Characterization of the Gene for mp20: A *Drosophila* Muscle Protein That Is Not Found in Asynchronous Oscillatory Flight Muscle

Agnes Ayme-Southgate, Paul Lasko, Cynthia French, and Mary Lou Pardue

Department of Biology, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139

Abstract. A *Drosophila melanogaster* gene encoding a muscle specific protein was isolated by differential screening with RNA from primary cultures of myotubes. The gene encodes a 20-kD protein, muscle protein 20 (mp20), that is not detected in the asynchronous oscillatory flight muscles, but is found in most, if not all, other muscles (the synchronous muscles). The sequence of the protein, deduced from the DNA, contains two regions of 12 amino acids with significant similarity to high-affinity calcium-binding sites of other proteins. This protein is easily extracted from the contractile apparatus and thus does not seem to be a tightly bound structural component. The gene (located in polytene region 49F 9-13) is unique in the *D. melanogaster* genome and yields two transcripts, 1.0 and 0.9 kb long. The levels of the two transcripts are regulated differently during development, yet the coding regions of the two transcripts are identical.

I N general, insect muscle is very similar to vertebrate striated muscle. Similarities are seen both in the ultrastructure of the myotubes and in the proteins that have been characterized. The best studied muscle proteins are those directly involved in the structure of the myofibrils. The genes for several myofibrillar proteins have been characterized in *Drosophila*. These proteins include actin (Fyrberg et al., 1980; Tobin et al., 1980), myosin heavy chain (Bernstein et al., 1983; Rozek and Davidson, 1983), two of the myosin light chains (Falkenthal et al., 1984; Parker et al., 1985; Toffenetti et al., 1987), and tropomyosin (Bautch et al., 1982). All of these *Drosophila* proteins show significant homology to the corresponding proteins of vertebrate muscle.

Despite the fact that all *Drosophila* muscles are striated, these muscles can be divided into two distinct types. The synchronous muscles show a 1:1 ratio between the frequency of excitation and the frequency of contraction as is seen in most vertebrate muscles. In contrast the asynchronous muscles contract with a much greater frequency; this rapid contraction is possible because these muscles are stretch activated (Pringle, 1978). All the larval muscles and most of the adult muscles are synchronous muscles. Only the fibrillar indirect flight muscles (IFM)[1] of the adult thorax are asynchronous; however the IFM make up a large fraction of the thoracic muscle of the adult. The physiological difference between synchronous and asynchronous muscle is reflected at the morphological level; the sarcoplasmic reticulum is very poorly represented in the IFM but is well developed in the synchronous muscles. Under the microscope, the IFM appear more precisely organized than the synchronous muscles; the numerous myofibrils are easily discernable and give the IFM a characteristic fibrillar organization. Generally the synchronous muscles have a high ratio of thin to thick myofilaments (9–12:1) compared to a lower ratio of 6:1 in the asynchronous IFM (for review see Elder, 1975; Crossley, 1978).

In this paper, we describe the cloning and characterization of a gene from *D. melanogaster* that encodes a protein, mp20, that is present only in the synchronous muscles. We have found no evidence that the IFM has any closely related protein and suggest that mp20 has a function that is not needed in the asynchronous muscles of the IFM. Although there are other muscle proteins with isoforms that are specific for particular muscle types (Falkenthal et al., 1985; Bernstein et al., 1986; Fyrberg et al., 1983; Basi et al., 1984), mp20 is the first muscle protein which seems to be lacking in any form in the asynchronous muscle.

Materials and Methods

RNA Analysis

RNA was extracted from primary muscle tissue culture cells as previously described (Storti et al., 1978). For the developmental stages, embryos were

---

P. Lasko's present address is Department of Genetics, University of Cambridge, Cambridge, England CB2 3EH. C. French's present address is Diagnostic Products Corp., Los Angeles, CA 90045.

---

1. *Abbreviations used in this paper*: HRP, horseradish peroxidase; IFM, indirect flight muscles; MHC, myosin heavy chain.

collected for a period of 12 h and then aged for the appropriate time. Larvae were further staged after hatching, and pupae staged after puparium formation. Total nucleic acids were extracted as previously described (Ayme and Tissieres, 1985). The RNA was further purified by precipitation with 3 M sodium acetate.

RNA gel electrophoresis was performed after either glyoxal denaturation (Thomas, 1980), or formaldehyde denaturation (Maniatis et al., 1982). Transfer to Hybond-N filter (Amersham Corp., Arlington Heights, IL) was done without pretreatment using 20× SSC as transfer buffer.

DNA Sequence Determination

The DNA sequence determination was performed using the dideoxy chain-termination technique (Sanger et al., 1977). Commercial mixes (New England Biolabs, Beverly, MA) were used. Nucleic acid and protein sequences were compared using programs developed by the National Biomedical Research Foundation, particularly ALIGN and FASTN, and BEST-FIT from the University of Wisconsin Genetics Computing Group (Devereux et al., 1984).

Genomic DNA Analysis

Genomic DNA was prepared from 1–5 g of embryos. After dechorionation, the embryos were homogenized into 5 ml of buffer I (0.25 M sucrose, 50 mM Tris pH 7.4, 1 mM EDTA, 0.5% Triton-X 100). The nuclei were pelleted by centrifugation at 6,000 rpm for 5 min and then resuspended at 9 ml/g of embryos in buffer II (100 mM NaCl, 50 mM Tris pH 7.4, 100 mM EDTA). The nuclei were lysed by adding SDS to a final concentration of 2% and mixing the solution gently. Proteinase K was added to a final concentration of 50 μg/ml. The solution was mixed and incubated at 50°C for 3–4 h. The proteins were extracted first with phenol then twice with chloroform/isoamyl alcohol (24:24:1). The whole extraction process was repeated at least two more times. The DNA solution was dialyzed extensively against TE (10 mM Tris, 1 mM EDTA, pH 8) before use.

Protein Gel Electrophoresis

One-dimensional SDS–polyacrylamide gel electrophoresis was performed as described by Laemmli (1970). Before loading, the samples were mixed with 0.2 Vol of 5× SDS sample buffer (0.625 M Tris-HCl, pH 6.8, 0.5% SDS, 50% glycerol, 0.025% bromphenol blue). For two-dimensional polyacrylamide gels, the procedure was essentially that of O'Farrell (1975). The proteins, mixed with AMPHO sample buffer (9.5 M urea, 5% β-mercaptoethanol, 2% NP-40, 2% ampholines [LKB Instruments, Inc., Gaithersburg, MD] pH 3.5–10), were separated by isoelectric focusing in the first dimension. The second dimension consisted of an SDS–polyacrylamide gel as above. Tube gels were equilibrated in 65 mM Tris HCl, pH 6.8, 10% glycerol, 5% β-mercaptoethanol, and 2% SDS for 1–2 h and loaded horizontally on top of the stacking gels.

