encode important functions of the GRH protein. To investigate the significance of the alternate use of exons 4 and 5, we examined when and where the two GRH isoforms N and O are expressed during development. The wild-type expression of GRH protein during embryonic development has been described previously (6). GRH is expressed in the embryonic CNS, epidermis, and small regions of the fore- and hindgut, all of which are derived from the ectoderm. We first extended the analysis of GRH expression to larval development by using a monoclonal antibody which recognizes all identified GRH isoforms (6). GRH expression persists in epidermal cells and parts of the foregut and hindgut during larval life, and its expression is also clearly detected in other cuticle-producing tissues, such as the tracheae, anterior spiracles, proventiculus, and the cells that secrete the head skeleton (Fig. 2A to D). In addition, GRH is expressed uniformly in the imaginal discs and in the optic lobes of the CNS (Fig. 2E and F), both of which derive as pockets of cells from the embryonic ectoderm. In the CNS, GRH expression is also detected in the neuroblasts (NBs) throughout larval life (Fig. 2E). This also is a continuation from embryonic development since GRH expression marks the persistent imaginal NBs at the end of embryogenesis. No expression was detected in mesodermally derived cells, the endoderm (including the anterior and posterior midgut rudiments), or the peripheral nervous system.

To establish the distribution of the two *grh* transcripts N and O in larvae, in situ hybridization was performed with two fragments, one corresponding to the exons specific to the O transcript (exons 4 and 5) and one common to all transcripts (DNA-binding domain regions underlined in Fig. 1C). Using the common probe, a strong signal was detected in all the tissues in which GRH protein had been observed: the larval epidermis, tracheae, foregut, and imaginal discs (data not shown), as well as the NBs and optic lobes (Fig. 3A). However, the mRNA detected by the O-specific fragment shows a different profile. The main site of expression is the CNS, where the O-isoform mRNA is present at high levels in the NBs but is not detectable in the developing optic lobes (Fig. 3B). Ex-

