Expression of a MyoD family member prefigures muscle pattern in Drosophila embryos

Alan M. Michelson, Susan M. Abmayr, Michael Bate, Alfonso Martinez Arias, and Tom Maniatis

Department of Biochemistry and Molecular Biology, Harvard University, Cambridge, Massachusetts 02138 USA; Department of Zoology, University of Cambridge, Cambridge CB2 3EJ UK

We have isolated a Drosophila gene that is expressed in a temporal and spatial pattern during embryogenesis, strongly suggesting an important role for this gene in the early development of muscle. This gene, which we have named nautilus (nau), encodes basic and helix–loop–helix domains that display striking sequence similarity to those of the vertebrate myogenic regulatory gene family. nau transcripts are initially localized to segmentally repeated clusters of mesodermal cells, a pattern that is reminiscent of the expression of the achaete-scute genes in the Drosophila peripheral nervous system. These early nau-positive cells are detected just prior to the first morphological evidence of muscle cell fusion and occupy similar positions as the later-appearing muscle precursors. Subsequently, nau transcripts are present in at least a subset of growing muscle precursors and mature muscle fibers that exhibit distinct segmental differences. These observations establish nau as the earliest known marker of myogenesis in Drosophila and indicate that this gene may be a key determinant of pattern formation in the embryonic mesoderm.

[Key Words: Myogenesis, mesoderm development; helix–loop–helix proteins; pattern formation]

Received August 8, 1990; revised version accepted September 24, 1990.

Morphological, genetic, and molecular studies of embryonic development in Drosophila melanogaster have established a conceptual framework for understanding the fundamental processes that govern early animal morphogenesis. These studies have shown that separate regulatory gene networks define characteristic patterns along the dorsoventral and anteroposterior axes of the blastoderm embryo, and that the proper hierarchical and progressive deployment of these factors is essential for the subsequent steps of gastrulation, segmentation, and organogenesis (for reviews, see Akam 1987; Anderson 1987; Gehring 1987; Nüsslein-Volhard et al. 1987; Ingham 1988; Manseau and Schupbach 1989; Martinez-Arias 1989). Superimposed on and emerging from the early patterning elements is the expression of the homeotic genes, whose function is the determination of region-specific identity [Akam 1987; Gehring 1987; Ingham 1988]. Through these mechanisms, a given cell acquires tissue-specific properties, as well as the specialized features that characterize its local identity and that distinguish it from cells of the same lineage situated elsewhere in the organism.

Integrated with these pathways are those that direct the formation of internal organs. For example, development of the nervous system involves the segregation of neuroblasts from a sheet of initially equivalent ectodermal cells, a process that is under complex genetic control and involves mechanisms such as progressive determination and cell–cell communication [Campos-Ortega 1985, 1988; Ghysen and Dambly-Chaudière 1989]. Since relatively little is known about the embryonic development of other organs in Drosophila, it is unclear to what extent these principles can be generalized or whether novel mechanisms of organogenesis might also exist.

These latter questions can be approached by examining the genetic and molecular determinants of myogenesis in Drosophila. Muscle precursors, identified as fused doublets and triplets of cells, segregate in a stereotyped pattern within the ventral mesoderm of extended germ-band embryos at the onset of shortening (Bate 1990). These derive from a subset of mononucleate mesodermal cells that previously may have undergone commitment as muscle founders (Bate 1990). By the end of germ-band shortening, the definitive muscle pattern is reflected in the segmental arrangement of muscle precursors. As is well-established for the Drosophila nervous system [Campos-Ortega 1985, 1988; Ghysen and Dambly-Chaudière 1989] and epidermis [Martinez-Arias 1989], the generation of muscle pattern may involve cell–cell interactions, in this case between mesodermal and ectodermal derivatives or possibly among cells of the mesoderm alone (Bate 1990). The molecular
detectors of these developmental processes are not known.

Recent work on myogenesis in vertebrates has led to the identification of a family of transcription factors that are both functionally and structurally related. MyoD was the first of these to be identified and cloned on the basis of its myoblast-specific expression and its ability to confer a myogenic phenotype on a variety of non-muscle cells (Davis et al. 1987, Weintraub et al. 1989). The MyoD protein contains a highly basic domain immediately adjacent to a region that has been postulated to form a helix-loop-helix (HLH) structure (Murre et al. 1989a). These two domains are both necessary and sufficient for determining the myogenic potential of MyoD, and are required for the sequence-specific binding of the MyoD protein to muscle gene regulatory elements (Tappscar et al. 1988; Lassar et al. 1989, Davis et al. 1990). Three other mammalian genes, myogenin (Edmondson and Olson 1989, Wright et al. 1989), Myf-5 (Braun et al. 1989), and MRF-4 (Rhodes and Konieczny 1989, also referred to as herculin by Miner and Wold 1990 and Myf-6 by Braun et al. 1990), were subsequently shown to have the capacity to induce muscle differentiation in tissue culture cells and have in common MyoD-like basic and HLH domains. Both MyoD and myogenin are expressed in the myotome and limb bud of mouse embryos, suggesting that these genes may function in vivo as determinants of muscle precursors (Sas soon et al. 1989).

To identify genes that control muscle development in Drosophila, we have taken advantage of the observation that essential domains of factors involved in developmental processes are often conserved in evolution (Dressler and Gruss 1988). Thus, we have isolated a Drosophila gene that contains basic and HLH domains that are strikingly similar in sequence to those of the vertebrate myogenic regulatory gene family. This region of the Drosophila homolog is as related to each of the vertebrate factors as the latter are to each other and is notably divergent from the others in the flanking amino acid sequences. We show further that the temporal and spatial expression of this gene is consistent with a role in the early patterning of at least a subset of the somatic muscles of Drosophila embryos. The present analysis of a gene that we have termed nautilus [nau] lays the groundwork for a detailed genetic and molecular investigation of myogenesis in an intact organism.

