A Novel Basic Helix-Loop-Helix Protein Is Expressed in Muscle Attachment Sites of the *Drosophila* Epidermis

PHILIPPE ARMAND, ANITA C. KNAPP, ALEC J. HIRSCH, ERIC F. WIESCHAUS,
AND MICHAEL D. COLE*

Department of Molecular Biology, Princeton University, Princeton, New Jersey 08544-1014

Received 16 August 1993/Returned for modification 28 September 1993/Accepted 25 February 1994

We have found that a novel basic helix-loop-helix (bHLH) protein is expressed almost exclusively in the epidermal attachment sites for the somatic muscles of *Drosophila melanogaster*. A *Drosophila* cDNA library was screened with radioactively labeled E12 protein, which can dimerize with many HLH proteins. One clone that emerged from this screen encoded a previously unknown protein of 360 amino acids, named *delilah*, that contains both basic and HLH domains, similar to a group of cellular transcription factors implicated in cell type determination. *Delilah* protein formed heterodimers with E12 that bind to the muscle creatine kinase promoter. In situ hybridization with the *delilah* cDNA localized the expression of the gene to a subset of cells in the epidermis which form a distinct pattern involving both the segmental boundaries and intrasegmental clusters. This pattern was coincident with the known sites of attachment of the somatic muscles to tendon cells in the epidermis. *delilah* expression persists in *snail* mutant embryos which lack mesoderm, indicating that expression of the gene was not induced by attachment of the underlying muscles. The similarity of this gene to other bHLH genes suggests that it plays an important role in the differentiation of epidermal cells into muscle attachment sites.

The basic helix-loop-helix (bHLH) family of proteins was initially discovered through genetic studies in *Drosophila melanogaster* and mice in which genes linked to the determination of specific cell types (neuronal and muscle) were found to encode a common motif that was conserved in each of the proteins, as well as more distantly related genes such as the *c-myc* oncogene (19, 40). The function of this motif (and its name) became clear when a DNA-binding protein implicated in B-cell gene expression was found to share the same conserved sequences (29). bHLH proteins have been linked to diverse biological processes such as neurogenesis (reviewed in reference 9), myogenesis (reviewed in reference 33), dosage compensation (14), segmentation (35), transcriptional regulation (34, 36), oncogenesis (2, 13, 27), and mesoderm formation (38). The majority of these studies have identified bHLH genes through genetic criteria, in which the gene responsible for a mutant phenotype was cloned and the sequence was subsequently found to encode an HLH motif. Other bHLH genes were cloned on the basis of DNA-binding activity or by homology to other family members (3, 29, 33). The striking frequency at which developmentally important genes have been found to encode a bHLH DNA-binding domain suggests that the majority of these genes will control essential cellular and developmental processes.

The HLH family can be divided into at least two groups, those with a closely linked basic region that bind DNA and those which lack a basic region and seem to serve as inhibitors of the former class. The HLH domain itself mediates protein dimerization, whereas the basic region is required for DNA binding (17, 41). However, the basic region only functions in dimerized proteins and hence is dependent on the HLH domain for activity. The HLH domain dictates whether a protein can dimerize with itself or whether it requires a heterologous partner, whereas the basic region appears to determine the sequence specificity of DNA binding. However,

subtle mutations in the basic region that do not alter the DNA recognition specificity can dramatically alter biological activity (18), suggesting that the basic region may also interact with other cellular components such as other transcription factors. Genes which lack a basic region, such as Id (3) and *emc* (20, 21), suppress DNA binding when heterodimerized with proteins with an intact basic domain (39, 44).

In this study, we wished to explore whether there are new members of the bHLH family that function in the differentiation of specific cell types in *D. melanogaster*. We employed a sensitive screening approach by using radioactive protein to select cDNAs from an expression library which encode proteins that form stable heterodimers with the mammalian bHLH protein E12. The resulting novel bHLH gene was found to be expressed only in epidermal cells that function as attachment sites for the somatic muscles. These findings suggest that protein screening can be used to identify numerous bHLH genes in diverse organisms and cell types, taking advantage of the fact that the proteins readily heterodimerize in stable and specific complexes.

MATERIALS AND METHODS

Cloning of bHLH proteins by protein screening. The mammalian E12 protein (29) was used as a probe for screening of a *Drosophila* cDNA expression library (a generous gift of K. Zinn). To provide a high-specific-activity labeling site, a 6xHis expression vector (QE-9; Qiagen) was modified with a synthetic oligonucleotide which encoded the amino acids RRASV (6, 25). A segment of E12 extending from an internal *NotI* site to the C terminus was cloned into the vector, and protein was isolated on a nickel chelate column. E12 protein was labeled with protein kinase from heart muscle (Sigma) and [γ - 32 P]ATP and was trichloroacetic acid precipitated to remove unincorporated ATP, washed in acetone, and resuspended in 6 M guanidinium HCl–50 mM NaPO₄ (pH 8). The denatured, labeled protein was then diluted into binding buffer and incubated at 4°C for 8 h with filter lifts of the λ gt11 library. Filters were washed four times in TNET (500 ml each wash) at

* Corresponding author. Electronic mail address: mcole@molbiol.princeton.edu.

4°C and autoradiographed at -80°C for 4 to 8 h. Further details will be presented elsewhere (23a).