Hybrid Selection and In Vitro Translation

The RNA was hybrid-selected as described in Ricciardi et al. (1979). The in vitro translation was performed in a micrococcal nuclease–treated rabbit reticulocyte lysate prepared by the method of Pelham and Jackson (1976) under the conditions described in Storti et al. (1980). 7 μl of the reaction was then added to 20 μl of either SDS or AMPHO sample buffers for analysis on one- or two-dimensional protein gels.

Dissections

Larvae: dissections were done in a depression microscope slide under PBS (10 mM sodium phosphate, pH 7.0, 0.15 NaCl). Larval body wall was hand dissected by decapitation of the larva with a pair of dissecting needles, followed by involution of the body wall, starting from the posterior end, and removal of the now exposed viscera. The cuticle and the muscle layer that make up the body wall were further separated by gentle homogenization in PBS+ (PBS containing 1% β-mercaptoethanol and 2 mM PMSF [Sigma Chemical Co., St. Louis, MO]) in a Dounce homogenizer (Kontes Glass Co., Vineland, NJ). After homogenization, the cuticles remain intact and can be removed, leaving the muscle in the homogenization buffer. The muscle fraction was fractionated into pellet (containing the nuclei, the mitochondria, and the myofibrils) and supernatant by a 5-min centrifugation at 12,000 g. All the fractions were homogenized in PBS+. Adult: body parts were hand dissected from cold-anesthetized live flies, boiled in PBS+ for

2 min, and homogenized. The IFM was isolated as described in Mogami et al. (1982). The rest of the thorax was used as the source of the synchronous flight muscles. The homogenization was done in PBS+, without boiling. For two-dimensional gels, thoraces were directly homogenized in AMPHO sample buffer and loaded as such on top of the first dimensional tube gel.

Expression of Chimeric Proteins

A 574-bp Xho II/Xho II fragment was isolated from the pupal cDNA clone (see Figs. 1 and 5). This fragment encodes the entire muscle protein minus 5 amino acids at the carboxy-terminal end but adds 13 amino acids at the amino-terminal end since the translation starts at the AUG of the bacterial protein in the vector and continues through 39 nucleotides 5' to the mp20 translation start (see Fig. 5). The Xho II/Xho II fragment was subcloned into two vectors: λ-gtll (lacZ fusion; Young and Davis, 1983; Huynh et al., 1985) and pATH10 (anthranilate synthetase [trpE] fusion; Spindler et al., 1984).

β-galactosidase–linked fusion proteins from the λ-gtll clones were expressed by growing a lysogenic strain of Escherichia coli to an OD 600 of 0.3 in Luria-Bertani broth supplemented with 50 μg/ml ampicillin at 30°C. The cells were heat-shocked at 45°C for 15 min with shaking, then isopropyl-β-thio-galactopyranoside was added to a final concentration of 5 mM. The cells were transferred to 37°C for 1–2 h to continue protein expression. Cells were harvested by centrifugation, resuspended in 0.01 vol of SDS sample buffer, and sonicated for 2 min at 60 W.

Anthranilate synthetase–linked fusion proteins from the pATH10 clones were expressed by growing a 10 ml culture of plasmid-containing cells in M9 medium, supplemented with 5 g/liter casamino acids (Difco Laboratories, Inc., Detroit, MI), 20 μg/ml tryptophan, 10 μg/ml thiamine-HCl, and 50 μg/ml ampicillin at 37°C. When the cells reached mid-logarithmic phase, the culture was diluted into 100 ml of the same medium, lacking tryptophan, and aerated well for 1 h. Indoleacrylic acid (1 mg/ml in ethanol; Sigma Chemical Co.) was added to a final concentration of 10 μg/ml, and the cultures were grown for 4 h more. The cells were collected by centrifugation, resuspended in SDS sample buffer containing 6 M urea, and sonicated as above.

Purification of Chimeric Proteins and Immunization of Rabbits

Chimeric proteins were purified by preparative polyacrylamide gel electrophoresis, followed by electroelution. Gels (7.5 or 10% SDS polyacrylamide) were loaded with 500 μl of solubilized cell lysate and run as described above. After electrophoresis, the gels were soaked in 0.25 M KCl to visualize the proteins. The gel region containing the fusion protein was excised with a razor blade, minced into small pieces, and the protein was recovered using an ISCO (Lincoln, NE) electroelution device. Electroelution was either at 3 W for 3 h, or at 1 W overnight. The protein was then recovered from the eluate by precipitation with 4 vol of acetone at −20°C for 1 h. Recovered protein was checked for purity on Coomassie Blue-stained analytical gels. Rabbits (New Zealand White, female, 5–7 kg) were immunized subcutaneously with 200–300 μg of antigen emulsified with complete Freund's adjuvant. They were then boosted twice more, at days 7 and 17 after the first injection, with the same amount of antigen emulsified with Freund's incomplete adjuvant. Blood (45 ml) was collected at day 24 after the initial injection, and weekly for two more weeks afterward. Serum was collected from the blood by allowing clotting at 37°C for 1 h, and removing the supernatant from a 10-min, 1,500 g centrifuge spin. Sera were stored for up to 2 wk at 4°C with the addition of 0.2% sodium azide, and for longer periods of time at −20°C in 1.5-ml aliquots.

Antibody Probing of Filter-bound Proteins (Western Blot)

Proteins were immobilized on nitrocellulose filters as described by Towbin et al. (1979). Proteins were transferred from gels to nitrocellulose by equilibrating the gel in transfer buffer (25 mM Tris, 150 mM glycine, pH 8.3, 0.1% SDS, 20% methanol) for 30 min, then transferring them electrophoretically in a Trans-blot apparatus (Bio-Rad Laboratories, Richmond, CA) in the same buffer overnight at 0.3 A at 4°C. Filters were air dried before probing.