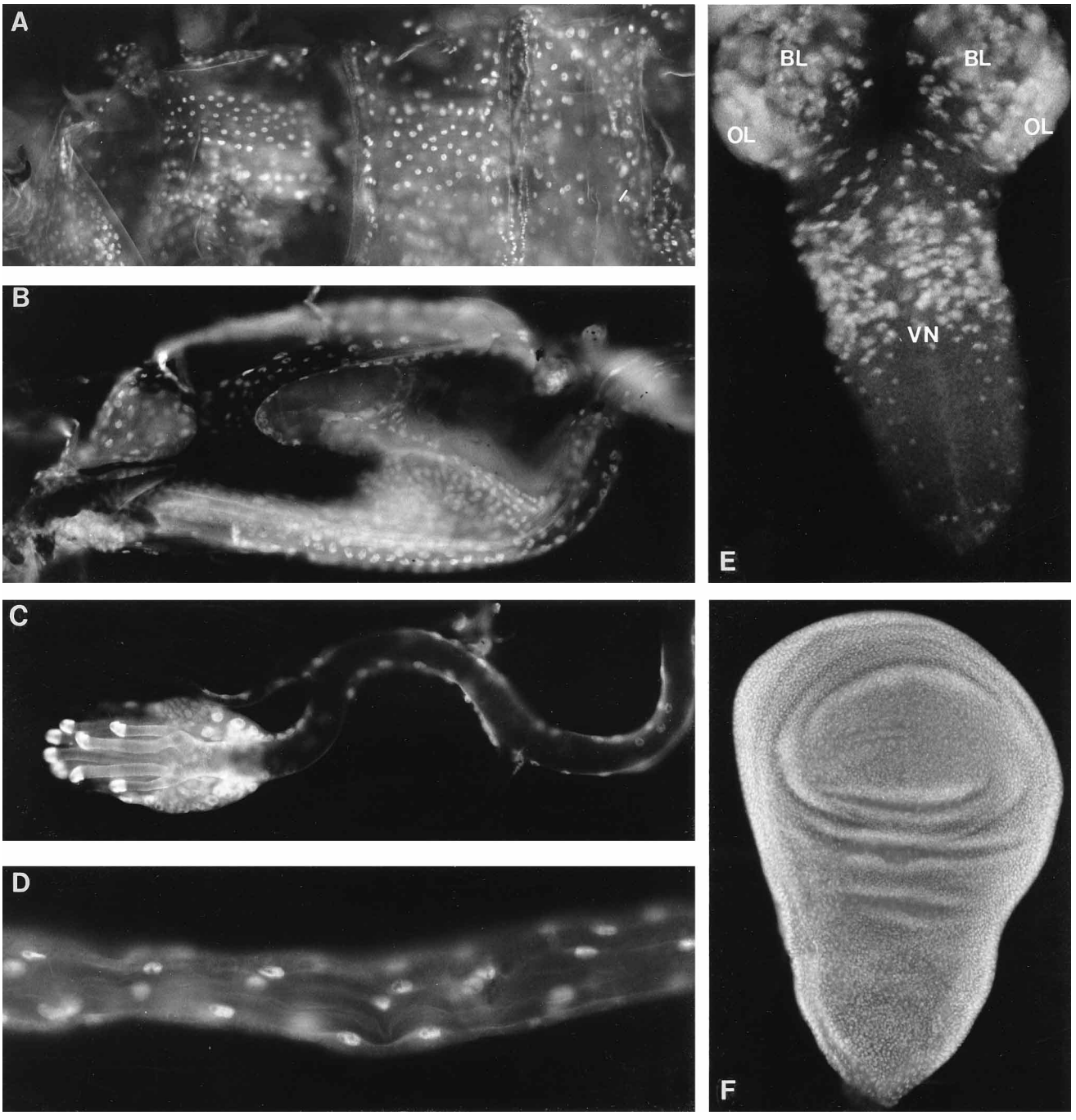

FIG. 2. GRH expression persists in cuticle-secreting cells, in the imaginal discs, and in the NBs of the CNS throughout larval life. (A to D) Nuclear GRH expression is detected in all cuticle-secreting tissues, such as the epidermis (A) (anterior part of the epidermis with underlying cuticle attached), the head skeleton (B), the anterior spiracle (left) and attached trachea (C), and the foregut (D). (E) In addition, GRH is present in specific cells in the CNS: in the NBs that lie in the brain lobes (BL) and ventral nerve cord (VN), and in the proliferating cells of the developing optic lobe (OL). (F) GRH is also expressed in all imaginal discs; shown here is a wing imaginal disc for which all the nuclei in the disc epithelium express GRH. Tissues were dissected from late- (A to D and F) or mid- (E) third-instar larvae.

pression of the O isoform in the other tissues is at low levels (discs and tracheae) or almost undetectable (epidermis). Similar results have been achieved in embryos, where O mRNA is predominantly detected in the NBs of the CNS (31a). Although it is difficult to determine by in situ hybridization whether the N and N′ *grh* mRNAs are also expressed in the NBs, the results presented below indicate that they are not. Thus, the observation that the O-specific transcript is only expressed in the NBs demonstrates that *grh* is spliced in the CNS in a cell-specific manner.

A novel mutation specifically disrupts the O isoform of GRH, resulting in absence of GRH protein from the NBs and larval-pupal lethality. The majority of *grh* alleles isolated (five of six) result in embryonic lethality. However, one allele, grh^{370}, differs in that the larvae containing it survive beyond embryogenesis. When a fly with this allele is crossed with

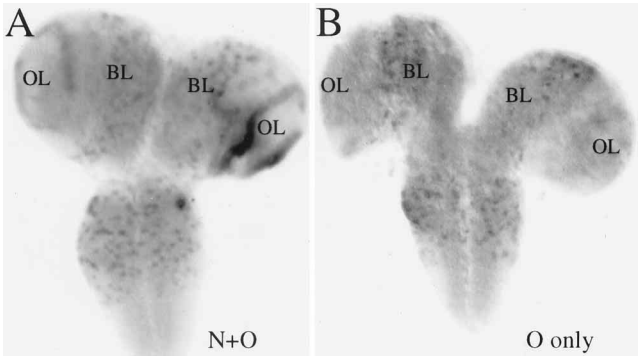

FIG. 3. Cell-specific splicing of *grh* in the CNS. The distribution of *grh* O mRNA in comparison to all *grh* mRNA species in the CNS of late-third-instar larvae was analyzed by in situ hybridization. OL indicates the developing optic lobes (note that at this stage the proliferating cells in the optic lobes are in rings, whereas earlier they appear as clusters [see, e.g., Figures 4 and 7]), and BL marks the brain lobes. (A) A probe detecting all *grh* mRNA splice forms reveals expression in the NBs and the proliferating zones of the optic lobes (OL). (B) In contrast, a probe corresponding to exons 4 and 5, which are exclusive to the O mRNA, reveals expression of the *grh* O mRNA in the NBs only.

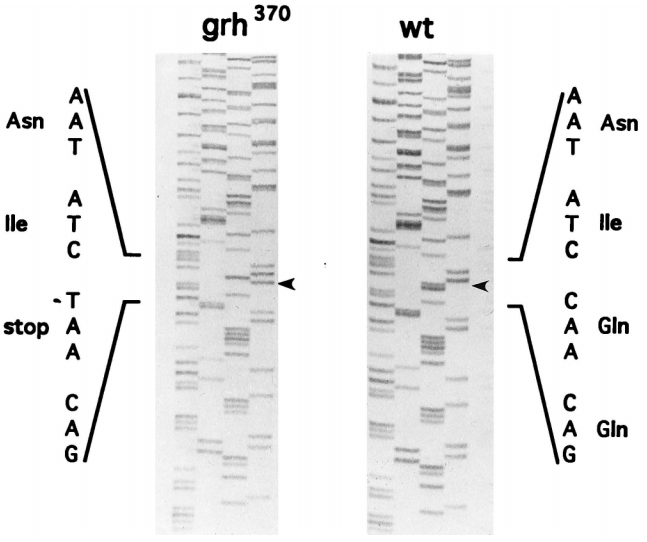

FIG. 5. The *grh*^{370} mutation introduces a stop codon within exon 5. Shown is a section from a sequencing gel revealing the mutation in *grh*^{370}. The bases corresponding to the mutant region are displayed adjacent to the relevant sector, and arrowheads indicate the single bases which are mutated. The sequence of exons 4 and 5 is available in the EMBL database. wt, wild type.