In vitro transcription and translation. The MyoD1 (pVZC13b) and E12 cDNA clones used for in vitro transcription were gifts from Andrew Lassar and Cornelis Murre, respectively, and have been described previously (19, 29). The *delilah* clone contained a full-length cDNA cloned into the *Eco*RI site of pBluescript (Stratagene). *delilah*-encoding RNA was synthesized in a 100- μ l reaction volume by using 1 μ g of linearized DNA as a substrate. The reaction was performed at 37°C for 1 h in the presence of T7 RNA polymerase (Gibco BRL) and RNasin (Promega). After RqDNAseI (Promega) and proteinase K digestion, the RNA was extracted with phenol-chloroform, precipitated, and resuspended in 50 μ l of water. One microliter of this suspension was used for in vitro translation with rabbit reticulocyte lysate in a 30- μ l reaction volume in the presence of [³⁵S]methionine. The reactions were allowed to proceed for 1 h at 30°C. To evaluate proper translation, sodium dodecyl sulfate-polyacrylamide gel electrophoresis was performed by standard methods. For electrophoretic mobility shift assay experiments, 2 μ l of RNA-programmed lysate, representing approximately equimolar amounts of each protein, was used.

Bacterial expression and DNA-binding assays. To produce Delilah protein in bacteria, the PCR-amplified *delilah* cDNA was cloned into the His6-HMK vector described above. The proper reading frame was confirmed by sequencing. The histidine fusion proteins were produced by using a nickel chelate affinity column as described by the manufacturer (Qiagen).

The interaction of Delilah with different bHLH proteins was examined by mixing together approximately 50 ng of bacterially expressed proteins in the presence of 0.5% dry milk. Wherever in vitro-translated proteins were used, dry milk was excluded from the reaction mixtures. The mixture first was incubated at 37°C for 10 min to allow dimer exchange. The reaction mixtures were then incubated at 4°C for 10 min after the addition of 1 ng of ³²P-labeled MCK probe (for buffer conditions, see reference 4). The sequence of the MCK DNA oligomer has been previously described (17). The MCK oligonucleotide was labeled by filling 5' overhangs with the Klenow fragment of DNA polymerase I (New England Biolabs) and [α -³²P]dCTP (Amersham) together with nonradioactive dATP, dGTP, and dTTP. Protein-DNA complexes were resolved on 6% acrylamide gels run at 4°C in 0.25 $\times$ TBE (1 $\times$ TBE is 90 mM Tris-borate and 2 mM EDTA, pH 8.0). The gel was fixed twice in 20% isopropanol-7% acetic acid and, after a drying, exposed to two sheets of X-ray film with an intensifying screen. Only the film near the screen is shown.

In situ hybridization. In situ hybridization to *Drosophila* embryos utilized the 1.3-kb cDNA fragment that had been digoxigenin labeled by established techniques (37a). Embryos were staged by morphology and by using timed collections. The micrograph of muscles was from larval preparations previously described (37). In situ hybridization to *Drosophila* polytene chromosomes maps the *delilah* gene to band 97B on the right arm of the third chromosome (data not shown).

RESULTS

Cloning of a novel bHLH cDNA. Since members of the bHLH family of proteins have been implicated in many stages of development and differentiation in *D. melanogaster*, we were interested in exploring whether there were additional members of the family that may play important roles in specific cell lineages. Two approaches to the identification of novel HLH

proteins have been utilized successfully. The simplest is to use the highly conserved residues in the two helical regions of the HLH domain to design hybridization or PCR primers that would bind to new genes and facilitate cloning, an approach that led to the identification of the Id gene (3) and new members of the myogenic family (33). However, there are very few contiguous residues that are conserved among the different bHLH genes, making the choice of primers with minimal redundancy very difficult. A second approach takes advantage of the requirement that all bHLH proteins dimerize to form a functional DNA-binding complex. If one protein is radioactively labeled, it can serve as a probe to screen bacteriophage cDNA expression libraries for other proteins that form stable dimers and hence allow direct cloning without prior knowledge of the sequence. To date, this approach has proven successful with proteins containing the leucine zipper dimerization domain or a combination bHLH-leucine zipper (5, 6). The latter approach seemed the most powerful for cloning novel HLH proteins, so we devised a simple and sensitive protocol with the mammalian bHLH proteins E12 and *myoD* as a preliminary test (see Materials and Methods).

The human E12 protein can heterodimerize with both mammalian and *Drosophila* HLH proteins (30), so E12 was used as a probe to screen a 9- to 12-h *Drosophila* embryo cDNA library. One positive clone was obtained from a screen of 2×10^5 phage. The subcloned 1.3-kb insert was sequenced, and the sequence was used to search the GenBank data base (Fig. 1). The cDNA encoded a single long open reading frame of 360 amino acids which initiated at a consensus *Drosophila* ATG (12). The ATG was preceded by an in-frame stop codon, suggesting that the entire protein coding region is present within the cDNA. On the basis of the expression pattern described below, this gene was named *delilah*. As anticipated from the screening protocol, the region of the predicted protein between amino acids 95 and 153 could be readily aligned with the bHLH motif found in all members of the HLH family (Fig. 2), and a computer search with this region selected other HLH proteins. No other region of the *delilah* protein had significant homology to proteins in GenBank.

The sequence of *delilah* was used in pairwise comparisons with other HLH proteins to determine if any features suggested particularly close relationships. Within the basic region, the arginines, lysines, and central asparagine were conserved at positions that were common to most bHLH proteins, but the sequence itself did not match with that of any other protein. Significantly, several positions differ between *delilah* and the basic region consensus for the *myoD*-related subgroup of HLH proteins, including threonine and methionine residues that are essential for myogenic activity (18, 43). Since the basic region is thought to be the major determinant of DNA recognition specificity, it is likely that Delilah protein will exhibit optimal binding to an E-box-related sequence (CANNTG) distinct from other HLH proteins. Within the HLH protein dimerization domain itself, the sequence of *delilah* was equally divergent from other members of the family. The only noteworthy feature was the length of the so-called loop region, which was longer than that of most HLH proteins (Fig. 2). The only HLH members with a longer loop region are the *achaete-scute* genes from *D. melanogaster* (40), but the rat homologs of *achaete-scute* do not retain this feature (23). Thus, the bHLH domain of *delilah* predicts that the protein will exhibit unique protein-DNA and protein-protein interactions that distinguish its function from that of other HLH family members.