The filters were blocked for 60 min in TBST (10 mM Tris-HCl, pH 7.5, 150 mM NaCl, 0.1% Tween-20 [Sigma Chemical Co.]), supplemented with 2% BSA. Antibody binding was in TBST + 2% BSA containing a 1:100

A

pEH2.2

160b
20C
5

1kb

B

100bp
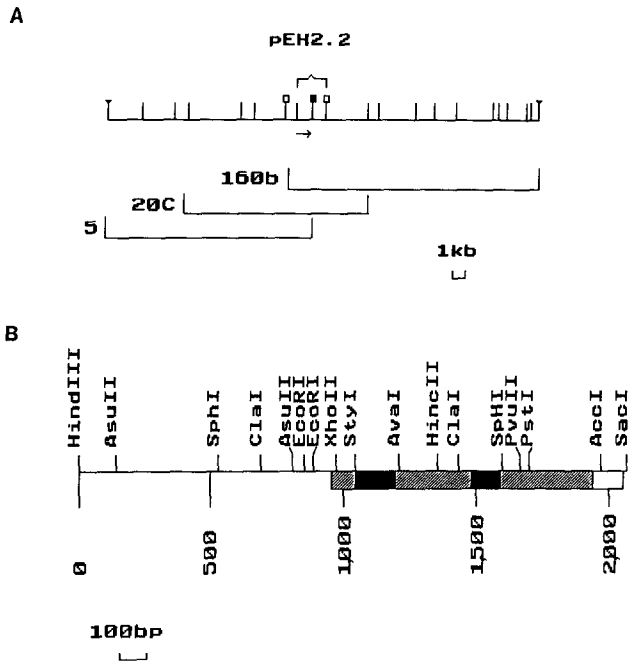
Figure 1. (A) Genomic map of the three overlapping mp20 clones. Only the Eco RI restriction sites are given (indicated by straight lines), except for the Hind III *(open squares)* and Sac I *(solid square)* sites respectively at the 5' and 3' ends of the muscle gene. The arrow, below the map, indicates the direction of transcription for the mp20 gene. *(B)* Detailed restriction map of the Hind III–Sac I fragment containing the mp20 gene. The box represents the transcribed region, the grey-striped segments represent the sequence contained in both cDNA clones, the white segment indicates the region present only in the pupal cDNA clone (and in the larger mp20 transcript). The intron regions are depicted by black segments.

4   12   18

Figure 2. Autoradiogram showing induction of the mp20 gene during differentiation in primary myogenic cultures. PolyA$^+$ RNA from primary cultures was extracted at 4, 12, and 18 h after embryonic cell plating (as described in Storti et al., 1978). The RNA was separated by agarose gel electrophoresis, transferred to nitrocellulose, and hybridized to ^{32}P-labeled insert from the pEH2.2 clone. Two transcripts, 1.0 and 0.9 kb long, are seen only in the 12- and 18-h cultures, after myotubes have formed. The mp20 mRNA is not detectable in the 4-h myoblast culture.

dilution of antiserum for 60 min. The filters were then washed three times for 10 min each in TBST, and then incubated with a 1:7,500 dilution of anti-rabbit IgG alkaline phosphate conjugate (Promega Biotec, Madison, WI) for 30 min in TBST + 2% BSA. The filters were again washed three times for 10 min each in TBST, then blotted damp dry, and placed into the color development substrate solution (0.3 mg/ml nitro blue tetrazolium, 0.150 mg/ml 5-bromo-4-chloro-3-indolyl phosphate [Promega Biotec]; in 100 mM Tris-HCl, pH 9.5, 100 mM NaCl, 5 mM MgCl$_2$). The color reaction was stopped by washing the filters with 20 mM Tris-HCl, pH 8.0, 5 mM EDTA. All incubations were done at room temperature.

Indirect Immunofluorescence

The larval gut was removed in PBS + 0.05% NP-40 and left for 1 min in that solution after removal from the animal. The gut was then fixed for 1 min in 3.7% formaldehyde in PBS, incubated in 45% acetic acid for 5–10 min, and squashed on a subbed slide. The preparation was frozen by immersion in liquid nitrogen and the coverslip was removed with a razor blade. The preparation was dehydrated in 95% ethanol, three times, 10-min each, and air dried. It was then washed two times, 5-min each in PBS, incubated 5 min in PBS + 0.1% Triton X-100, followed by two washes, 10-min each, in PBS. The preparation was blocked in PBS + 0.1% Tween 20 + 10% FCS for 30 min at room temperature. The primary antibody was incubated overnight at 4°C at a dilution of 1:100 in PBS + 10% FCS. After washing two times, 15 min each, in PBS, the secondary antibody (goat anti-rabbit-FITC conjugate) was incubated for 30 min at room temperature at a dilution of 1:300 in PBS. The secondary antibody had been previously absorbed on fixed *Drosophila* Kc cell extract to eliminate nonspecific binding.

Total Protein Detection

Proteins were transferred to nitrocellulose filter as described above. The

membrane was then extensively washed three times, 60 min each, in BT solution (BT: 50 mM Na$_2$B$_4$O$_7$.10 H$_2$O, 0.2% Tween-20, pH 9.3). The washing solution was then replaced with 100 ml of BT containing 20 µl of NHS-biotin (*N*-hydroxysuccinimide biotinate; Bio-Rad Laboratories) for 15 min with constant agitation. The filter was washed with BT alone two times for 5 min each. After this step of protein modification, a regular Western blot was performed starting at the BSA blocking step. After completion of the Western blot analysis a picture was taken as a record of the location of the antigen. The membrane was then treated in order to detect all the proteins modified by the NHS-biotin. 100 ml of TTBS solution (TTBS: 20 mM Tris, 0.5 M NaCl, 0.2% Tween-20) containing 100 µl of avidin-HRP (avidin-linked horseradish peroxidase; Bio-Rad Laboratories) was added to the filter and incubated for 1 h. The membrane was washed two times, 5 min each, in TTBS and two times, 5 min each, in TBS (TTBS without Tween-20). The HRP color development was done with a 0.015% H$_2$O$_2$ solution containing the HRP color development substrate (4-chloro-1-naphthol) in TBS with 20% methanol. All incubations were done at room temperature.

Results

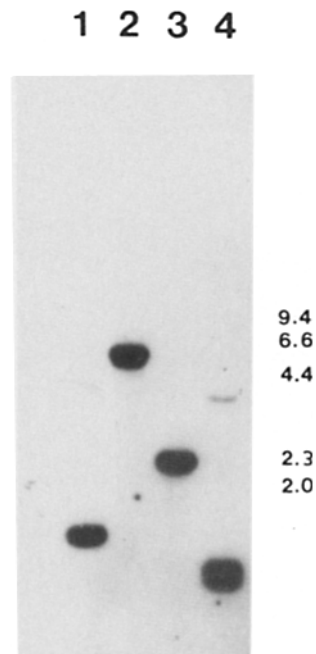### Isolation and Characterization of the mp20 Gene