Results

Cloning of the Drosophila nau gene

We determined initially that Drosophila DNA contains sequences related to the vertebrate myogenic regulatory gene family by performing low-stringency genomic Southern blots with the HLH regions of mouse MyoD and rat myogenin as the hybridization probes. Both probes detect a family of related sequences in the Drosophila genome (data not shown). Comparable low-stringency hybridization and wash conditions were then used to screen a Drosophila genomic library (Maniatis et al. 1978). Clone 5b, which contained sequences detected with both MyoD and myogenin probes, was selected for further analysis. At high stringency, the MyoD/myogenin homologous region of clone 5b hybridizes to a single fragment in the Drosophila genome, although additional bands are detected at low stringency (data not shown).

The DNA sequence of the MyoD/myogenin cross-hybridizing region of clone 5b revealed a striking homology to the HLH domains of all members of the vertebrate myogenic regulatory gene family (Figs. 1A and 2). However, on either side of this 120-bp segment, the Drosophila genomic sequence is completely divergent from the mammalian cDNA sequences, and the HLH open reading frame was lost. Since this region of clone 5b did not appear to encode a complete protein, and since the presence of consenus 5' and 3' splice sites on the appropriate sides of the putative HLH domain suggested the existence of introns at these locations (Fig. 1A), we isolated a homologous cDNA clone for additional studies. The sequence of this cDNA revealed a long open reading frame that begins with a perfect match to the Drosophila translational initiation consensus sequence (Caven 1987) and is capable of encoding a 332-amino-acid protein. Although this is not the first ATG codon in the 5' end of the cDNA, it most likely represents the translational start site because it marks the beginning of the longest open reading frame that includes the HLH homology. In addition, an in-frame termination codon is located immediately upstream of this ATG. The existence of this stop codon has been verified by partial sequencing of an independent cDNA clone. The apparent coding region is flanked by 5'- and 3'-untranslated regions of at least 397 and 135 bp, respectively; the precise lengths cannot be assigned because the exact boundaries of the transcription unit remain to be determined. However, the cDNA clone appears to be nearly full length since its size is comparable to that of the corresponding mRNA (Fig. 3).

The nau gene contains conserved basic and HLH domains

The complete cDNA sequence not only confirmed the existence of a separate HLH-encoding exon, as suggested by the analysis of the genomic clone, but also revealed the presence of a basic domain that is highly similar to and in precisely the same location as the basic regions of the vertebrate myogenic regulatory factors (Figs. 1B and 2). The sequence similarity extends for an additional six amino acids on the amino-terminal side of the first basic residue, resulting in the total conservation of 72 contiguous amino acids. Although introns precisely demarcate the boundaries of the HLH domain, there is no intron on the immediate 5' side of the basic region (data not shown).

Pairwise sequence comparisons indicate that the nau basic and HLH domains are as closely related to the corresponding regions of each of the vertebrate muscle-specific factors as the latter are to each other (Fig. 2). Fur-
thermore, nau is much more related to the myogenic class of HLH proteins than to any of the other HLH family members, including those of Drosophila. The nau protein sequence completely diverges from those of the vertebrate myogenic factors on either side of the 72-amino-acid conserved domain, with the exception of several short blocks of similarity, as originally described by Braun et al. (1989), for which the significance is currently unclear. Thus, nau cannot be considered a homolog of any particular member of the vertebrate myogenic regulatory gene family; rather, what appears to be critical for this family is the inter- and intraspecies conservation of the basic and HLH domains, a region that is both necessary and sufficient for the DNA-binding and myogenic activation functions of MyoD (Tapscott et al. 1988; Lassar et al. 1989; Davis et al. 1990).

Chromosomal localization of nau

The cytological location of the nau gene was determined by in situ hybridization of a genomic nau fragment to the polytene chromosomes of third-instar Drosophila larvae. Hybridization was detected at a single site in the distal region of chromosome band 95A on the right arm of the third chromosome [data not shown]. To date, no mutations that map to this site have been reported.

Temporal expression of nau

Hybridization of a Northern blot containing poly(A)+ RNAs from various stages of Drosophila development with a nau cDNA probe revealed that the nau gene encodes a single 1.7-kb transcript (Fig. 3). This transcript is first detected at a low level in 4- to 8-hr embryos. In contrast to the maternally derived MyoD RNA that is present in Xenopus embryos (Hopwood et al. 1989), no nau sequences are found in purified Drosophila ovarian RNA [data not shown]. nau-encoded transcripts continue to increase during embryogenesis and reach a peak at 12–16 hr. Of note, embryonic accumulation of nau mRNA precedes that of Drosophila sequences are found in purified Drosophila ovarian RNA [data not shown].

Figure 1. Partial genomic and complete cDNA sequence of nau. [A] A partial sequence of the genomic fragment including the exon that encodes the HLH region is shown. The sequences that are present in cDNA have been boxed. The remaining sequences are presumed to represent introns because they have not been found in CDNA's and do not generate a protected fragment when used as a probe for RNase protection analysis. As underlined and indicated in boldface type, consensus 5' and 3' splice sites are present. [B] The DNA sequence derived from the cDNA isolate with the longest insert is shown. This cDNA was sequenced as described in Materials and methods. The domain corresponding to the basic regions of the vertebrate myogenic family has been indicated by a box. The positions of known introns are indicated by arrowheads.