one with a chromosomal deficiency which removes *grh* [*Df(2R)Pcl*^{7B}], most of the mutant larvae die during early third instar (>75%), a few survive to pupation, and 1 to 2% eclose as adults; these die shortly after eclosion and appear to be unable to coordinate their movements, as demonstrated by their persistently shaking appendages. This late lethality suggested to us that *grh*^{370} could be a mutation in a larval- and pupal-specific function, and since the GRH O isoform has a limited distribution, we decided to test whether the *grh*^{370} lesion specifically affects this mRNA. The pattern of GRH expression in the *grh*^{370} mutants is consistent with this hypothesis, since the GRH protein is specifically absent in the NBs (Fig. 4). In other places, such as the epidermis, head skeleton, tracheae, anterior spiracles, discs, and foregut, GRH levels appear normal compared to those in wild-type larvae (data not shown). Furthermore, in the CNS, the protein is still present in the optic lobes even though it is completely absent from the NBs (Fig. 4B). The lack of expression by NBs in *grh*^{370} mutants is not due to the failure of NB to develop, since staining with an antibody against the Asense protein, which marks all the NBs (5), demonstrates that the number and position of the NBs are normal in these mutants (data not shown).

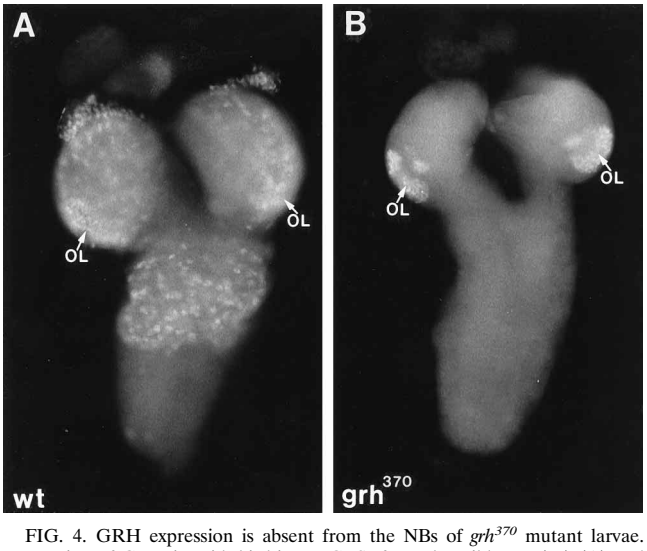

FIG. 4. GRH expression is absent from the NBs of *grh*^{370} mutant larvae. Expression of GRH in mid-third-instar CNSs from the wild type (wt) (A) and *grh*^{370} homozygous mutants (B). GRH expression is clearly absent from the imaginal NBs of *grh*^{370} mutants, although the protein is still present in the optic lobes (OL; arrows).

To test whether the *grh*^{370} mutation specifically affects the O mRNA, we sequenced the O-specific exons, exons 4 and 5 (including the surrounding intron sequences), from the mutant and parental stocks. A genomic fragment spanning exons 4 and 5 was produced from *grh*^{370} larval DNA by PCR, and several clones from independent PCR reactions were sequenced. This analysis revealed a base pair change in exon 5 of *grh*^{370} (Fig. 5) which causes a change in the corresponding codon from a glutamine to a stop codon and confirms that the mutation in *grh*^{370} specifically affects the O form of GRH. The absence of detectable GRH in the NBs of *grh*^{370} larvae also demonstrates that only O-class mRNAs are present in these cells. Translation of the mutant *grh*^{370} O transcripts would produce a protein truncated at the end of exon 5. This protein would lack the identified DNA-binding and dimerization domains of GRH and thus be unlikely to have any residual function.

GRH displays a tissue-specific activity in the CNS which correlates with isoform distribution. To test whether both forms of GRH can activate transcription in vivo, we placed multiple copies of two GRH binding elements (be1 and be2 [Fig. 6D], both derived from the *Ddc* gene [reference 18 and unpublished data]) upstream of a *lacZ* reporter gene and used P-element-mediated transformation to introduce these reporter constructs into the genomes of flies. Both reporter genes are expressed in GRH-expressing cells during embryonic and larval development. However, while expression of reporter be1 is detected in all GRH-expressing cells, reporter be2 is expressed only in the CNS (Fig. 6A and B). Furthermore, when the CNS is dissected from a third-instar larva, reporter be1 activity is detected in the optic lobes, NBs, and tracheael cells whereas reporter be2 is detected only in the NBs (Fig. 7A and B). This restricted expression of reporter be2 occurs in many independent transformant lines, indicating that it reflects a