Delilah protein binds to DNA as a heterodimer. The presence of a basic region preceding the HLH domain indicated that Delilah would bind to DNA, but it was of interest to

FIG. 1. Nucleotide sequence and translation of *delilah* cDNA. The first ATG is consistent with *Drosophila* translation initiation sites (12). There is a termination codon at nucleotide 16 which is in the same reading frame as the ATG, suggesting that the coding region is complete. The nucleotide sequence is numbered to the right, and the amino acid sequence is numbered on the left.

determine if the protein could homodimerize like E12 or required a partner. Delilah protein was produced both by in vitro translation and by cloning into a bacterial expression vector. In vitro translation yielded a 40-kDa polypeptide after electrophoresis on a denaturing gel (not shown), in agreement with the predicted open reading frame. We also produced Delilah protein by using a bacterial expression vector. However, there were numerous bands ranging in molecular mass from 18 to 40 kDa in the final preparations, suggesting that the protein is being rapidly degraded when produced at a high concentration. Nevertheless, functional protein was present in these preparations as shown below.

Since the basic region of *delilah* was similar, although not identical, to *myoD*, we tested for DNA-binding activity with an oligonucleotide from the MCK promoter. For comparison, we also used both in vitro-translated and bacterially produced E12 and in vitro-translated myoD. Somewhat surprisingly, under

conditions that demonstrated easily detectable DNA-binding activity for E12 homodimers (Fig. 3, lane 16) and very strong binding by myoD-E12 heterodimers (lane 18), we observed no DNA-binding complexes formed with in vitro-translated Delilah either alone (lane 14) or in combination with E12 (lane 15). Furthermore, the intensity of the E12 homodimer band was undiminished when coincubated with Delilah (compare lanes 15 and 16), suggesting that none of the E12 homodimers was being recruited into heterodimers that lacked DNA-binding activity. A similar coincubation of E12 with MyoD gave complete conversion of the E12 homodimers into heterodimers with strong DNA-binding activity (lane 18). For all of the in vitro-translated proteins, the total amount of each protein produced was comparable as judged by [³⁵S]methionine incorporation (not shown). In contrast to the results with in vitro-translated Delilah, the addition of bacterially produced Delilah protein to the in vitro-translated E12 protein produced

	basic	Helix I	loop	Helix II
Delilah	YRRKTANARETRMREINTAFETLRHCVPEAIKGEDAANT.....	NEKLTKITTLRLAMKYITMLTDS		
da	ERRQANNARERIRIRDINEALKELGRMCMTLKSDEP.....	QTKLGILNMAVEVIMTLEQQ		
nau	DRRKAATMRERRRLRKVNEAFEILKRRSTSNP.....	NQRLPKVEILRNAIEYIESLEDL		
mouse MyoD	DRRKAATMRERRRLSKVNEAFEILKRRSTSNP.....	NQRLPKVEILRNAIRYIEGLQAL		
E12	ERRVANNARERLVRDINEAFKELGRMCQLHLNSEKP.....	QTKLLILHQAVSVILNLEQQ		
twist	NQRVMANVRERQRTQSLNDAFKSLQQTITL.....	PSDKLSKIQTLLKATRYIDFLCRM		
As-T5	IRR...NARERNRVKQVNNGFSQLRQHIPAAVIADLSNGRRGIGPGANKKLSKVSTLKMAYVIRRLQKV			

FIG. 2. Amino acid sequence alignment of the *delilah* basic and HLH domains with *daughterless* (da) (11), *nautilus* (nau) (28), mouse Myo D1 (19), E12 (29), *twist* (38), and *achaete-scute* T5 (As-T5) (40).

FIG. 3. DNA-binding activity of Delilah dimerized with E12. DNA-binding activity was monitored by electrophoretic mobility shift assay, using the MCK E box as a probe. The proteins utilized are indicated above each lane as follows: bE, bacterial E12; bD, bacterial Delilah; iE, in vitro-translated E12; iD, in vitro-translated Delilah; iM, in vitro-translated MyoD. In lanes 1 and 3 to 10, constant amounts of bE were analyzed by using the MCK oligomer in a mobility shift assay. When mixed together with bD (increasing amounts of bD in lanes 3 to 7), a new specific complex formation (indicated by arrow b) is obtained. bD alone shows undetectable binding activity (lane 2). Competition was performed with the Delilah-E12 heterodimer with increasing amounts of the MCK oligomer at a 40-fold (lane 5), 80-fold (lane 6), and 120-fold (lane 7) molar excess. In lanes 8 to 13, the indicated in vitro-translated proteins (iE, iD, and iM) were premixed with bE (lanes 8 to 10) or bD (lanes 11 to 13). iD was also analyzed alone (lane 14) or together with E12 (lane 15) or MyoD (lane 19). E12 (iE) was also analyzed for DNA binding by itself (lane 16) or together with Myo D (lane 18). A specific Delilah-E12 DNA-binding complex was evident in lane 11 (arrow a). In vitro-translated MyoD alone exhibits weak DNA-binding activity (lane 17). Note that the autoradiograph only represents the ^{32}P -labeled protein-DNA complex and not [^{35}S]methionine-labeled proteins, and the free DNA has run off the gel.

a new DNA-binding complex consistent with heterodimer formation (Fig. 3, lane 11), and the E12 homodimer band is largely eliminated (compare lanes 11 and 16). These results suggest that heterodimer formation between Delilah and E12 is inefficient and heterodimers only form with the higher concentration of Delilah afforded by the bacterial protein. It is also possible that the larger Delilah protein produced by in vitro translation dimerizes less effectively, although this seems less likely.