Messenger RNA was isolated from cultures of pulsating myotubes (Storti et al., 1978) and used to synthesize a labeled cDNA probe. The probe was then depleted of "housekeeping" mRNA by prehybridization to mRNA from Schneider L-2 cells (an undifferentiated cell line). The unhybridized fraction of the probe, therefore enriched for myotube-specific cDNA, was used to screen a library of Charon 4 lambdaphage containing genomic *Drosophila* DNA (Maniatis et al., 1978). The genomic clones identified by this probe were then rescreened individually by hybridization to filters containing RNA extracted from myoblast and myotube cells. All the clones hybridizing specifically to the myotube RNA
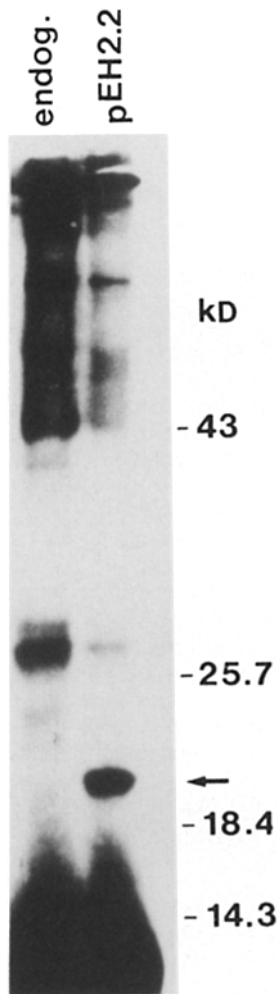
Figure 3. Autoradiograph of [³⁵S]-methionine-labeled proteins translated in a rabbit reticulocyte lysate without (lane *endog.*) and with (lane *pEH2.2*) RNA hybrid selected on DNA from the pEH2.2 plasmid carrying the mp20 gene. The hybrid-selected RNA directs translation of a 20-kD peptide (*arrow*). The lane containing products of the lysate without hybrid-selected RNA (lane *endog.*) has been heavily overloaded to demonstrate the clear lack of the 20-kD peptide in the absence of the mp20 coding sequence. The sizes of molecular mass markers (in kilodaltons) are on the right.

Figure 4. Autoradiograph of a Southern blot of *D. melanogaster* genomic DNA cleaved with several restriction enzymes and hybridized with ³²P-labeled Eco RI–Sac I fragment from the pEH2.2 clone. The lanes are loaded with DNA digested respectively with: (*1*) Eco RI + Sac I; (*2*) Eco RI; (*3*) Eco RI + Hind III; and (*4*) Pvu II + Eco RI. In each case the hybridization detects only the fragments predicted if the genome contained a single copy of the mp20 gene. The faint band at 4 kb in lane *4* is produced by a Pvu II site ∿4 kb from the end of the gene.

were considered as containing presumptive muscle-specific genes. Further studies have shown these clones to include genes coding for tropomyosin (Bautch et al., 1982), and myosin light chain 2 (Toffenetti et al., 1987).

The mp20 gene described here is contained in one of these presumptive muscle-specific clones (referred to here as 160b, see Fig. 1). The transcribed region(s) of the clone were mapped by probing restriction digests of the phage 160b DNA with end-labeled RNA from 18–24 h myotubes. Only one site of RNA hybridization was mapped to a 1.2-kb Eco RI/Sac I fragment. Because this fragment is located at the end of the 160b insert it seemed possible that the 160b clone might not include all sequences of interest. Therefore the *Drosophila* genomic library was rescreened with the Eco RI/Sac I fragment. Only two additional genomic clones were identified (named 20C and 5); both overlapped the initial clone. The three clones represent together 33 kb of contiguous *Drosophila* sequence. The restriction map of this region is depicted in Fig. 1. Analysis of corresponding cDNA clones (see below) later showed, however, that the complete transcribed region of the gene is included within the Eco RI/Sac I fragment of the original clone. A 2.2-kb Eco RI/Hind III fragment, including the transcribed region, was subcloned into pBR322 to yield plasmid pEH2.2 (see Fig.

1). RNA complementary to this subclone encodes the protein that we now refer to as mp20 (see below). There is only one copy of the mp20 gene in the 33 kb of cloned DNA.

The pattern of expression of the mp20 gene corresponds to that expected of a muscle-specific gene. DNA from pEH2.2 was used to probe polyadenylated RNA isolated from primary embryonic culture cells, before and after muscle differentiation (Storti et al., 1978). A significant accumulation of mp20 RNA is seen in cells that have undergone differentiation to myotubes but no complementary RNA is detected in undifferentiated myoblasts (4 h after plating, Fig. 2). Two transcripts, of ∿1.0 and 0.9 kb, can be distinguished and both are found only after myotube differentiation has begun.