2088 GENES & DEVELOPMENT
Figure 2. Sequence alignment of nau basic and HLH domains with those of the vertebrate myogenic regulatory gene family. (A) The 72-amino-acid regions that include the basic and HLH domains of nau have been aligned with those of all members of the vertebrate myogenic regulatory gene family. Genes included in this comparison are nau, MyoD (Davis et al. 1987), Myf-5 (Braun et al. 1989), myogenin (Wright et al. 1989), MRF4 (Rhodes and Konieczny 1989), CMD1 (Lin et al. 1989), qmfl (de la Brousse and Emerson 1990), and Xm25 (Hopwood et al. 1989). The species corresponding to embryos transcripts in Drosophila nau, (B) myogenin (Wright et al. 1989), MRF4 (Rhodes and Konieczny 1989), CMD1 (Lin et al. 1989), qmfl (de la Brousse and Emerson 1990), and Xm25 (Hopwood et al. 1989), and Xm25 (Scales et al. 1990, also named XMMyoDa; Harvey 1990). The species corresponding to embryos transcripts in Drosophila nau, MyoD(x), and MyoD(m) are indicated in parentheses: (d) human; (m) mouse; (c) chicken; (r) rat; (q) quail; (x) Xenopus. The two forms of Xenopus MyoD likely represent different alleles [Harvey 1990; Scales et al. 1990]. Not included in this table are (1) human homologs of MyoD [Myf-3], myogenin [Myf-4], and MRF4 [Myf-6] (Braun et al. 1990) and (2) a mouse homolog of MRF4 [herculin, Miner and Wold 1990]. Of the total 72 amino acids, 52 positions are conserved in genes from all members, including nau. These amino acids are indicated by boxes. Conservative changes (Glu/Asp, Arg/Lys, and Ser/Thr) have been accepted. In addition, a careful examination of the nonhomologous regions reveals that these amino acids are conserved among many of the members. (B) This table shows a pairwise comparison of the 72-amino-acid domains of all genes in A. Percentages were derived from the number of homologous amino acids, including the conservative changes described above.

Localization of nau transcripts in Drosophila embryos

nau mRNA was localized in Drosophila embryos of various developmental stages using a whole-mount in situ hybridization protocol [Tautz and Pfeifle 1989]. nau transcripts are first detected ~7 hr postfertilization, corresponding to the end of the fully extended germ band stage (Fig. 4A). These transcripts are initially found in segmentally repeated clusters of cells in both lateral and medial positions within the ventral somatic mesoderm. These early nau-containing clusters are composed of three to five cells that have not yet undergone fusion (Fig. 4A). At this stage, twice as many clusters are present in medial as compared with lateral positions of each metameric unit.

Just prior to the onset of germ-band shortening, the complexity of the nau pattern increases as many new expressing cell clusters appear between the initial medial and lateral groups [Fig. 4B]. However, when viewed from the ventral side, it is apparent that these intermediate clusters are specific to the abdominal segments [Fig. 4C]. Additional differences among the thoracic segments are also apparent before shortening begins (Fig. 4C).

As the germ band retracts, there is a change in the pattern of nau expression. Expression in medial cells declines, new clusters appear laterally, and there is an evolution of distinct dorsoventral and anteroposterior patterns [Fig. 4D]. Thus, three rows emerge along the length of the embryo, one each in the dorsal, ventral, and pleural regions. These correspond to the locations of the three major groups of embryonic somatic muscles [Campos-Ortega and Hartenstein 1985]. By the completion of germ-band retraction, these rows have characteristic appearances [Fig. 4E]. The dorsal set includes cells in all segments with apparently greater expression anteriorly. The reciprocal is true of the ventral set, with expression seen only in the abdominal segments A1–A7. Very weak expression occurs in the pleural row and seems to fade posteriorly.

By the time the mature muscle pattern is complete, nau transcripts are present in a small number of segmentally repeated ventral abdominal muscles that have a characteristic elongated appearance due to their stretching between epidermal attachment sites [Fig. 4F]. Segmental differences are also seen at this stage in the dorsal and pleural muscle groups. Despite the dynamic changes that occur in nau expression between various embryonic stages, a consistent feature is the restricted segmental and subsegmental distribution of expressing cells in the somatic mesoderm. Of note, nau mRNA is not detected at any developmental stage in cells of the 

GENES & DEVELOPMENT 2089
visceral mesoderm. Staining has been transiently observed in cells anterior to thoracic segment T1, the significance of which remains unclear.

**The location of nau-expressing cells coincides with that of early muscle precursors**

The first sign of muscle differentiation in *Drosophila* embryos is the formation of ventral muscle precursors through the fusion of certain mesodermal cells overlying the central nervous system, an event that occurs at the onset of germ band retraction ([Bate 1990]). Although the earliest expression of *nau* just precedes fusion, there is a striking correspondence between the location of the medial *nau*-positive cell clusters and that of the early muscle precursors ([Fig. 5A,B]). Furthermore, the increased complexity of the *nau* pattern that occurs before shortening ([Fig. 4B,C]) correlates with the appearance of increased numbers of muscle precursors at similar positions within the mesoderm ([Bate 1990]). The lateral clusters of cells containing *nau* transcripts coincide with muscle precursors underlying the lateral epidermis.

**nau expression in dorsoventral polarity mutants**

Mutations in any of the dorsal group of *Drosophila* maternal-effect genes result in a dorsalized phenotype. That is, ventrally derived structures including mesoderm fail to develop and are instead replaced by derivatives of cells that occupy a more dorsal location in the blastoderm embryo ([Anderson 1987]). *dorsal* is the prototypical member of this gene class and probably functions at the end of the maternal-effect pathway ([Nüsslein-Volhard et al. 1980; Anderson 1987]). If *nau* expression is associated with developing muscle as suggested by the above analysis, then *nau* transcripts should not be present in homozygous *dorsal* mutant embryos. This result was indeed obtained using a sensitive RNase protection assay to analyze RNA from *dl/l* embryos that had developed to a point at which both heterozygous and wild-
A *Drosophila* homolog of the MyoD gene family

Figure 4. Localization of *nau* mRNA in whole-mount embryos. The spatial distribution of *nau* mRNA in embryos was determined by in situ hybridization using a digoxigenin-labeled probe as described by Tautz and Pfeifle (1989) and modified as discussed in Materials and methods. The presence of *nau* mRNA is indicated by the product of the alkaline phosphatase reaction. Stages were evaluated based on Campos-Ortega and Hartenstein (1985). In all views, anterior is to the left. Dorsal is at the top in D–F.