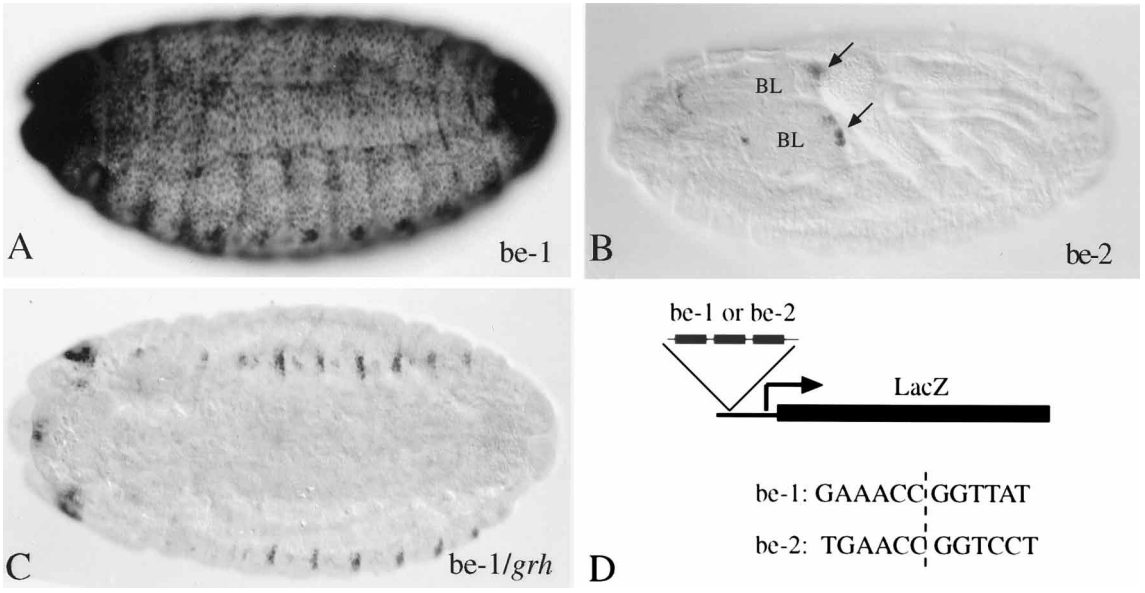

FIG. 6. Differential expression of two GRH-responsive genes (reporters be1 and be2). (A and B) The embryonic expression pattern of each reporter gene (be1 and be2) was detected in early stage 16 embryos with an antibody recognizing β-galactosidase. (A) Expression of reporter be1 can be seen in the epidermis and in the tracheae underneath the epidermis. (B) In contrast to reporter be1, be2 is only expressed in the CNS (arrows mark the NBs in the brain lobes [BL], where the strongest levels are detected at this stage) and is absent from other GRH-expressing tissues. Anterior is to the left. (C) In *grh* mutant embryos (grh^{B37}/grh^{B37}; ventral view of a stage 14 embryo is shown), expression of reporter be1 is greatly reduced, demonstrating that GRH is responsible for most aspects of its expression. Residual expression present in localized patches in the epidermis results from enhancers at the site of insertion of the reporter gene. (D) Schematic illustration of GRH reporter constructs and binding sites (not drawn to scale). The black arrows indicate the transcription start site in the *hsp70* minimal promoter, the black rectangle represents the *lacZ* gene, and the three small stippled boxes represent the GRH binding sites which are inserted upstream of the promoter. The DNA sequences of the two GRH binding elements, be1 and be2, are shown with the stippled line marking the division between the two half-sites in each binding element (dotted line marks the center). Both sequences are derived from the *Ddc* gene, and be2 was previously referred to as element I (18).

property of the regulatory sequence. This restricted expression implies either that it is recognized by a different protein or that it is only recognized by GRH proteins in specific cells.

The expression of both reporters, be1 and be2, is dependent on GRH function, as little or no β-galactosidase is detected when the reporter genes are crossed into mutants in which disruption of all GRH isoforms occurs or into others with deficiencies which remove the entire *grh* gene (Fig. 6C and data not shown). Similarly, in grh^{370} mutant larvae, the expression from the NBs is missing, as expected from the effect of this mutation on GRH; no expression of reporter be2 is detected in grh^{370} larvae (Fig. 7D), and expression of reporter be1 is absent from NBs but persists in the optic lobes and in other tissues (Fig. 7C). Thus, we conclude that GRH can activate the expression of reporter be2 in a tissue-specific manner which correlates with the expression pattern of the O isoform of GRH.

DISCUSSION

The *Drosophila* transcription factor GRH is one of an emerging group of transcription factors which includes human factors that affect the transcription of HIV and other viruses (23, 28, 34). These proteins share a distinct DNA-binding domain, form oligomers, and have been implicated in both the activation (see, e.g., references 12, 23, 27, and 34) and the repression of transcription (16, 22). Our analysis of the genomic structure of *grh* further supports an evolutionary relationship between these proteins, since the exon structure within the DNA-binding domain is largely conserved between the human and *Drosophila* proteins (references 23, 28, and 33 and this work). The extent of the functional similarity between the human and *Drosophila* proteins is less clear. However, it has been suggested that the activity of at least one of the human relatives, CP2, is modulated in a tissue- and stage-specific manner by interactions with other factors (17). Here we suggest that a similar mechanism may be significant in the tissue-specific function of *grh* and that the alternate splicing of the *grh* gene to produce isoforms with different cellular distributions could account for these different cellular functions.