The protein encoded by the pEH2.2 clone was initially identified by in vitro translation of the RNA hybrid-selected by this clone DNA. The RNA was translated in a rabbit reticulocyte lysate, and the products were separated on a one-dimensional protein gel (Fig. 3). When RNA complementary to pEH2.2 DNA is added to the lysate, we detect the synthesis of a new polypeptide, with an apparent molecular mass of ∿20 kD. This polypeptide is not found in products of control lysates, even when increased amounts of radioactivity are loaded on the gel (Fig. 3, lane *endog.*). When the products of the in vitro translation are separated by two-dimensional gel electrophoresis, the mp20 protein migrates with an apparent molecular mass of ∿20 kD and a pI of ∿5.5 (data not shown); however, this in vitro–translation product differs in pI from the in vivo–synthesized protein (see below) and thus must be differently modified in vitro compared with in vivo. On two-dimensional gels, we detect a second product of the in vitro translation. This polypeptide has the same pI as the 20-kD polypeptide but is significantly smaller. It is not resolved from the endogenous labeling of the in vitro translation system on a one-dimensional gel. Repetitions of the RNA selection and in vitro–translation experiments with

```
                    .     -600          .        .        .        .     -550          .
              gtacatgtcatttttgttaaactttttatgcagaaagtcctttaataagaaattgccaatgtcgctgatctg
                    .            -500        .        .        .        .            -450
atctgataactgtgaataattgtgaggatgtccatgtcatcataaatttcgctgtaaattctggaggatgtgtcgctcgatgggaaccaa
              .        .        .        -400        .        .        .
tttgctgccatttttaattatgacccaatgcatgccgatggaaattacgggcaacgaatgcacgcgcccgaaatgagttcaaaatgacag
        -350        .        .        .        .        -300        .        .
caggacagcaagtcccaaaaaatgcggctaacatgcaaatgagcagtactttgaagagcacaatctacacacactcgcacacacacacac
        -250  .        .        .        .        .        -200        .
acacacataccatcgattggaaatgtgtgaaaataatgaagcatttggatttttccttttcctttttctcaaccgatgatgatctaaaattataac
                -150        .        .        .        .        .        -100    EcoRI
gcacgattcatacgcgggagacgggcaaagtcggcgtttcgaatatatagaacgcccacgctgcgaaatcttcacacactcgactccgaattc
                .EcoRI -50        .        .        .        .        .
ggatatattttggaaaattcttgaacgaattcgtcggagaacgctctctcgcgctttaaagccgcatacgccccgttggctggc
1        .    T        .        .XhoII        .        50        .        .        .
ATCAGTGTTAGCTTGTCAGAGCTTTAGTGAAGATCCCGCAGGACCCGAAACCAAAAACCAAGAATCAAACATGTCTCTTGAGCGTGCCGTT
                                                                MetSerLeuGluArgAlaVal
        100        .        .        .        .        150        .        .
CGTGCCAAGgtgaatactcaatcagtgtagaaaaaagatccttaagaaaacacacatataaacgaaaatatatgcaagaaatgtgaaaca
ArgAlaLys
                200        .        .        .        .        G   250        .
aagagttttgtgaaatattgtgcctaaaatatgaattttttcttatccgatccacctagATTGCCAGCAAGCGGAATCCCGAGATGGAC
                                                                IleAlaSerLysArgAsnProGluMetAsp
                        300        .        .        .        T   350        .
AAGGAGGCCCAGGAGTGGATCGAGGCCATCATTGCCGAGAAGTTCCCCGCCGGCCAGTCCTACGAGGATGTGCTCAAGGACCGTCAGGTG
LysGluAlaGlnGluTrpIleGluAlaIleIleAlaGluLysPheProAlaGlyGlnSerTyrGluAspValLeuLysAspGlyGlnVal
                .'.        .        400        .        .        .        .        450
CTGTGCAAACTGATCAACGTGCTGTCGCCCAATGCCGTGCCCAAGGTCAACTCCTCGGGCGGCCAGTTCAAGTTCATGGAGAACATCAAC
LeuCysLysLeuIleAsnValLeuSerProAsnAlaValProLysValAsnSerSerGlyGlyGlnPheLysPheMetGluAsnIleAsn
                        .        500        .        .        .        .
AACTTCCAGAAGGCCCTGAAGGAGTACGGTGTGCCCGACATCGATGTCTTCCAGACCGTCGATCTGTACGAGAAGAAGGATATTGCCAAC
AsnPheGlnLysAlaLeuLysGluTyrGlyValProAspIleAspValPheGlnThrValAspLeuTyrGluLysLysAspIleAlaAsn
        550        .        .        .        .        600        .        .
GTTACCAACACCATCTTCGGTTTGGGCCGTGCCgtaagtttcaaatcgccttctatcccattgtcatccaggctaatccttgaattgcggc
ValThrAsnThrIlePheAlaLeuGlyArgAla
        650        .        .        .        .        700        .        .
tcattgcagACCTACAAGCATGCCGACTTCAAGGGTCCCTTCCTGGGCCCCAAGCCCGCCGATGAGTGCAAGCGCGATTTCACCGAAGAG
        ThrTyrLysHisAlaAspPheLysGlyProPheLeuGlyProLysProAlaAspGluCysLysArgAspPheThrGluGlu
                        750        .        .        .        .        800
CAGCTGAAGGCTGGCCAGACCATTGTGGGTCTGCAGGCCGGTTCCAACAAGGGAGCCACCCAGGCCTGGCCAGAACCTCGGCGCTGGCCGC
GlnLeuLysAlaGlyGlnThrIleValGlyLeuGlnAlaGlySerAsnLysGlyAlaThrGlnAlaGlyGlnAsnLeuGlyAlaGlyArg
XhoII    .        .        .        850        .        .        .        900
AAGATCCTGCTCGGCAAGTAAGCGCCAAAGGATGGCCAGGATGTCCACACCCTTTTCTACACTTATGCTAAGTGAACACACCCATATATA
LysIleLeuLeuGlyLys *
        .  T        .        .        .        950↓  ↓        .        .        .
TTTTGTATGACGAAAAACATGAAGAACACAATGATATTCCTCACCAAAGCAAAACTCAACACCAATGAACACACACTCACACAAAACAAC
    1000        .        .        .        .        1050        .        .
ACATAAAAGCAACGCAAAGGCTTGTCGGCCTTATATTAAGTATACGCGTAAGTTAATCTATTTATGAGTACAATAAAGTAATTCAAAAAAC
            1100        .    SacI
CCAGTTTTCCACTTTCcgatcacctctataaccttgagctc
```

Figure 5. DNA and protein sequence of mp20. The capital letters represent the cDNA sequence, the lower case letters represent the untranscribed and intron regions. Nucleotide +1 is the beginning of the two cDNA clones. Some of the major restriction sites are marked. The putative TATA box and the polyadenylation signal are underlined. The end of the shorter cDNA clone is indicated by an arrow. The amino acid sequence is shown below the DNA sequence. The underlined amino acids are part of the two calcium binding sites. The sequence shown for the transcribed regions is taken from the cDNA clones (which have no nucleotide differences in the regions in which they overlap). There are four nucleotide differences between the genomic and cDNA sequences; the changes for the genomic clone are indicated above the cDNA sequence. Two of the four nucleotide changes are within the coding region: (*a*) a change from GGC in genomic DNA to AGC in the cDNA, with a change from glycine to serine (amino acid 13); and (*b*) a change from CTT in genomic DNA to CTC in the cDNA (amino acid 45). The second change is silent at the translational level since both codons encode leucine. Both the genomic and cDNA clones were derived from the Canton-S wild-type strain, suggesting that this variation represents a polymorphism present in that population.

the same RNA preparation give very different ratios of the two products and it seems likely that the smaller polypeptide is a premature termination product of the reticulocyte lysate, as has been proposed for other genes (Toffenetti et al., 1987).

The gene encoding mp20 appears to be a single copy gene with no closely related sequences in the *D. melanogaster* genome. To show this, the Eco RI/Sac I fragment from pEH2.2 was used to probe genomic *Drosophila* DNA that had been digested by several different restriction enzymes. As shown in Fig. 4, only a single band, whose size corresponds to that predicted by the restriction map of the phage clones, is detected with any of the cleavages, except when enzymes that cut within the mp20 region are used (see Fig. 1 for the restriction map). The hybridization stringency used was that which allowed cross-hybridization of the multiple members within the two tubulin gene families (Mischke and Pardue, 1982) and thus should have detected any gene(s) closely related to mp20, if such genes existed.

In situ hybridization of pEH2.2 reveals complementary sequences at a single site on larval salivary gland polytene chromosomes in region 49F 7-15 (data not shown). Using a set of new deficiencies within the vestigial region (Lasko and Pardue, 1988), the hybridization was narrowed to the 49F 9-13 region.