(A) Dorsal view of an embryo at late stage 11, just prior to germ-band shortening at ~7 hr postfertilization. Arrows indicate locations of the tracheal pits. (B) Dorsal view of an embryo at early stage 12, immediately prior to the onset of germ-band shortening at ~7.5 hr postfertilization. (C) Ventral view of the same embryo as in B. (D) Lateral view of an embryo at late stage 12, at ~9 hr postfertilization. (E) Lateral view of an embryo at stage 13, following completion of germ-band shortening at 10 hr postfertilization. (F) Lateral view of an embryo at early stage 14, at ~10.5 hr postfertilization.

type embryos contained levels of *nau* mRNA that are readily detected (Fig. 6A).

The phenotype of the *dl* allele not only includes loss of the mesoderm but also of more lateral ectodermal derivatives (Anderson 1987). However, in *twist* mutants, only the mesoderm and its derivatives fail to develop. Therefore, we examined expression of *nau* by in situ hybridization of embryos homozygous for a null allele of *twist*. The typical patterns of *nau* transcript localization are not observed in embryos with the characteristic *twist* phenotype (Fig. 6B; data not shown). Embryos having this appearance represented 24% of the 370 embryos examined, whereas the expected *nau* patterns were observed in the normal embryos in this population [data not shown]. Thus, *nau* mRNA indeed accumulates within the mesoderm and this expression is dependent on the establishment of correct dorsoventral polarity through the combined action of maternal and zygotic determinants.

Discussion

The *Drosophila* *nau* gene was identified and cloned on the basis of its sequence similarity to the known func-
The basic and HLH domains of MyoD and its mammalian homologs are both necessary and sufficient for activating myogenic commitment and differentiation in tissue culture cells and for binding to the regulatory regions of muscle-specific genes [Tapscott et al. 1988; Lassar et al. 1989, Davis et al. 1990]. Members of this family are expressed solely in skeletal muscle [Davis et al. 1987, Braun et al. 1989, 1990; Edmondson and Olson 1989, Rhodes and Konieczny 1989, Wright et al. 1989; Miner and Wold 1990] and in regions of the mouse [Sassoon et al. 1989; Wright et al. 1989] and quail [de la Brousse and Emerson 1990] embryos from which muscle is derived. nau is not only equally well conserved at the protein level within the functional domains of each of these vertebrate factors but also has striking parallels in its patterns of expression. [1] nau transcripts are detected in embryos immediately before the earliest sign of muscle differentiation and prior to transcription of muscle-specific genes such as MHC. [2] There is a second peak of nau expression during formation of the adult musculature. [3] nau transcripts are first detected in unfused mesodermal cells that occupy similar positions as the slightly later-appearing muscle precursors in extended germ-band embryos. [4] After germ-band shortening, nau-positive cells are located in dorsal, pleural, and ventral positions that are analogous to those of the precursors of the major groups of somatic muscles. [5] In even later embryos, nau RNA is found in at least a subset of mature somatic muscles. [6] nau transcripts are greatly enriched in RNA from isolated legs of adult flies, structures that are largely composed of muscle. [7] There is no expression of nau in the visceral mesoderm or its derivatives, analogous to the lack of expression of the vertebrate homologs in smooth and cardiac muscle. The interspecies sequence conservation and the somatic muscle lineage-specific expression of the nau gene are strong associations that implicate it in some aspect of myogenesis in Drosophila. This possibility can only be established by genetic analysis of the nau locus and by the identification of targets of nau action, work which is in progress.

Potential myogenic function of the nau gene

If nau is involved in muscle development, what might be its specific role? The earliest sign of myogenesis in Drosophila is the appearance of doublets and triplets of fused cells in the ventral mesoderm overlying the developing central nervous system of late extended germ-band embryos [Fig. 5A; Bate 1990]. These fused cells, referred to as muscle precursors, continue to grow into mature muscles by fusion with other mesodermal cells that are located in their vicinity. Since the complete muscle pattern is evident by the end of germ-band shortening, the mechanisms of muscle patterning must be operative prior to this time [Bate 1990]. One interpretation of these findings is that the initial event in muscle pattern formation is the specification of unfused muscle founder cells at appropriate positions in the embryo. The unique properties of a given muscle would then be determined by its particular founder [Bate 1990]. Cells containing nau transcripts are first observed in positions

---

Figure 5. Relation of nau-positive cells to early muscle precursors. (A) Embryos were dissected and stained with toluidine blue for morphological examination of muscle precursors. Arrows indicate binucleate (fused) muscle precursors. Note their positions relative to the levels of the tracheal pits indicated by stars; the actual tracheal pits are situated more laterally, beyond the presented field of nau and the midline indicated with an inverted triangle. (B) An enlarged view of a digoxigenin-labeled embryo at a stage comparable to that shown in A. Arrows refer to the nau-staining cells in medial positions. Note the position of these cells in comparison with the binucleate cells in A and relative to the tracheal pits and the midline indicated as in A. In this view, the tracheal pits are at the indicated positions but are not in exact alignment because the embryo is slightly rotated. Anterior is at the top in both A and B.
similar to those of the earliest muscle precursors. Thus, nau expression appears to prefigure some of the definitive embryonic muscles. Furthermore, the earliest nau-containing cells are mononucleate, raising the possibility that these are the hypothetical muscle founders and that nau is involved in their commitment. Alternatively, pattern formation in the mesoderm may depend not on the segregation of single muscle founder cells but, instead, on the direct definition of precursors. According to this model, the earliest clustered expression of nau specifies the muscle precursors themselves. A third possibility is that the cells expressing nau are those that are competent to fuse with precursors and that the latter are independently defined. The resolution of the current whole-mount in situ hybridization technique does not allow an unequivocal resolution of these three possibilities for the nature of the early nau-positive cells.