Tissue-specific splicing to modulate transcription factor functions. We have found that there is highly regulated splicing of GRH to produce an isoform containing two extra exons (*grh* exons 4 and 5) as the sole GRH isoform in the NBs. Other examples of alternate splicing modulating the activities of transcription factors fall into two types. The most dramatic effects occur when DNA binding or dimerization is affected, as, for example, with I-POU (32) and ICER (25). These proteins can titrate the active sibling form such that the balance between the isoforms is critical in determining whether active transcription complexes can be formed. In contrast, the alternate splicing of homeotic genes such as *Ultrabithorax* (*Ubx*) results in isoforms with small differences in the amino acid sequence outside the critical homeodomain (20, 26). Careful analysis of the *Ubx* isoforms indicated that they have different functions in vivo (29), although the proteins behave identically in vitro and, when expressed at high levels, they largely have similar effects (8). These subtle differences in activity more closely resemble the effects on GRH of alternative splicing, as we do not detect any difference in the activities of the isoforms in vitro and the alternate exons are inserted adjacent to the DNA-binding and dimerization domains. However, both the conservation of the alternate exons in a distantly related species, *D. virilis*, and the fact that a nonsense mutation which abolishes the function of the NB GRH isoform is lethal suggest that alternate splicing

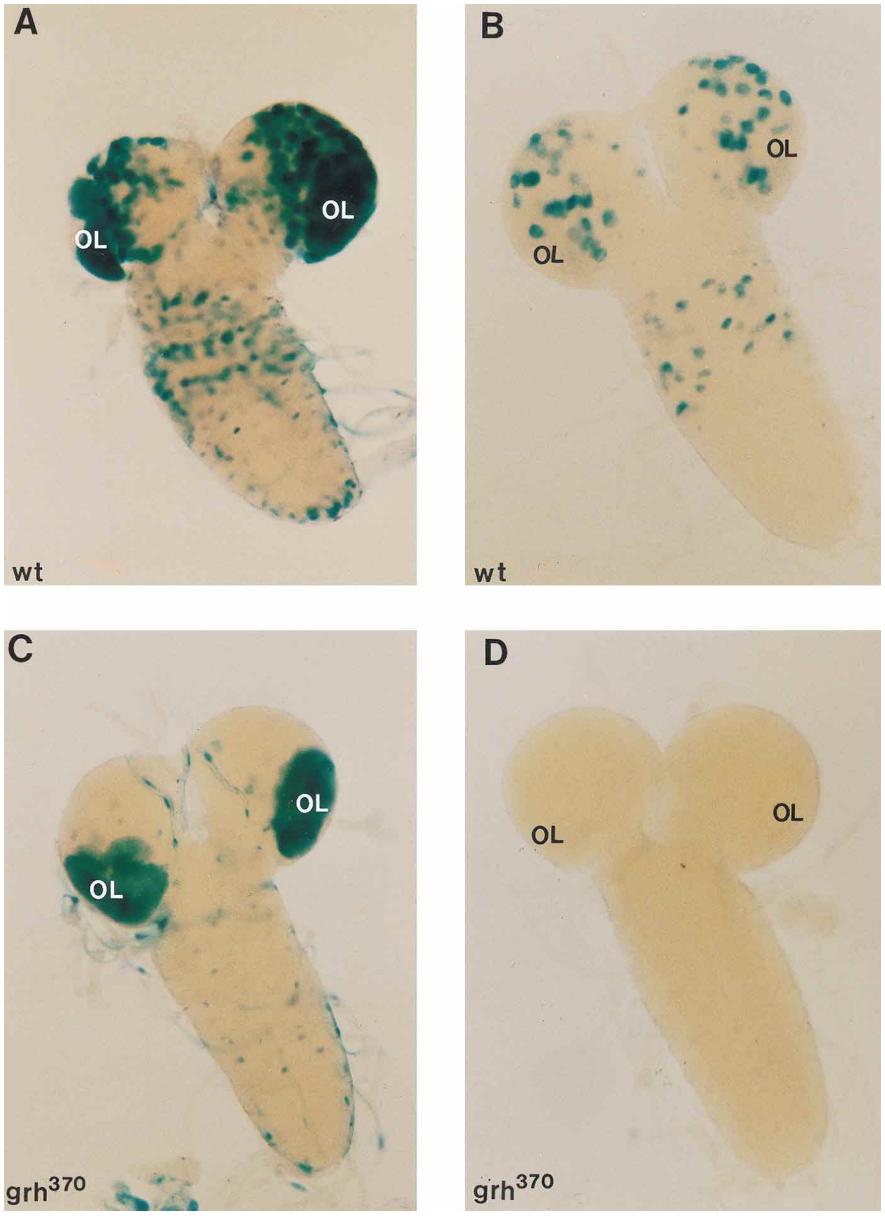

FIG. 7. Reporter gene expression in NBs is lost in *grh*^{370} mutants. (A and B) The differential expression pattern of the reporter genes (be1 and be2) is maintained in the larval CNS (detected here in CNS from mid-third instar by histochemical staining with X-Gal). Reporter be1 (A) is detected in the NBs (large cells), the optic lobes (OL), and tracheael cells, whereas reporter be2 (B) is expressed only in the NBs. wt, wild type. In *grh*^{370} mutant larvae, NB expression is absent from both reporter be1 (C) and reporter be2 (D). With reporter be1 (C), expression is still detectable in the optic lobes and in tracheal cells, although it is absent from the NBs.

does confer some specific function to GRH that may be relevant to the cellular context in the nervous system.