Heterodimer formation between Delilah and E12 was also demonstrated by using both proteins produced in bacteria. In this case, the heterodimers migrated faster than the small E12 homodimers (Fig. 3, lanes 3 and 4), almost certainly because of the degradation evident in the Delilah protein preparation. This experiment was instructive because large amounts of each protein could be added to the binding reaction mixture. No DNA-binding activity was observed with Delilah protein alone (lane 2), regardless of concentration, indicating that Delilah does not homodimerize. Furthermore, the DNA-binding activity of the heterodimer was readily inhibited by cold oligonucleotide (lanes 5 to 7). Thus, the DNA-binding activity of Delilah requires dimerization with an E12-like protein, similar to most members of the HLH family.

delilah is expressed in muscle attachment cells. To gain

some insight into the function of *delilah*, it was of interest to determine the cell type and developmental timing of *delilah* expression. Embryos at different developmental stages were stained by in situ hybridization to the *delilah* cDNA to localize the cells with highest levels of expression. Early embryos (0 to 6 h) exhibited little staining, indicating that *delilah* is likely to be zygotically expressed rather than maternally supplied. Late-stage embryos (12 to 24 h) exhibited a very reproducible pattern of epidermal cell staining (Fig. 4). In the thoracic and abdominal segments, the majority of the expressing cells were localized to the segmental boundaries. Staining of cells along the boundaries was not uniform, however, and the heaviest staining occurred in discontinuous stripes along each furrow. A higher magnification of the pattern illustrates that the stripes of staining are mainly localized to a single row of cells (Fig. 5E). Furthermore, although the most predominant staining occurs at the segment boundary, there were also clusters and individual cells within each segment that stained reproducibly. The pattern of staining suggested that *delilah* expression was localized to a specific subset of epidermal cells.

After considering the possible functions of different epidermal cells, we noted a striking similarity between the pattern of *delilah* expression and the sites to which the underlying muscles attach. Muscle syncytia form from the fusion of mesoderm-

FIG. 4. Embryonic staining pattern of *delilah* expression. Lateral view of a late-stage embryo (14 to 18 h) hybridized with the *delilah* cDNA. The embryo is oriented, with the anterior end to the right and the dorsal side up. Prominent staining of the pharyngeal muscle attachment sites (p) is indicated. The arrow indicates a *delilah*-expressing epidermal cell along the dorsal midline that is anterior to the segment boundary.

derived cells beginning approximately 8 h into embryonic development (1). The muscle fibers then attach to the overlying epidermis at characteristic positions, and these sites of attachment are often referred to as tendon cells. Consistent with this assignment of *delilah* expression to muscle attachment sites, we also found prominent staining of the mouthparts, in bands of cells that correspond to the positions of pharyngeal muscle attachment (Fig. 4). The muscles and attachment sites of the mouthparts share a common embryological origin with the epidermal tendon cells.

To illustrate the correlation between *delilah* expression and the underlying muscles, we photographed embryos stained with the *delilah* probe and compared the pattern with a schematic diagram of the expected muscle attachment sites (Fig. 5). A diagram of the muscles in the abdominal segments (Fig. 5A) was prepared from previous descriptions (1, 16, 37), using the numbering system for individual muscles of Bate (1). Since there was some variation in previous descriptions of the points at which muscles attach, we confirmed the positions shown in Fig. 5A by examining phalloidin-stained larvae (not shown). The most common orientation for embryonic and larval muscles is along the horizontal axis of the embryo in three major bands labeled I, II, and III. There is a slight anterior shift of the endpoints of group III near the lateral midline. Most of the longitudinal muscles attach to the epidermis at the segmental boundaries, but some muscles (such as muscles 19 and 20) have one end at the boundary and the other at a position within the segment. To facilitate discussion, we have labeled the muscle attachment sites either anterior (a) or posterior (p) for muscles running longitudinally and dorsal (d) or ventral (v) for those which parallel the segment furrow. For example, the anterior end of muscle 20 attaches within the broad band labeled I and terminates within the segment (20p). The attachment site of each muscle is marked with a dark oval, and attachment sites for the clusters of muscles which constitute the major bands are marked with an extended dark bar. Since there are 30 muscles per abdominal segment, we have

not attempted to label all of the muscles, but we have indicated the positions of all expected attachment sites.

A close examination of the *in situ* hybridization pattern for *delilah* shows a nearly perfect correlation with the expected muscle attachment sites shown in Fig. 5A. The prominent stripes of staining along each segment boundary correlate with the major bands of muscles (labeled I, II, and III in Fig. 5B and C). The gaps in staining occur at positions where there are no muscles attached to the epidermis. One particularly striking feature was the *delilah* staining pattern at the dorsal midline (Fig. 5E). The relatively long row of cells that stain along the segmental boundary corresponding to the attachment sites for muscle group I was consistently found to stop short of the midline by 1 or 2 cells. A small cluster of cells immediately anterior to the segment boundary on the midline was reproducibly stained. Remarkably, this feature correlates with the structure of muscle 1 (Fig. 5D), which parallels the dorsal midline and has its major attachment site on the segment boundary. However, a portion of this muscle always splits near its posterior end and attaches to a second site anterior to the boundary, at a position that corresponds precisely with the epidermal cells that hybridize to *delilah* on the dorsal midline.