Isolation of cDNA Clones and Determination of the Encoded Protein Sequence

The entire insert from the phage 20C was purified and used to isolate clones from two λ-gt10 cDNA libraries made from polyadenylated RNA of late embryos or late pupa (Poole et

al., 1985). Two cDNA clones were isolated from both the pupal and embryonic libraries. Both of the pupal inserts are ∼900 bp long as determined by PAGE, while the embryonic inserts are each ∼150 bp shorter. The nucleotide sequence of the pupal and embryonic clones, along with 1.7 kb of the corresponding genomic DNA, has been determined and is diagrammed in Fig. 5.[2]

Both the embryonic and pupal cDNAs begin at the same nucleotide at the 5' end, suggesting that they are in fact full-length representations of their corresponding messages. RNase protection experiments confirm that this nucleotide is the position of the transcription start site (data not shown). Moreover, a TTTAAAA sequence, which can be used as a TATA box in *Drosophila* (Eveleth et al., 1986) lies 25–31 bp upstream from the start of the cDNAs (see Fig. 5). When the cDNA sequences are compared with that of genomic DNA, it is evident that the gene contains two introns: one of 140 bp between amino acids 10 and 11 of the open reading frame, and one of 67 bp between amino acids 121 and 122. The splicing pattern of the pupal clone is identical to that of the embryonic clone, and both introns possess consensus junctions (Mount, 1982). The clones differ only at the 3' end, the embryonic clone terminating 142–146 bp upstream from the end of the pupal clone (since nucleotides 142–146 are all A residues, it is impossible to determine precisely where the addition of the poly(A) tail begins). The sequences of both
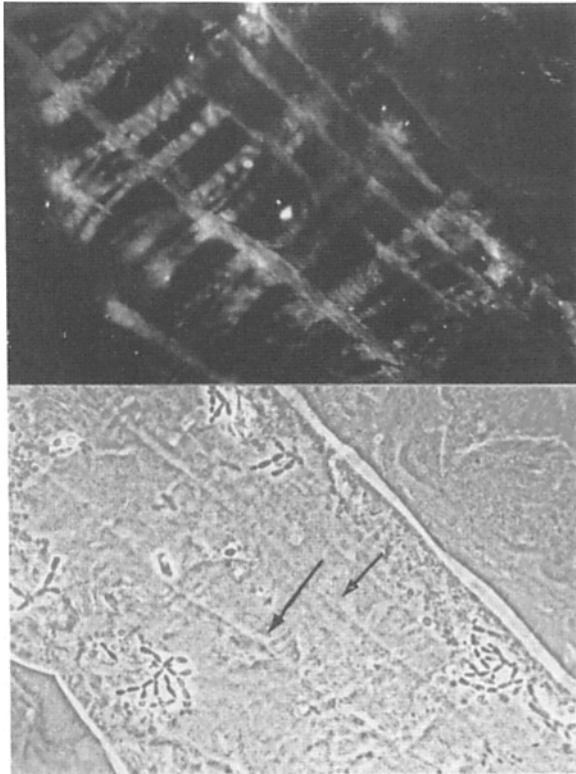
Figure 6. Immunofluorescence demonstrates the presence of the mp20 protein in the muscle fibers but not in the parenchymal cells of the gut. The larval gut was squashed on a slide and treated as described in Materials and Methods. The mp20 protein was localized by indirect immunofluorescence using an antibody directed against an in vitro–synthesized mp20 protein. The top picture represents the FITC fluorescence, showing the protein localization to the meshwork of muscle fibers surrounding the gut. The bottom picture is a phase micrograph of the same region, showing the presence of a layer of parenchymal gut cells which do not bind the mp20 antibody. The muscle fibers are faintly visible (*arrows*).

the embryonic and pupal cDNA clones possess a polyadenylate stretch at the 3′ end, implying that the different 3′ ends are, in fact, representative of the actual *Drosophila* messages. The pupal polyadenylation occurs 30 bases from a consensus AATAAA polyadenylation site (Proudfoot, 1982), while the shorter embryonic transcript does not have a precise match at the appropriate location. The region 15–30 bases upstream of the termination site of the shorter transcript is highly A rich, however, and has numerous near matches to the polyadenylation consensus sequence.

It is noteworthy that the difference in size of the two types of cDNA clones is similar to the length difference between the two transcripts seen by hybridization to myotube RNA. We have confirmed the correspondence of the cDNA clones to the transcripts by Northern blot hybridization to larval RNA which has good amounts of both transcripts (see Developmental Expression below). From the structure of the cDNA clones, a region of 142 nucleotides is present only at the 3′-untranslated region of the larger cDNA clone. When a Northern blot containing RNA extracted from third instar larvae was hybridized with a probe corresponding to the 142

nucleotides unique to the large cDNA, only the large transcript showed hybridization (data not shown). As a control, a similar filter was hybridized with a probe of the same length but originating from the coding region. Hybridization of the second probe was detected over both transcripts, showing that either transcript could be detected by a small probe if the probe sequence was present in the transcript (data not shown). From this analysis we conclude that the two cDNA clones are representative of the two mRNAs and that the two transcripts differ only in length of the 3′ region. The longer transcript is the more abundant one in the pupal stages so it is not surprising that both cDNA clones from the pupal library represent this transcript.

The open reading frame found in the mp20 sequence would code for a polypeptide of 184 amino acids, beginning from the first AUG codon (Fig. 5). This protein would have a molecular mass of 20,166 D. Two stretches of 12 amino acids that match the proposed consensus sequence for the calcium binding site of calcium-binding proteins (the so-called E-F hand; Tufty and Kretsinger, 1975) are found between residues 20–31 and 93–104 (Fig. 5, underlined amino acids). Extensive computer analysis did not reveal any other significant degree of homology with any protein sequence present in the database that we used (NBRF protein database).

Characterization of the mp20 Protein Product

To raise antisera against mp20, we used two different fusion proteins, made by linking most of the mp20 coding sequence to parts of the sequence of two bacterial proteins. Four different antisera were obtained and tested by binding to *Drosophila* thoracic proteins immobilized on filters. All four sera recognize the same protein of 20 kD, the size predicted from both the DNA sequence and the in vitro–translation product (data not shown). Although some of the antisera show other sites of minor binding on the filters, the 20-kD protein is always the most prominent binding and the only one shared by all four sera.

The muscle specificity of mp20 has been demonstrated further by indirect immunofluorescence using these antisera on squashes of various tissues. A dramatic example is shown in Fig. 6 where antibody binding is detected only in the muscle fibers surrounding the larval gut but not in any of the intestinal parenchymal cells that make up the gut.