The expression of two GRH reporter genes in vivo also suggests that the two GRH isoforms have different actions. Reporter genes containing palindromic (be1) and nonpalindromic (be2) GRH binding sites have different profiles of activation in vivo, and that of reporter be2 correlates specifically with the distribution of the GRH O isoform. This restriction of reporter be2 expression to the NBs is not a threshold effect, since GRH protein levels are lower in the CNS than in epidermal cells (6) (data not shown) and reporter be1 expression (detected in all GRH-expressing cells) is weaker in the NBs than in the other GRH-expressing tissues. The NB (O) isoform of GRH does not demonstrate significantly better binding to the be2 site; thus, a different mechanism for the response must be invoked. We cannot rule out the possibility that it is a consequence of the tissue-specific distribution of other factors (e.g., an inhibitor present in all tissues except the NBs) rather than the alternate splicing of GRH per se (Fig. 8). However, alternate splicing could modulate the binding affinity of GRH indirectly, and one explanation for the observed in vivo selectivity is that the O isoform can form a heterodimer with another transcription factor that results in high-affinity be2 binding (Fig. 8). The regulation of globin genes by CP2 appears to require it to form a complex with a second, tissue-specific factor, which is as yet not cloned but which confers tissue and stage specificity on CP2 function (17). Such partners could include other GRH-like proteins, since the two different

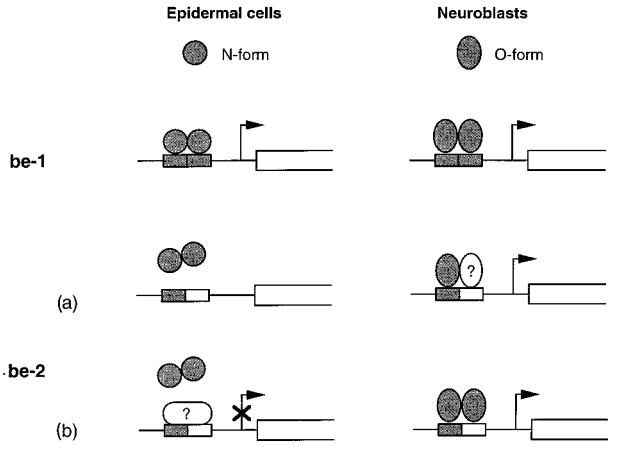

FIG. 8. Scheme illustrating differential activation of GRH targets in epidermis and NBs. The left panel illustrates interactions between GRH and the target sites be1 and be2 in epidermal cells, where the N form of GRH (grey circles) is expressed. The right panel illustrates interactions between GRH and its targets in the NBs, where the O isoform (grey ellipses) is expressed. Activation of the reporter gene is indicated by an arrow, while repression of transcription is symbolized by an X on the arrow. (Left panel) In epidermal cells, the N form of GRH (grey circles) activates transcription from be1, which consists of two optimal GRH half-sites (grey boxes). However, be2, which contains one optimal and one suboptimal GRH half-site, does not promote activation, either because the N-form homodimer does not have a high enough affinity to occupy this site (a) or because this binding site is blocked by an unknown repressor (rounded rectangle) in epidermal cells (b). (Right panel) The O isoform of GRH (grey ellipse) in the NBs is able to activate transcription from be1, through binding in a homodimer complex. The O isoform can also activate transcription from be2 either because it forms a heterodimeric complex with another partner (open ellipse) acquiring high affinity for the be2 site (a) or because the absence of an epidermal repressor allows access of the O-isoform homodimers (b).

human GRH-related proteins, CP2 and LBP-1a, have been shown to form heterodimers on the HIV promoter (34). The recognition sequences for CP2-LBP dimers encompass both palindromic and direct repeats (28, 33), suggesting that, as with the retinoic acid receptor family, there is considerable potential for variation in the target sequences depending on the combination of partners and the promoter context. Thus, we propose that be1 and be2 represent targets for homo- and heterodimers of GRH, respectively, and that the O isoform preferentially forms heterodimers (Fig. 8).

The tissue-specific distribution of GRH mRNAs implies that the splicing mechanism of NBs is different from that of the other tissues. In spite of the complex exon structure of GRH and the fact that CP2 and LBP-1 exhibit splicing variants that alter the DNA binding and dimerization capabilities directly, the only regulated splicing detected in our analysis affects exons 4 and 5. There appears to be a single target for this regulation, since both exons are always present together. This could involve either positive effectors enhancing the use of the 3′ acceptor site at the start of exon 4 in neuroblasts, as seen with the neural-specific exon in the *myc* transcript (24) and the female-specific exon in the *Drosophila doublesex* gene (31), or negative factors suppressing the exon 4 acceptor in other tissues. If the former is correct, we would expect to find a neural-specific factor(s) which promotes this splicing event and which would provide a mechanism for coordinately modifying the activity of a selective array of transcription factors.