In addition to the major bands of muscles that run horizontally along the larvae, several other muscles attach at positions within each segment that do not lie along the furrow. A group of five traverse muscles (muscles 18 and 21 to 24 in Fig. 5A) extend along the dorsoventral axis and attach to the epidermis under the major muscle groups II and III. The positions of their attachment sites are labeled 21d to 23d and 21v to 23v, etc., and the *delilah* probe reproducibly hybridizes to cells that correspond to each of these attachment sites (brackets in Fig. 5B and C). The hybridization pattern frequently can be resolved into three closely spaced cells or clusters which correlate with the attachment sites of muscles 21, 22, and 23. The latter three muscles parallel each other and attach at adjacent positions on the dorsoventral axis. *delilah* staining that corresponds to the attachment sites of the other traverse

FIG. 5. (A) Schematic diagram of the muscles and expected attachment sites in a *Drosophila* embryo. The positions of individual muscles were taken from previous descriptions (1, 16, 37), and the attachment sites are indicated by dark bars or ovals. Attachment sites are labeled only for muscles that fall within each segment and hence can be resolved from the broad bands (I, II, and III) that correspond to the attachment of many adjacent muscles. (B) Lateral view of two abdominal segments after in situ hybridization with the *delilah* cDNA. The dorsal midline (dm) is at the top, and the photo encompasses only the top half of the embryo (bands I and II in panel A as indicated). The three patches of *delilah* staining that coincide with the dorsal attachment sites 21d to 23d are bracketed, and the corresponding set is shown in panel A. The arrow indicates a *delilah*-expressing epidermal cell along the dorsal midline that is anterior to the segment boundary. (C) Lateral view of abdominal segments hybridized with the *delilah* cDNA, showing the lower half of the embryo. The ventral midline (vm) is at the bottom, and bands corresponding to II and III in panel A are indicated. The patches of staining that coincide with dorsal and ventral attachment sites 21d to 23d and 21v to 23v, respectively, are indicated by brackets. A region of staining that does not correspond to known muscle attachment sites is labeled with an asterisk. (D) Phase-contrast micrograph of muscles at the dorsal midline (d), corresponding to region labeled 1p' in panel A. The attachment sites of muscle 1 along the segmental furrow (f) are indicated by leftward arrows, and the shorter, split section of muscle 1 which attaches anterior to the furrow is shown by rightward arrows. The nuclei of individual epidermal cells are visible as faint circles. (E) In situ hybridization pattern with the *delilah* cDNA along the dorsal midline. The arrows indicate the staining of epidermal cells both along the furrow (f) and slightly anterior to the furrow along the midline. The same pattern of staining is indicated by an arrow in panel B and in Fig. 4.

muscles is also observed, but it is more difficult to resolve the individual stained epidermal cells because of the close juxtaposition of sites. A detailed examination of all sections of the *delilah* staining pattern shows that there is a complete correlation with the embryonic muscle pattern, suggesting a functional relationship between *delilah* expression and the differentiation of epidermal cells into muscle attachment sites. However, we did note one cell or cluster of cells that hybridize

to the *delilah* probe yet do not correspond to any known muscle attachment site (Fig. 5C, *). The latter cells have no obvious functional relationship to the tendon cells that are otherwise the exclusive sites of *delilah* expression.

The localization of *delilah* expression to epidermal tendon cells was further confirmed by sectioning of stained embryos (Fig. 6A). Phase-contrast micrographs allow the visualization of both the epidermal layer (e) and the underlying muscles

FIG. 6. (A) Longitudinal cross-section of the epidermal (e) layer showing the position of the cells that express *delilah* (arrows). The muscles (m) form a layer just beneath the epidermis and attach at discrete positions that are coincident with the sites of *delilah* expression. (B and C) *delilah* expression in early embryos (~9 to 10 h). Note that the expression initiates in a small cluster of intrasegmental cells at the extended germ band stage. (D and E) *delilah* expression in *snail* mutant embryos.

(m). The sites of *delilah* staining coincide with the cells at the junction between epidermis and muscle, which is the location of the tendon cells.

***delilah* expression is temporally correlated with the positions of muscle attachment.** The late embryonic pattern of *delilah* expression appears at approximately 12 h after egg laying (stage 15) and remains unchanged until hatching. Muscle syncytium formation begins earlier (8 h), but recruitment of mesodermal cells into the growing muscle syncytia continues

until 12 to 14 h after egg laying (1). The precise timing of individual muscle-epidermal cell attachment has not been well documented, but the final attachment to the epidermis occurs gradually between 10 and 16 h of development (31). *delilah* expression appears to precede or at least coincide with the timing of muscle attachment. This temporal correlation was reinforced by an analysis of the earliest pattern of *delilah* expression, which occurs 9 to 10 h after egg laying at the extended germ band stage (Fig. 6B and C). The earliest cells

which express *delilah* are not positioned along the segmental boundaries but begin as a small cluster of cells in the middle of the segment. These cells may well correspond to epidermal cells which attach to the traverse muscles (i.e., muscles 18 and 21 to 24 in Fig. 5A). We speculate that these muscles might initiate their attachments at a slightly earlier stage when their attachment sites would not be obscured by the overlying dorsal lateral muscles. We also find a very low level of diffuse staining in the mesoderm, which is visible in Fig. 6B. This staining disappears about the time that epidermal expression begins.

***delilah* expression is not induced by muscle attachment.** The correlation between *delilah* expression and muscle attachment suggested two alternate models for the involvement of this gene in the differentiation of epidermal cells into functional tendon cells. The first is that *delilah* actively participates in the targeting of muscles to their proper attachment. Muscles develop from the mesoderm and find and attach to specified positions on the overlying epidermis (1). It is likely that some cell surface marker and/or morphological feature marks the individual epidermal cells as attachment sites, but it is unknown whether specific attachments are positional or involve a more complex cell-cell code. An alternate model for *delilah* expression is that the gene is induced by cell-cell adhesion as the muscles find their appropriate attachment sites. Muscles appear to send out small processes as they migrate to their appropriate locations (1, 31), so it is difficult to determine precisely when the first contact between muscle and tendon cell occurs.