Whether the cells that express nau are founders, precursors, or fusion-competent cells, we propose that this gene has a myogenic determination and differentiation function in Drosophila that is analogous to that suggested for MyoD in mouse, frog, and quail embryos (Hopwood et al. 1989; Sassoon et al. 1989; de la Brousse and Emerson 1990). That is, each of these genes confers the myogenic phenotype on multipotent mesodermal cells. However, a unique feature of the pattern of nau expression is that only a subset of all myogenic cells within a particular segment express the gene at any given time. This is true both in extended germ-band embryos when muscle precursors first appear and considerably later when nau transcripts are found in a small number of mature muscles. Excluding trivial explanations for the failure to detect nau in a larger population of cells (see Results), this raises the intriguing possibility that the expression of nau reflects muscle cell identity even at the earliest stages of myogenesis. Thus, nau may not only determine the general features of the myogenic lineage, but it may also contribute to the specific properties of particular cells. This is in contrast to the expression of MyoD in mouse, frog, and quail embryos in which cells in the entire myotome of all somites appear to contain this factor, at least at the resolution of current methods (Hopwood et al. 1989; Sassoon et al. 1989; de la Brousse and Emerson 1990).

The differentiation of a particular cell not only involves the acquisition of the general properties of its lineage but also the features that delineate its unique identity. Thus, all muscles must express a common set of contractile proteins [although even here isoform variations may occur (Emerson and Bernstein 1987)], as well as products that confer particular size, shape, location, and attachment. One possibility is that the locations of the early myogenic cells contribute to subsequent muscle identity through the reading of previously encoded positional information. Cellular interactions, such as those that occur in the developing nervous system, may also influence this process. This implies that nau expression responds either directly or indirectly to genes controlling segmentation, dorsoventral polarity, and cell-cell communication. This could explain how metamerism of the internal organs is integrated with that of the entire organism.

Multiple independent myogenic genes have been identified in all mammalian species examined (Davis et al. 1987; Braun et al. 1989, 1990; Rhodes and Konieczny...
Michelson et al.

1989; Wright et al. 1989; Miner and Wold 1990). Muscle regulatory loci other than nau may also exist in Drosophila, because at any given time the expression of nau can account for only a subset of the embryonic muscles or their precursors. In all species including Drosophila, the continued expression of myogenic factors in mature muscles implies that they have a secondary maintenance function.

Evolution of the HLH superfamily

The HLH domain of nau is encoded by a single 120-bp exon. Given the presence of this domain in a variety of proteins having diverse developmental functions in different species [Murre et al. 1989a], it is possible that the nau HLH exon represents an ancestral genomic organization that facilitated the evolution of this superfamily through exon shuffling [Gilbert 1978]. This hypothesis is particularly intriguing in view of the recent characterization of several factors that contain HLH but not adjacent basic regions [Benezra et al. 1990; Ellis et al. 1990; Garrell and Modolell 1990].

Different members of the HLH superfamily are capable of heterodimerization in vitro [Murre et al. 1989b; Brennan and Olson 1990], and recent evidence suggests that such interactions may have physiological consequences [Benezra et al. 1990]. For example, this appears to be the case for MyoD and E12 [A. Lassar and H. Weintraub, pers. comm.], a factor originally identified on the basis of its binding to an immunoglobulin gene regulatory region [Murre et al. 1989a]. This association is of particular interest in Drosophila since the mutant phenotype of daughterless (da), a gene that is closely related to mammalian E12, includes muscle defects [Caudy et al. 1988a,b]. Perhaps lack of appropriate interactions between nau and da proteins is responsible for this phenotypic effect. Thus, various members of the HLH superfamily may have evolved in concert to preserve regulatory interactions.

nau and mesodermal pattern formation

Independent of the precise function of nau, the present work establishes this gene as a valuable early marker for the muscle lineage. For example, it should be informative to examine the effects on nau expression of other genes that influence muscle development (Hooper 1986).

One of the intriguing features of the early pattern of nau-positive cells is that these cells strictly occur in metameric clusters. This is reminiscent of the clustering of founder cells in the Drosophila embryonic peripheral nervous system [Ghysen and Dambly-Chaudière 1989], a phenomenon that is reflected in the expression pattern of different members of theachaete-scute (AS-C) gene complex [Cabrera et al. 1987; Romani et al. 1987]. The spatial restriction of clusters of ectodermal cells containing AS-C gene transcripts defines the sense organ pattern much as we have proposed for nau in muscle development. Just as nau expression accounts for only a subset of muscles or their precursors at any given time, the expression of a single AS-C transcription unit does not define the complete neuroblast pattern [Cabrera et al. 1987; Romani et al. 1987]. Another similarity between the AS-C and nau genes is that both encode HLH proteins, a structural feature that may prove to be a hallmark of such patterning elements.

Whether these descriptive parallels between myogenesis and neurogenesis will apply at a functional level must await a detailed genetic study of muscle development in Drosophila. Having identified the nau gene as a prime candidate for involvement in this developmental pathway, we are now in a position to undertake this analysis.

Materials and methods

Library screening

To isolate a genomic nau fragment, 250,000 clones from a Drosophila genomic library [Maniatis et al. 1978] were initially screened at low stringency with the entire mammalian MyoD cDNA [Davis et al. 1987] as the hybridization probe. The conditions for hybridization were 6 x SSC, 5 x Denhardt’s reagent, 100 µg/ml salmon sperm DNA, 1 mM EDTA, 10 mM sodium phosphate buffer [pH 7.0], and 0.5% SDS. Hybridization was carried out at 60°C for a minimum of 18 hr, with a probe labeled by random priming [Feinberg and Vogelstein 1983] to a specific activity of 1 x 10^6 to 2 x 10^6 cpm/µg at a concentration of 2 x 10^4 cpm/ml. The filters were washed in 2 x SSC, 10 mM sodium phosphate buffer [pH 7.0], 1 mM EDTA, and 0.5% SDS. Successive washes were done at 37°C, 45°C, and 50°C. The filters were then exposed overnight with an intensifying screen. Multiple positive plaques were obtained and re-screened with subfragments of the mammalian cDNA. The clones that hybridized to the HLH domain were analyzed further. Of these, one clone (5b) with a significant degree of homology was characterized further.