GRH activates transcription in vivo. GRH binds DNA and activates transcription from its cognate sites in vitro. Our data confirm that GRH acts as a transcriptional activator in the developing organism. Embryos carrying a reporter gene containing multiple GRH be1 sites express β-galactosidase in a pattern which resembles that of GRH expression itself. This expression is greatly reduced by mutations in *grh*, confirming that it is dependent on GRH function and thus demonstrating that GRH has the capacity to activate transcription in at least the majority of cells in which it is expressed. This is consistent with observations from in vitro experiments in which GRH (also called NTF-1) interacted directly with one of the proteins present in the TFIID complex and activates transcription from a variety of promoters (11, 12). However, these positive effects on transcription contrast with the function reported for GRH during early embryogenesis, at which time it appears to act as a repressor (16, 22). The effects of grh^{370} mutations on the activity of the reporter genes in the CNS argue that both splice forms of GRH function as activators in vivo; thus, alternate splicing cannot account for the conversion of GRH between activator and repressor of transcription. In vitro, GRH is phosphorylated by MAP kinase (22), indicating that posttranscriptional modifications may be significant in altering GRH activity in addition to the alternate splicing of mRNA described here.

ACKNOWLEDGMENTS

We thank members of our lab, in particular Chris Thompson, for helpful discussions and encouragement. We also thank Sarah Lazar for carrying out the DNA sequencing of exons 4 and 5, Rick Tearle for giving us the grh^{370} mutant flies, Jesus de Celis for some of the pictures of GRH staining in wild-type tissues, and Nick Brown and Jose de Celis for helpful comments on the manuscript.

This work was supported by an MRC project grant, and Anne Uv was the recipient of a fellowship from N.A.V.F.

REFERENCES

1. **Anderson, S. O.** 1990. Sclerotization of insect cuticle, p. 133–155. *In* E. Ohnishi and H. Ishizaki (ed.), Molting and metamorphosis. Japan Scientific Societies Press, Tokyo.
2. **Ashburner, M.** 1989. *Drosophila*: a laboratory manual. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y.
3. **Attardi, L. D., and R. Tjian.** 1993. *Drosophila* tissue-specific transcription factor NTF-1 contains a novel isoleucine-rich activation domain. Genes Dev. **7:**1341–1353.
4. **Biggin, M. D., and R. Tjian.** 1988. Transcription factors that activate the Ultrabithorax promoter in developmentally staged extracts. Cell **53:**699–711.
5. **Brand, M., A. P. Jarman, L. Y. Jan, and Y. N. Jan.** 1993. Asense is a *Drosophila* neural precursor gene and is capable of initiating sense organ formation. Development **119:**1–17.
6. **Bray, S. J., B. Burke, N. B. Brown, and J. Hirsh.** 1989. Embryonic expression pattern of a family of *Drosophila* proteins that interact with a central nervous system regulatory element. Genes Dev. **3:**1130–1145.
7. **Bray, S. J., and F. C. Kafatos.** 1991. Developmental function of Elf-1: an essential transcription factor during embryogenesis in *Drosophila*. Genes Dev. **5:**1672–1683.
8. **Busturia, A., I. Vernos, A. Macias, J. Casanova, and G. Morata.** 1990. Different forms of Ultrabithorax proteins generated by alternative splicing are functionally equivalent. EMBO J. **9:**3551–3555.
9. **Church, G. M., and W. Gilbert.** 1984. Genomic sequencing. Proc. Natl. Acad. Sci. USA **81:**1991–1995.
10. **Duncan, I.** 1986. Control of Bithorax complex functions by the segmentation gene *fushi terazu*. Cell **47:**297–309.
11. **Dynlacht, B., L. Attardi, A. Admon, M. Freeman, and R. Tjian.** 1989. Functional analysis of NTF-1, a developmentally regulated *Drosophila* transcription factor that bind neuronal *cis*-elements. Genes Dev. **3:**1677–1688.
12. **Dynlacht, B. D., T. Hoey, and R. Tjian.** 1991. Isolation of coactivators associated with the TATA-binding protein that mediate transcriptional activation. Cell **66:**563–576.
13. **Glaser, R. L., M. F. Wolfner, and J. T. Lis.** 1986. Spacial and temporal pattern of hsp 26 expression during normal development. EMBO J. **5:**747–754.
14. **Henikoff, S.** 1984. Unidirectional digestion with exonuclease III creates targeted breakpoints for DNA sequencing. Gene **28:**351–359.
15. **Hiromi, Y., and W. J. Gehring.** 1987. Regulation and function of the *Drosophila* segmentation gene *fushi tarazu*. Cell **50:**963–974.
16. **Huang, J. D., T. Dubnicoff, G. J. Liaw, Y. Bai, S. A. Valentine, J. M. Shirokawa, J. A. Lengyel, and A. J. Courey.** 1995. Binding sites for transcription factor NTF-1/Elf-1 contribute to the ventral repression of *decapentaplegic*. Genes Dev. **9:**3177–3189.
17. **Jane, S. M., A. W. Neinhuis, and J. M. Cunningham.** 1995. Hemoglobin