To address these issues, we took advantage of the developmental mutants *snail* and *twist* that block the formation of muscles. In these mutants, epidermal and neuronal differentiation is unaltered, but the mesoderm fails to form and all mesoderm-derived cells, including muscles, are absent. Figures 6C and D show that mutant *snail* embryos continue to express *delilah* in localized patches or rows of cells, and similar results were obtained with *twist* mutants (not shown). The distorted shape of the *snail* embryo makes it difficult to orient individual *delilah*-expressing cells with their wild-type counterparts, but it is clear that *delilah* expression is not dependent on the presence of an underlying mesoderm and muscle fibers. We cannot determine if *delilah* expression is entirely unperturbed in *snail* embryos, but it would appear that most of the *delilah* expression pattern is programmed within the epidermis. These observations are consistent with a function for *delilah* in the differentiation of specific epidermal cells into muscle attachment sites or tendon cells.

One other developmental mutant was relevant to the issue of muscle-tendon cell junctions, the mutant *lethal myospheroid*. This mutation occurs within the β subunit of the integrin cell surface receptor (24, 26), and the muscle-tendon cell junctions rupture when muscles first contract (32, 45). The pattern of *delilah* expression was unaltered in *lethal myospheroid* mutants (not shown), indicating that expression is not dependent on continued cell-cell contact.

DISCUSSION

The bHLH family of proteins has frequently been implicated through genetics in the determination of specific cell types in *D. melanogaster*. The most thoroughly studied members of the family are the *achaete-scute* complex and *daughterless* (11), both of which were identified as essential for proper neuronal cell differentiation. *achaete-scute* and *daughterless* form heterodimers (30) and presumably function as transcription factors to control a group of target genes that are critical for the differentiation process (8). Loss of either gene function impairs

the ability of epidermal cells that have been committed to the neuronal lineage to differentiate into mature neurons (reviewed in reference 9). The expression of the *achaete-scute* complex and that of *daughterless* differ markedly, however, with the former being transcribed primarily in neuroblasts, whereas *daughterless* is expressed ubiquitously. By analogy, we speculate that the bHLH gene *delilah* described in this study may be involved in the differentiation of epidermal cells into the tendon cells that form the attachment sites for all muscles in *D. melanogaster*. Since *delilah* cannot homodimerize or bind to DNA alone, we presume that some other bHLH protein, perhaps the ubiquitously expressed *daughterless* (15), serves as its partner.

The tendon cells serve as the connectors of the somatic muscles to the chitinous exoskeleton. The attachment between tendon cell and muscle involves a desmosome-like junction and an extensive extracellular matrix that is laid down between the apposing membranes (31, 32). The formation of this junction takes place over an extended period of development (between 8 and 16 h after egg laying). In the early stages, clusters of fusing myoblasts are closely apposed to the overlying epidermis, and these cells send out filopodia that appear to initiate the appropriate cell-cell junctions. The interdigitation of the tendon and muscle cell membranes proceeds over the succeeding hours, and the mature junctions with a full component of extracellular matrix are not complete until 14 to 16 h (31). In addition to the juxtaposition of cell membranes, the attachment site is also the origin of numerous filaments that extend through the tendon cell and may join to the plasma membrane on the cuticle side. It is not known whether these filaments are equivalent to intermediate filaments that are common to many cells or whether they are a specialized feature of tendon cells that provides added rigidity to withstand the forces generated by muscle contraction.

The integrins are critical to the tendon cell-muscle junction and to many other cell-cell junctions in *D. melanogaster* (46) and other organisms (22). Mutation of the integrin β subunit leads to tendon cell-muscle separation in the mutant *lethal myospheroid* (24, 26). Although the defect in this mutant is in a membrane receptor, loss of the integrin β subunit also substantially delays the deposition of the extracellular matrix material at the desmosome-like attachment sites (32). Further evidence for the involvement of integrins in this junction comes from immunolocalization, in which the PS1 and PS2 integrins are localized in the apposed tendon cell and muscle membranes, respectively (7, 24). However, functional integrins are apparently not essential for the initial localization and attachment of muscles to their respective tendon cells, since the early formation of the epidermal and muscle pattern is unaltered in *lethal myospheroid* embryos, and the defect only leads to a rupture of the junction as muscle contraction begins (32, 45). Therefore other signals that promote the specific cell-cell interactions must exist, and this signal could derive from *delilah* or perhaps from the regulatory molecules that initially establish the specific pattern of *delilah* expression. Although the PS1 integrin is highly expressed in tendon cells, it is unlikely to be regulated by *delilah* because it is also expressed in many other tissues (24, 42).

Given the vast differences in *D. melanogaster* and mammalian development, it does not seem likely that attachment of muscles per se would be mediated by similar functions, since mammalian muscles are usually attached to tendons which are mesenchymally derived. However, the features of the *Drosophila* tendon cell-muscle junctions are strikingly similar to the attachment of basal epidermal cells to the basement membrane in mammalian skin. In both cases, cell adhesion is

mediated at least in part by integrins that localize to a desmosomal or hemidesmosomal structure linked to an extracellular matrix (for example, see reference 10 and references therein). An even more striking analogy is evident in the arrector pili muscles, which are responsible for movement of the hair follicles and which connect to the basement membrane opposite the basal cells of the epidermis. It will be interesting to determine if a mammalian homolog of *delilah* exists and if it participates in epidermal cell differentiation.

ACKNOWLEDGMENTS

We thank Kai Zinn, Cornelius Murre, Andrew Lassar, and Claire Cronmiller for providing the cDNA library and plasmids. We also thank Yash Hiromi, Krishna Bhat, Ted Wright, Andy Fire, Susan Abmayr, and Alan Michelson for helpful discussions.