Mogami et al. (1982) have published photographs of two dimensional gels of *Drosophila* thoracic proteins or IFM proteins that serve as a standard for identifying *D. melanogaster* muscle proteins. We have used similar two-dimensional gels to determine which spot of the Mogami nomenclature is recognized by the antisera against the mp20 fusion protein. Thoracic proteins were run on a two-dimensional gel, transferred to nitrocellulose filters, and tested by Western blot. The antibody detects a single major spot on two-dimensional gels of ∼20 kD. To identify the spot corresponding to mp20 it was necessary to be able to detect the other proteins on the same filter after the antibody binding was performed. In our experiments, we are working with small amounts of many proteins and we have developed a very sensitive technique for allowing detection of the remainder of the proteins after the antibody reaction (see Materials and Methods for a complete description). By comparison with the gel pattern of Mogami et al. (1982) we have assigned the mp20 protein as spot 170 (Fig. 7). The pI of spot 170 is
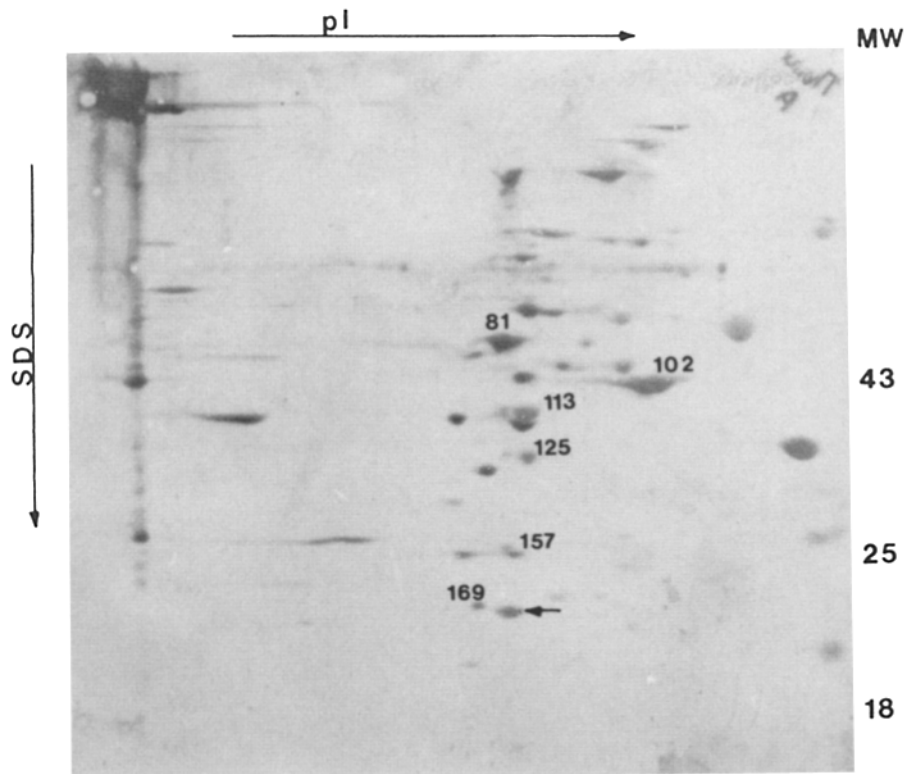
Figure 7. Identification of the mp20 protein. Proteins from two *D. melanogaster* thoraces were run on a two-dimensional gel and transferred to nitrocellulose as described in Materials and Methods. The immobilized proteins were modified with NHS-biotin. A regular Western blot analysis was then performed with serum raised against the anthranilate synthetase fusion protein of mp20 and the position of the mp20 spot is indicated by the arrow. The total modified proteins were then revealed by reaction with horseradish peroxidase–linked avidin. A few proteins are numbered according to the system of Mogami et al. (1982). The positions of known molecular mass markers (in kilodaltons) are indicated on the right. The mp20 binds to spot 170 of Mogami et al. (1982). It has an apparent molecular mass of 20 kD and a pI of 7.0.

determined by Mogami as ~7, therefore it seems that the in vivo product differs in pI from the in vitro–translation product (pI 5.5, see above), possibly due to posttranslational modifications either in vivo or in the lysate.

In an attempt to provide clues to mp20 function, we have used antibody binding to study both the partition of mp20 and its distribution into various types of muscles. In these experiments a monoclonal antibody raised against the myosin heavy chain (MHC) was used as a control for the purity of the fractions and for the total amount of muscle protein present (myosin heavy chain is localized into the myofibrillar structure of all types of muscle). The partition of mp20 has been analyzed by differential extraction of the muscles from the larval body wall. Fig. 8 shows the analysis of different protein fractions prepared from third instar larvae. Each lane of the gel is loaded with material derived from an equal number of larvae and the relative amounts of mp20 and MHC are assayed by binding of their respective antibodies. The mp20, together with the MHC, is present in the body wall (composed mainly of cuticle and associated hypodermal muscles). After the subfractionation of the body wall, both mp20 and MHC are found in the muscle homogenate. However, when the muscle cells are disrupted and further fractionated by differential centrifugation, all of the mp20 is present in the supernatant fraction whereas the MHC is associated predominantly with the pellet that contains the nuclei, the mitochondria, and the myofibrillar apparatus. From this we conclude that mp20 is not tightly associated with the structural apparatus of the myofibril and is not a nuclear or mitochondrial protein but is either a cytoplasmic or membrane-associated protein.

The larvae contains only synchronous muscles but the adult has, in addition to many synchronous muscles, the asynchronous indirect flight muscles. To investigate the distribution of mp20 in these two different muscle types we have used antibody binding to different adult body parts. The MHC antibody was used this time as a control for the proportion of muscle in each body fraction. mp20 is present in equal amounts in the head and legs, with approximately a twofold increase in the thorax and a fourfold increase in the abdomen. MHC follows approximately the same distribution, except in the thoraces where the amount of MHC is at least ten times greater than in any of the other fractions (data not shown). Comparison of the amounts of MHC and mp20 detected in these experiments indicates that mp20 is significantly underrepresented in the thoracic muscles. Thoraces were further dissected to free the IFM from the other structures of the thorax, including the synchronous flight muscles. Both the IFM and the remainder were analyzed by Western blotting for the presence of mp20. As shown in Fig. 9, most of the mp20 of the thorax is recovered in the synchronous muscle fraction; there is no mp20 in the IFM fraction, although parallel gels show much myosin in the IFM. We conclude from these experiments that mp20 is not present in the asynchronous indirect flight muscles in any significant amount and that mp20 must therefore be a synchronous muscle-specific protein.