Three independent cDNA clones were isolated from a 3- to 12-hr embryonic agt10 cDNA library [Poole et al. 1985], kindly provided by H. Vassein (University of California, San Francisco). A total of 1.5 x 10^6 clones from the E7 pool were screened at high stringency using previously isolated Drosophila nau clones. Restriction mapping indicated that these three clones overlapped. The longest insert, which corresponded approximately to the length of the mRNA, was sequenced.

DNA sequencing

Genomic restriction fragments were cloned into a pSP73 vector. These included the 0.8-kb BamHI fragment, which contained the HLH domain, and two subclones of 478 and 415 nucleotides that were generated by HindIII digestion of the 0.8-kb fragment. Double-stranded plasmids were sequenced by the deoxynucleotide termination procedure using Sequenase [U.S. Biochemicals], initiating from the T7 and SP6 primers in the vector.

Overlapping series of BAL-31 deletions from both the 5' and 3' ends of the longest cDNA clone were generated and subcloned into a pSP73 vector. Double-stranded deletions were sequenced as above. Sequences were determined from both strands. Synthetic oligonucleotides were used to fill in gaps.
Drosophila stocks

Flies were raised on standard yeast—cornmeal—agar medium at 25°C, except where indicated.

The dpI allele is described in Nüsslein-Volhard et al. (1980) and was provided by R. Lehmann (MIT). This stock was maintained at 18°C. The twiDP6 allele was generated by Nüsslein-Volhard and co-workers (Tearle and Nüsslein-Volhard 1987) and provided by the Gelbart laboratory (Harvard University).

RNA preparation

For 0- to 4-hr, 4- to 8-hr, and 8- to 12-hr embryonic RNAs, embryos were collected on yeasted grape juice plates from population cages at 4-hr intervals and aged at 25°C. For 12- to 16-hr and 16- to 20-hr RNAs, embryos were collected at 2-hr intervals and aged at 25°C until the midpoint of the indicated stage. They were harvested, dechorionated in 50% Clorox, frozen on dry ice, and stored at ~80°C. First-instar larvae were harvested at ~36 hr of development. Wandering third-instar larvae were isolated. Pupae were isolated on day 2 (early) and day 4 (late) of pupal development. Adults were harvested 2-3 days after eclosion. After harvesting, flies at all developmental stages were frozen on dry ice and stored at ~80°C. For examination of naul in dorsal mutants, 0- to 12-hr embryos were obtained from population cages of hand-sorted homozygous dorsal females.

Without prior thawing, tissues were pulverized on dry ice with a mortar and pestle, and quickly transferred to guanidinium solution (4 M guanidinium thiocyanate, 25 mM sodium citrate, 0.5% Sarkosyl, 0.1 M β-mercaptoethanol). The RNA was then processed according to standard procedures (Sambrook et al. 1989). Poly(A)+ RNA was isolated by two passages over oligo(dT)-cellulose as described (Sambrook et al. 1989).

RNA analysis

Northern analysis utilized ~4 μg of poly(A)+ RNA from the indicated developmental stages. The RNA was fractionated on a 1% agarose gel containing 6.7% formaldehyde, transferred to Biotrans (ICN) nylon membrane, hybridized, and washed under standard conditions (Sambrook et al. 1989). The DNA probes were generated by random priming (Feinberg and Vogelstein 1983) from isolated fragments in low-melting-temperature agarose. Probes included a 1.6-kb naul cDNA that is lacking a small portion of the 5′-untranslated region, a 640-nucleotide EcoRI–HindIII fragment from RP49 (O’Connell and Rosbash 1984), and a 2069-nucleotide HindIII fragment that includes exons 2, 3a, and 3b from the gene encoding MHC (Wassenberg et al. 1987).

RNase protections utilized 50 μg of total RNA and were carried out using standard procedures (Sambrook et al. 1989). The naul probe was generated from a 478-bp HindIII–BamHI genomic fragment that includes the exon encoding the HLH region, the following exon, and the intron between them (see Fig 1A). Thus, this probe generates protected fragments of 120 and 75 nucleotides, corresponding to the lengths of the two small exons. The a1-tubulin control was generated from a 400-bp XbaI–SalI genomic fragment (Kalayan and Wensink 1982) and cloned into pSP65. It gives a protected fragment of ~150 nucleotides.

Whole-mount embryo analysis

In situ hybridization to whole embryos was carried out using the Genius DNA labeling and detection kit (Boehringer Mannheim) according to the procedure of Tautz and Pfeffer (1989). The following modifications of the standard protocol were introduced based on our experience as well as that of B. Edgar (University of California, San Francisco) and D. Andrew (University of Colorado at Boulder). First, after synthesis with the Genius kit according to the manufacturer’s instructions, the probe was diluted 2.5-fold with H2O, spun over a 1-ml Sephadex G-50 column to remove unincorporated nucleotides, and lyophilized to 10 μl. Second, the embryos were treated with 0.03% H2O2 [freshly diluted from a 30% stock] for 2 min at room temperature in 70% ethanol, prior to the rehydration and proteinase K digestion steps. Third, the hybridization was carried out at 48°C for 18–24 hr with a digoxigenin-labeled probe concentration of 0.5 μg/ml (with respect to the original template), followed by four washes at 48°C. These included two washes with 100% hybridization solution, one wash with 4 : 1 hybridization solution/PBT, and one wash with 3 : 2 hybridization solution/PBT. PBT is PBS containing 0.1% Tween 20. The hybridization solution consisted of 5 × SSC, 50% formamide, 10 mM sodium phosphate buffer (pH 7.0), 1 × Denhardt’s reagent, 0.1% Tween 20, and 100 μg/ml tRNA. Finally, the alkaline phosphatase reaction was stopped by washing the embryos in PBT. Individual embryos were then mounted nonpermanently in 90% glycerol and examined immediately. When mounted in this fashion, the embryos could be rotated on the slide to examine multiple views.

Flat embryo preparations

The flat preparations of dissected, toluidine blue-stained embryos shown in Figure 5A have been described (Bate 1990).