This work was supported by grants from the NIH to M.D.C. and to E.F.W.

REFERENCES

- Bate, M. 1990. The embryonic development of larval muscles in *Drosophila*. *Development* **110**:791–804.
- Begley, C. G., P. D. Aplan, S. M. Denning, B. F. Haynes, T. A. Waldmann, and I. R. Kirsch. 1989. The gene *SCL* is expressed during early hematopoiesis and encodes a differentiation-related DNA-binding motif. *Proc. Natl. Acad. Sci. USA* **86**:10128–10132.
- Benezra, R., R. L. Davis, D. Lockshon, D. L. Turner, and H. Weintraub. 1990. The protein *Id*: a negative regulator of helix-loop-helix DNA binding proteins. *Cell* **61**:49–60.
- Berberich, S. J., and M. D. Cole. 1992. Casein kinase II inhibits the DNA-binding activity of Max homodimers but not Myc/Max heterodimers. *Genes Dev.* **6**:166–176.
- Blackwood, E. M., and R. N. Eisenman. 1991. Max: a helix-loop-helix zipper protein that forms a sequence-specific DNA-binding complex with myc. *Science* **251**:1211–1217.
- Blanan, M. A., and W. J. Rutter. 1992. Interaction cloning: identification of a helix-loop-helix zipper protein that interacts with c-fos. *Science* **256**:1014–1018.
- Bogaert, T., N. Brown, and M. Wilcox. 1987. The *Drosophila* PS2 antigen is an invertebrate integrin that, like the fibronectin receptor, becomes localized to muscle attachments. *Cell* **51**:929–940.
- Cabrera, C. V., and M. C. Alonso. 1991. Transcriptional activation by heterodimers of the *achaete-scute* and *daughterless* gene products of *Drosophila*. *EMBO J.* **10**:2965–2974.
- Campuzano, S., and J. Modolell. 1992. Patterning of the *Drosophila* nervous system: the *achaete-scute* gene complex. *Trends Genet.* **8**:202–208.
- Carter, W. G., P. Kaur, S. G. Gill, P. J. Gahr, and E. A. Wayner. 1990. Distinct functions for integrins $\alpha 3 \beta 1$ in focal adhesions and $\alpha 6 \beta 4$ /bullous pemphigoid antigen in a new stable anchoring contact (SAC) of keratinocytes: relation to hemidesmosomes. *J. Cell Biol.* **111**:3141–3154.
- Caudy, M., H. Vassin, M. Brand, R. Tuma, L. Y. Jan, and Y. N. Jan. 1988. *Daughterless*: a *Drosophila* gene essential for both neurogenesis and sex determination has sequence similarities to *myc* and the *achaete-scute* complex. *Cell* **55**:1061–1068.
- Cavener, D. R., and S. C. Ray. 1991. Eukaryotic start and stop translation sites. *Nucleic Acids Res.* **19**:3185–3192.
- Chen, Q., J. T. Cheng, L. H. Tsai, N. Schneider, G. Buchanan, A. Carroll, W. Crist, B. Ozanne, M. J. Siciliano, and R. Baer. 1990. The *tal* gene undergoes chromosome translocation in T cell leukemia and potentially encodes a helix-loop-helix protein. *EMBO J.* **9**:415–424.
- Cline, T. W. 1989. The affairs of *daughterless* and the promiscuity of developmental regulators. *Cell* **59**:231–234.
- Cronmiller, C., and C. A. Cummings. 1993. The *daughterless* gene product in *Drosophila* is a nuclear protein that is broadly expressed throughout the organism during development. *Mech. Dev.* **42**:159–169.
- Crossley, A. C. 1978. The morphology and development of the *Drosophila* muscular system, p. 499–560. In M. Ashburner and T. R. F. Wright (ed.), *The genetics and biology of Drosophila* 2b. Academic Press, London.
- Davis, R. L., P. F. Cheng, A. B. Lassar, and H. Weintraub. 1990. The myoD DNA binding domain contains a recognition code for muscle-specific gene activation. *Cell* **60**:733–746.
- Davis, R. L., and H. Weintraub. 1992. Acquisition of myogenic specificity by replacement of three amino acid residues from myoD into E12. *Science* **256**:1027–1030.
- Davis, R. L., H. Weintraub, and A. B. Lassar. 1987. Expression of a single transfected complementary DNA converts fibroblasts to myoblasts. *Cell* **51**:987–1000.
- Ellis, H. M., D. R. Spann, and J. W. Posakony. 1990. *extramacrochaetae*: a negative regulator of sensory organ development in *Drosophila* defines a new class of helix-loop-helix proteins. *Cell* **61**:27–38.
- Garrell, J., and J. Modolell. 1990. The *Drosophila extramacrochaetae* locus: an antagonist of proneural genes that like these genes encodes a helix-loop-helix protein. *Cell* **61**:39–48.
- Hynes, R. O. 1992. Integrins: versatility, modulation, and signaling in cell adhesion. *Cell* **69**:11–25.
- Johnson, J. E., S. J. Birren, and D. J. Anderson. 1990. Two rat homologues of *Drosophila achaete-scute* specifically expressed in neuronal precursors. *Nature (London)* **346**:858–861.
- Knapp, A. C., and M. D. Cole. Unpublished data.
- Leptin, M., T. Bogaert, R. Lehmann, and M. Wilcox. 1989. The function of PS integrins during *Drosophila* embryogenesis. *Cell* **56**:401–408.
- Li, B.-L., J. A. Langer, B. Schwartz, and S. Pestka. 1989. Creation of phosphorylation sites in proteins: construction of a phosphorylatable human interferon α . *Proc. Natl. Acad. Sci. USA* **86**:558–562.
- MacKrell, A. J., B. Blumberg, S. R. Haynes, and J. H. Fessler. 1988. The *lethal myospheroid* gene of *Drosophila* encodes a membrane protein homologous to vertebrate integrin beta subunits. *Proc. Natl. Acad. Sci. USA* **85**:2633–2637.
- Mellentin, J. D., S. D. Smith, and M. L. Cleary. 1989. Lyl-1: a novel gene altered by chromosomal translocation in T cell leukemia codes for a protein with a helix-loop-helix DNA binding motif. *Cell* **58**:77–84.
- Michelson, A. M., S. M. Abmayr, M. Bate, A. M. Arias, and T. Maniatis. 1990. Expression of a myoD family member prefigures muscle patterns in *Drosophila* embryos. *Genes Dev.* **4**:2086–2097.
- Murre, C., P. S. McCaw, and D. Baltimore. 1989. A new DNA binding and dimerization motif in immunoglobulin enhancer binding, *daughterless*, *myoD*, and *myc* proteins. *Cell* **56**:777–784.
- Murre, C., P. S. McCaw, H. Vaessin, M. Caudy, L. Y. Jan, Y. N. Jan, C. V. Cabrera, J. N. Buskin, S. D. Hauschka, A. B. Lassar, H. Weintraub, and D. Baltimore. 1989. Interactions between heterologous helix-loop-helix proteins generate complexes that bind specifically to a common DNA sequence. *Cell* **58**:537–544.
- Newman, S. M., Jr. 1978. An ultrastructural analysis of myogenesis in wild type and mutant embryos of *Drosophila melanogaster*. Ph.D. thesis. University of Virginia, Charlottesville.
- Newman, S. M., Jr., and T. R. F. Wright. 1981. A histological and ultrastructural analysis of developmental defects produced by the mutation *lethal(1)myospheroid* in *Drosophila melanogaster*. *Dev. Biol.* **86**:393–402.
- Olson, E. N., and W. H. Klein. 1994. bHLH factors in muscle development: dead lines and commitments, what to leave in and what to leave out. *Genes Dev.* **8**:1–8.
- Ruezinsky, D., H. Beckmann, and T. Kadesch. 1991. Modulation of the IgH enhancer's cell type specificity through a genetic switch. *Genes Dev.* **5**:29–37.
- Rushlow, C. A., A. Hogan, S. M. Pinchin, K. M. Howe, M. Lardelli, and D. Ish-Horowicz. 1989. The *Drosophila* hairy protein acts in both segmentation and bristle patterning and shows homology to N-myc. *EMBO J.* **8**:3095–3104.
- Schlissel, M., A. Voronova, and D. Baltimore. 1991. Helix-loop-helix transcription factor E47 activates germ-line immunoglobulin heavy-chain gene transcription and rearrangement in a pre-T-cell line. *Genes Dev.* **5**:1367–1376.
- Szabad, J., T. Schupbach, and E. Wieschaus. 1979. Cell lineage and development in the larval epidermis of *Drosophila melanogaster*. *Dev. Biol.* **73**:256–271.
- Tautz, D., and C. Pfeifle. 1989. A non-radioactive *in situ* hybridization method for localization of specific RNAs in *Drosophila*