switching in man and chicken is mediated by a heteromeric complex between the ubiquitous transcription factor CP2 and a developmentally specific protein. EMBO J. **14:**97–105.
18. **Johnson, W. A., C. A. McCormick, S. J. Bray, and J. Hirsh.** 1989. A neuronal specific enhancer of the *Drosophila* dopa decarboxylase gene. Genes Dev. **3:**676–686.
19. **Konrad, K. D., and J. L. Marsh.** 1987. Developmental expression and spatial distribution of dopa decarboxylase in *Drosophila*. Dev. Biol. **122:**172–185.
20. **Kornfeld, K., R. Saint, P. Beachy, P. Harte, D. Peattie, and D. Hogness.** 1989. Structure and expression of a family of *Ultrabithorax* mRNAs generated by alternative splicing and polyadenylation in *Drosophila*. Genes Dev. **3:**243–258.
21. **Kramer, K. J., and T. L. Hopkins.** 1987. Tyrosine metabolism for insect cuticle tanning. Arch. Insect Biochem. Physiol. **6:**279–301.
22. **Liaw, G. J., K. M. Rudolph, J. D. Huang, T. Dubnicoff, A. J. Courey, and J. A. Lengyel.** 1995. The *torso* response element binds GAGA and NTF-1/Elf-1, and regulates *tailless* by relief of repression. Genes Dev. **9:**3163–3176.
23. **Lim, L. C., S. L. Swendeman, and M. Sheffery.** 1992. Molecular cloning of the α-globin transcription factor CP2. Mol. Cell. Biol. **12:**828–835.
23a. **Lindsley, D. L., and G. G. Zimm.** 1992. The genome of *Drosophila melanogaster*. Academic Press, Inc., San Diego, Calif.
24. **Min, H., R. C. Chan, and D. L. Black.** 1995. The generally expressed hnRNP is involved in a neural-specific pre-mRNA splicing event. Genes Dev. **9:** 2659–2671.
25. **Molina, C. A., N. S. Foulkes, E. Lalli, and P. Sassone Corsi.** 1993. Inducibility and negative autoregulation of CREM: an alternative promoter directs the expression of ICER, an early response repressor. Cell **75:**875–886.
26. **O'Connor, M. B., R. Binari, L. A. Perkins, and W. Bender.** 1988. Alternative products from the Ultrabithorax domain of the bithorax complex. EMBO J. **7:**435–445.
27. **Parada, C. A., J. B. Yoon, and R. G. Roeder.** 1995. A novel LBP-1 mediated restriction of HIV-1 transcription at the level of elongation in vitro. J. Biol. Chem. **270:**2274–2283.
28. **Shirra, M. K., Q. Zhu, H.-C. Huang, D. Pallas, and U. Hansen.** 1994. One exon of the human LSF gene includes conserved regions involved in novel DNA-binding and dimerization motifs. Mol. Cell. Biol. **14:**5076–5087.
29. **Subramaniam, V., H. M. Bomze, and A. J. Lopez.** 1994. Functional differences between Ultrabithorax protein isoforms in *Drosophila melanogaster*: evidence from elimination, substitution and ectopic expression of specific isoforms. Genetics **136:**979–991.
30. **Tautz, D., and C. Pfeifle.** 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.
31. **Tian, M., and T. Maniatis.** 1993. A splicing enhancer complex controls alternative splicing of *doublesex* pre-mRNA. Cell **74:**105–114.
31a. **Thompson, C., T. A. Weaver, and S. J. Bray.** Unpublished data.
32. **Treacy, M. N., L. I. Neilson, E. E. Turner, X. He, and M. G. Rosenfeld.** 1992. Twin of I-POU: a two amino acid difference in the I-POU homeodomain distinguishes an activator from an inhibitor of transcription. Cell **68:**491–505.
33. **Uv, A. E., C. R. L. Thompson, and S. J. Bray.** 1994. The *Drosophila* tissue-specific factor Grainyhead contains novel DNA-binding and dimerization domains which are conserved in the human protein CP2. Mol. Cell. Biol. **14:**4020–4031.
34. **Yoon, J.-B., G. Li, and R. G. Roeder.** 1994. Characterization of a family of related cellular transcription factors which can modulate human immunodeficiency virus type 1 transcription in vitro. Mol. Cell. Biol. **14:**1776–1785.
35. **Zhong, F. M., S. L. Swendeman, M. Popik, P. M. Pitha, and M. Sheffery.** 1994. Evidence that levels of the dimeric cellular transcription factor CP2 play little role in the activation of the HIV-1 long terminal repeat in vivo or following superinfection with herpes-simplex virus type-1. J. Bio. Chem. **269:**21269–21276.