Acknowledgments

We thank D. Caron for excellent technical assistance. We are also grateful to A. Lassar, H. Weintraub, W. Wright, and H. Arnold for providing cDNAs, in addition to information prior to publication. Finally, we would like to thank members of the Gelbart laboratory for numerous discussions and technical advice, and C. Jennings for helpful discussions and critical reading of the manuscript. A.M.M. is a Burroughs Wellcome Fund Fellow of the Life Sciences Research Foundation. S.M.A. is a Damon Runyon–Walter Winchell Cancer Research Fund Fellow. This work was supported by grants from the Wellcome Trust to M.B. and A.M.A., by National Institutes of Health (NIH) grant 5RO1GM29379 to T.M., and by NIH Program Grant 5P01GM29301.

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

Note added in proof

The nucleotide sequence data reported will appear in the EMBL Nucleotide Sequence Database under accession number X56161.

References

Akam, M. 1987. The molecular basis for metameric pattern in the Drosophila embryo. Development 101: 1-22.

Anderson, K.V. 1987. Dorsal-ventral embryonic genes of Drosophila. Trends Genet. 3: 91–97.

Bate, M. 1990. The embryonic development of larval muscles in Drosophila. Development [in press].
Michelson et al.

Benezra, R., R.L. Davis, D. Lockshon, D.L. Turner, and H. Weintraub. 1990. The protein Id: A negative regulator of helix-loop-helix DNA binding proteins. *Cell* 61: 49–59.

Braun, T., G. Buschhausen-Denker, E. Bober, E. Tannich, and H.H. Arnold. 1989. A novel human muscle factor related to but distinct from MyoD1 induces myogenic conversion in 10T1/2 fibroblasts. *EMBO J.* 8: 701–709.

Braun, T., E. Bober, B. Winter, N. Rosenthal, and H.H. Arnold. 1990. Myf-6, a new member of the human gene family of myogenic determination factors: Evidence for a gene cluster on chromosome 12. *EMBO J.* 9: 821–831.

Brennan, T.I. and E.N. Olson. 1990. Myogenin resides in the nucleus and acquires high affinity for a conserved enhancer element on heterodimerization. *Genes Dev.* 4: 582–595.

Cabrera, C.V., A. Martinez-Arias, M. and M. Bate. 1987. The gene essential for myogenesis in Caenorhabditis elegans encodes a helix-loop-helix protein. *Nature* 335: 25–34.

Campos-Ortega, J.A. and S.I. Bernstein. 1987. Molecular genetics of the myogenic determination program. *Science* 235: 1245–1252.

Edmondson, D.G. and E.N. Olson. 1989. A gene with homology to the myc similarity region of MyoD1 induces myogenic conversion in C. elegans. *Cell* 58: 823–831.

Edmondson, D.G. and E.N. Olson. 1989. A gene with homology to the myc similarity region of MyoD1 is expressed during myogenesis and is sufficient to activate the muscle differentiation program. *Genes Dev.* 3: 628–640.

Edmondson, D.G. and E.N. Olson. 1989. A gene with homology to the myc similarity region of MyoD1 is expressed during myogenesis and is sufficient to activate the muscle differentiation program. *Genes Dev.* 3: 628–640.

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

Emerson, C.P. and S.I. Bernstein. 1987. Molecular genetics of myosin. *Annu. Rev. Biochem.* 56: 695–726.

Feinberg, A.P. and B. Vogelstein. 1983. A technique for radiolabeling DNA restriction endonuclease fragments to high specific activity. *Anal. Biochem.* 132: 6–13.

Garrels, J. and J. Modolell. 1990. The *Drosophila* *extraction* chaete locus, an antagonist of proneural genes that, like these genes, encodes a helix-loop-helix protein. *Cell* 61: 39–48.

Gehring, W.J. 1987. Homeoboxes in the study of development. *Science* 236: 1245–1252.

Gilbert, W. 1978. Why genes in pieces? *Nature* 271: 501.

Ghysen, A. and C. Dambly-Chaudière. 1989. Genesis of the *Drosophila* peripheral nervous system. *Trends Genet.* 5: 251–255.

Harvey, R. 1990. The Xenopus MyoD gene: An unlocalised maternal mRNA predates lineage-restricted expression in the early embryo. *Development* 108: 669–680.

Herr, W., R.A. Sturm, R.G. Clerc, L.M. Corcoran, D. Baltimore, P.A. Sharp, H.A. Ingraham, M.G. Rosenfeld, M. Finney, G. Ruvkun, and H.R. Horvitz. 1988. The POU domain: A large conserved region in the mammalian pit-1, oct-1, oct-2, and *Caenorhabditis elegans* unc-86 gene products. *Genes Dev.* 2: 1513–1516.

Hooper, J. 1986. Homeotic gene function in the muscles of *Drosophila* larvae. *EMBO J.* 5: 2321–2329.

Hopwood, N.D., A. Pluck, and J.B. Gurdon. 1989. MyoD expression in the forming somites is an early response to mesoderm induction in Xenopus embryos. *EMBO J.* 8: 3409–3417.

Ingham, P.W. 1988. The molecular genetics of embryonic pattern formation in *Drosophila*. *Nature* 335: 25–34.

Kaliszan, L. and P.C. Wensink. 1982. Developmental regulation of *Drosophila* a-tubulin genes. *Cell* 29: 91–98.

Kiehart, D.P., M.S. Lutz, D. Chan, A.S. Ketchum, R.A. Laymon, B. Nguyen, and L.S.B. Goldstein. 1989. Identification of the gene for fly non-muscle myosin heavy chain: *Drosophila* myosin heavy chains are encoded by a gene family. *EMBO J.* 8: 913–922.

Lassar, A.B., J.N. Buskin, D. Lockshon, R.L. Davis, S. Apone, S.D. Hauschka, and H. Weintraub. 1989. MyoD is a sequence-specific DNA binding protein requiring a region of myc homology to bind to the muscle creatine kinase enhancer. *Cell* 58: 823–831.