- embryos reveals translational control of the segmentation gene *hunchback*. *Chromosoma* **98**:81–85.
38. **Thisse, B., C. Stoetzel, C. Gorostiza-Thisse, and F. Perrin-Schmitt.** 1988. Sequence of the *twist* gene and nuclear localization of its protein in endomesodermal cells of early *Drosophila* embryos. *EMBO J.* **7**:2175–2183.
 39. **Van Doren, M., H. M. Ellis, and J. W. Posakony.** 1992. The *Drosophila* extramacrochaetae protein antagonizes sequence-specific DNA binding by daughterless-achaete-scute protein complexes. *Development* **113**:245–256.
 40. **Villares, R., and C. V. Cabrera.** 1987. The *achaete-scute* gene complex of *Drosophila melanogaster*: conserved domains in a subset of genes required for neurogenesis and their homology to *myc*. *Cell* **50**:415–424.
 41. **Voronova, A., and D. Baltimore.** 1990. Mutations that disrupt DNA binding and dimer formation in the E47 helix-loop-helix protein map to distinct domains. *Proc. Natl. Acad. Sci. USA* **87**:4722–4726.
 42. **Wehrli, M., A. DiAntonio, I. M. Fearnley, R. J. Smith, and M. Wilcox.** 1993. Cloning and characterization of α_{PS1} , a novel *Drosophila melanogaster* integrin. *Mech. Dev.* **43**:21–36.
 43. **Weintraub, H., V. J. Dwarki, I. Verma, R. Davis, S. Hollenberg, L. Snider, A. Lassar, and S. J. Tapscott.** 1991. Muscle-specific transcriptional activation by MyoD. *Genes Dev.* **5**:1377–1386.
 44. **Wilson, R. B., M. Kiledjian, C. P. Shen, R. Benezra, P. Zwollo, S. M. Dymecki, S. V. Desiderio, and T. Kadesch.** 1991. Repression of immunoglobulin enhancers by the helix-loop-helix protein Id: implications for B-lymphoid-cell development. *Mol. Cell. Biol.* **11**:6185–6191.
 45. **Wright, T. R. F.** 1960. The phenogenetics of the embryonic mutant, *lethal myospheroid* in *Drosophila melanogaster*. *J. Exp. Zool.* **143**:77–99.
 46. **Zusman, S., R. S. Patel-King, C. Ffrench-Constant, and R. O. Hynes.** 1990. Requirements for integrins during *Drosophila* development. *Development* **108**:391–402.