Lin, Z-Y., C.A. Dechesne, J. Eldridge, and B.M. Paterson. 1989. An avian muscle factor related to MyoD1 activates muscle-specific promoters in nonmuscle cells of different germ-layer origin and in BrdU-treated myoblasts. *Genes Dev.* 3: 986–996.

Liu, W.-E., L. Glazer, D. Segal, J. Schlessinger, and B.-Z. Shilo. 1985. The *Drosophila* EGF receptor gene homolog: Conservation of both hormone binding and kinase domains. *Cell* 40: 599–607.

Maniatis, T., R.C. Hardison, E. Lacy, J. Lauer, C. O’Connell, D. Quon, G.K. Sim, and A. Efstratiadis. 1978. The isolation of structural genes from libraries of eucaryotic DNA. *Cell* 15: 687–701.

Manseau, I.J. and T. Schupbach. 1989. The egg came first, of course Trends Genet. 5: 400–405.

Martinez-Arias, A. 1989. A cellular basis for pattern formation in the insect epidermis. *Trends Genet.* 5: 262–267.

Massague, J. 1987. The TGF-β family of growth and differentiation factors. *Cell* 49: 437–438.

Miner, J.H. and B. Wold. 1990. Herculin, a fourth member of the MyoD family of myogenic regulatory genes. *Proc. Natl. Acad. Sci.* 87: 1089–1093.

Murre, C., P.S. McCaw, and D. Baltimore. 1989a. A new DNA binding and dimerization motif in immunoglobulin enhancer binding, daughterless, MyoD, and myc proteins. *Cell* 56: 777–783.

Murre, C., P.S. McCaw, H. Vaessen, M. Caudy, L.Y. Jan, Y.N. Jan, C.V. Cabrera, J.N. Buskin, S.D. Hauschka, A.B. Lassar, H. Weintraub, and D. Baltimore. 1989b. Interactions between heterologous helix-loop-helix proteins generate complexes that bind specifically to a common DNA sequence. *Cell* 58: 537–544.
Nüsslein-Volhard, C., M. Lohs-Schardin, K. Sander, and C. Cremer. 1980. A dorso-ventral shift of embryonic primordia in a new maternal-effect mutant of Drosophila. Nature 283: 474–476.

Nüsslein-Volhard, C., H.G. Frohnhöfer, and R. Lehmann. 1987. Determination of anteroposterior polarity in Drosophila. Science 236: 1675–1681.

O’Connell, P. and M. Rosbash. 1984. Sequence, structure, and codon preference of the Drosophila ribosomal protein 49 gene. Nucleic Acids Res. 12: 5495–5513.

Poole, S.J., L.M. Kauvar, B. Drees, and T. Kornberg. 1985. The engrailed locus of Drosophila: Structural analysis of an embryonic transcript. Cell 40: 37–43.

Rhodes, S.J. and S.F. Konieczny. 1989. Identification of MRF4: A new member of the muscle regulatory factor gene family. Genes Dev. 3: 2050–2061.

Romani, S., S. Campuzano, and J. Modeell. 1987. The achaete-scute complex is expressed in neurogenic regions of Drosophila embryos. EMBO J. 7: 2085–2092.

Rozek, C.E. and N. Davidson. 1983. Drosophila has one myosin heavy-chain gene with three developmentally regulated transcripts. Cell 32: 23–34.

Sambrook, J., E.F. Fritsch, and T. Maniatis. 1989. Molecular cloning: A laboratory manual. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York.

Sassoon, D., G. Lyons, W.E. Wright, V. Lin, A. Lassar, H. Weintraub, and M. Buckingham. 1989. Expression of two myogenic regulatory factors myogenin and MyoD1 during mouse embryogenesis. Nature 341: 303–307.

Scales, J.B., E.N. Olson, and M. Perry. 1990. Two distinct Xenopus genes with homology to MyoD1 are expressed before somite formation in early embryogenesis. Mol. Cell. Biol. 10: 1516–1524.

Scott, M.P., J.W. Tamkun, and G.W. Hartzell. 1989. The structure and function of the homeodomain. Biochim. Biophys. Acta. 989: 25–48.

Struhl, K. 1989. Helix-turn-helix, zinc finger, and leucine zipper motifs for eukaryotic transcriptional regulatory proteins. Trends Biochem. 14: 137–140.

Tapscott, S.J., R.L. Davis, M.J. Thayer, P.-F. Cheng, H. Weintraub, and A.B. Lassar. 1988. MyoD1: A nuclear phosphoprotein requiring a myc homology region to convert fibroblasts to myoblasts. Science 242: 405–411.

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.

Tearle, R. and C. Nüsslein-Volhard. 1987. Tubingen mutants and stocklist. Dros. Inf. Serv. 66: 209–269.

Wassenberg, D.R., II, W.A. Kronert, P.T. O'Donnell, and S.I. Bernstein. 1987. Analysis of the 5' end of the Drosophila muscle myosin heavy chain gene. Alternatively spliced transcripts initiate at a single site and intron locations are conserved compared to myosin genes of other organisms. J. Biol. Chem. 262: 10741–10747.

Weintraub, H., S.J. Tapscott, R.L. Davis, M.J. Thayer, M.A. Adam, A.B. Lassar, and A.D. Miller. 1989. Activation of muscle-specific genes in pigment, nerve, fat, liver, and fibroblast cell lines by forced expression of MyoD. Proc. Natl. Acad. Sci. 86: 5434–5438.

Wright, W.E., D.A. Sassoon, and V.K. Lin. 1989. Myogenin, a factor regulating myogenesis, has a domain homologous to MyoD. Cell 56: 607–617.
Expression of a MyoD family member prefigures muscle pattern in Drosophila embryos.

A M Michelson, S M Abmayr, M Bate, et al.

*Genes Dev.* 1990, 4:

Access the most recent version at doi:10.1101/gad.4.12a.2086

References

This article cites 58 articles, 16 of which can be accessed free at:

http://genesdev.cshlp.org/content/4/12a/2086.full.html#ref-list-1

License

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