Developmental Expression of the mp20 Gene

The temporal pattern of expression of the mp20 gene is consistent with the temporal pattern of muscle development in *Drosophila*. Fig. 10 shows that transcripts become detectable to a high level only partway into embryonic development and
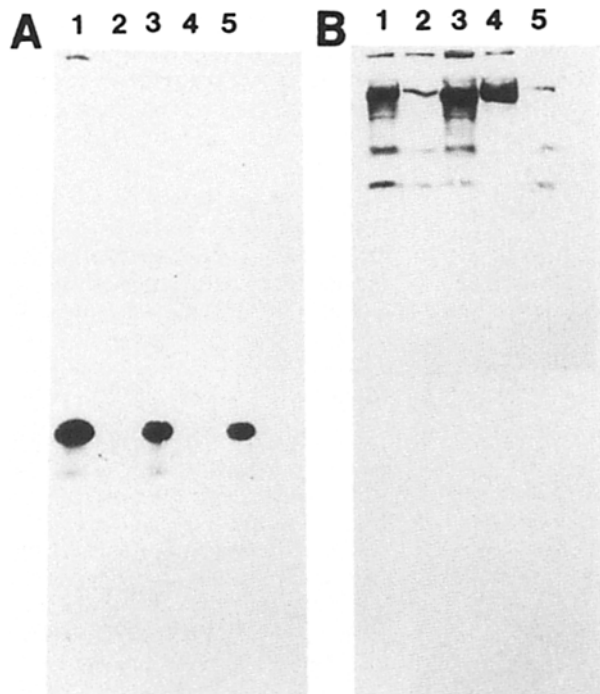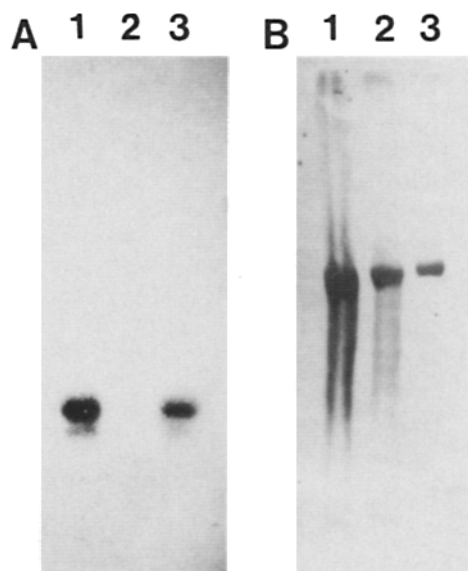
Figure 8. Western blot analysis of different fractions of larval body wall performed with (*A*) the antibody against mp20 and (*B*) the antibody against myosin heavy chain. Aliquots of the same protein fractions were run on a 12 (*A*) or a 10% (*B*) polyacrylamide gel. Materials from two larvae was loaded on each lane. Lanes show: (*1*) larval body wall, (*2*) cuticle, (*3*) hypodermal muscle, (*4*) pellet from fractionated muscle containing the nuclei, mitochondria, and myofibrillar apparatus, and (*5*) supernatant from fractionated muscle. The mp20 antibody (*A*) detects a 20-kD peptide present in the larval body wall (which contains cuticle plus muscle) and in the hypodermal muscle. When muscle is fractionated, the mp20 is extracted into the supernatant (lane *5*) rather than being bound to the nuclei, mitochondria or myofibrillar apparatus (lane *4*). The myosin heavy chain antibody detects peptides of ∼200 kD (plus some cleavage products) in the larval body wall and the hypodermal muscle. The low level of antibody binding in lane *2* indicates that most of the muscle was removed in preparing the cuticle. As expected of a myofibrillar protein, when muscle is fractionated (lanes *4* and *5*), most of the myosin is detected in the myofibrillar pellet rather than the cytosolic supernatant from the cells.

continue to be found at a moderately high level throughout the larval development. Transcripts regress in prepupae, disappear in 1-, 2-, and 3-d pupae, and reappear in 4- and 5-d pupae. It is also evident in Fig. 10 that the two mp20 transcripts are regulated differently during development. The two mRNAs are present in approximately equal amounts in all the larval stages. In the embryos, pupae, and adult the two transcripts are present but the amount of the 1.0-kb transcript is much greater than that of the 0.9-kb transcript.

Discussion

The data presented here on the distribution of mp20 and its mRNA demonstrate the muscle-specific expression of the mp20 gene. Transcription in primary culture muscle cells becomes detectable only after differentiation to myotubes has begun; no mp20 mRNA is detected in an undifferentiated cell


Figure 9. Western blot analysis of the different thoracic muscles performed with (*A*) the antibody against mp20 and (*B*) the antibody against myosin heavy chain. Protein fractions were run on a 12 (*A*) or a 6% (*B*) polyacrylamide gel. The equivalent of two adults was used per lane in *A* and *B*. Lanes contain: (*1*) complete thorax, (*2*) IFM, and (*3*) thorax after the removal of the IFM (synchronous tubular muscles mainly). The mp20 antibody detects a 20-kD protein that is present in the thorax but only in the synchronous (tubular) muscles. The myosin antibody detects peptides of ∼200 kD in both the IFM and the synchronous tubular muscle.

line. In cytological analysis of mixed cell types, antibody binding to the protein is detected only in muscle cells. When larval body parts are analyzed biochemically, the protein is detected only in fractions that contain muscle. The profile of mp20 RNA accumulation during development is consistent with the periods of muscle differentiation and formation; mp20 transcripts are abundant during the embryonic stages when larval muscle is synthesized and during the late pupal stages when adult muscle is made. Our analysis of the developmental pattern of expression of the mp20 gene was performed on RNA from whole organisms and thus represents an average of all muscles. The pattern is comparable to that determined for other muscle-specific proteins such as actin, tropomyosin, and myosin light chain 2 (Bautch et al., 1982; Fyrberg et al., 1983; Parker et al., 1985; Toffenetti et al., 1987). Among the actin genes, the act87E gene, encoding the actin isoform found in all muscles except the thoracic muscles, shows a pattern of expression which is very much like that of the mp20 gene (Fyrberg et al., 1983).

Two different mRNAs encode mp20; the relative ratio of the two transcripts varies during development. These RNAs correspond to the two different cDNA clones that were isolated and therefore the transcripts differ only in the length of their 3'-noncoding region and both yield the same primary translation product. The significance, if any, of this variation in length of the nontranslated region is unknown; however it appears to be common for *D. melanogaster* muscle protein genes to have identical mRNAs differing only in the choice of their polyadenylation sites (Basi et al., 1984; Parker et al., 1985; Bernstein et al., 1986; Bond and Davidson, 1986;
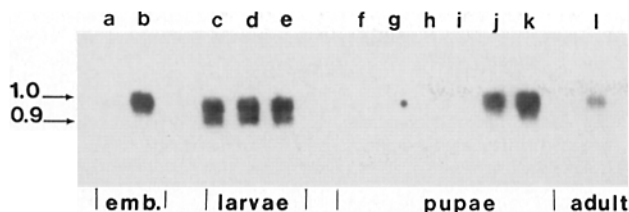
Figure 10. Developmental Northern blot analysis of the mp20 gene. Total RNA was extracted from various developmental stages and run on a 1.5% agarose-formaldehyde gel. The filter was hybridized with the oligolabeled Eco RI–Sac I fragment from pEH 2.2. The lanes have RNA from: (*a*) 0–12-h (*b*) 12–24-h embryos, (*c*) 1st instar larvae, (*d*) 2nd instar larvae, (*e*) 3rd instar larvae, (*f*) prepupae, (*g*) 1-d-, (*h*) 2-d-, (*i*) 3-d-, (*j*) 4-d-, (*k*) 5-d pupae, (*l*) adult. Only the relevant section of the gel is shown. Two transcripts of 1.0 and 0.9 kb (*arrows*) are detected. Transcripts are detected in the later embryonic stages and throughout larval life. No mp20 mRNA is seen in early pupae but transcripts again become abundant in the last 2 d of pupal life. The levels of the two transcripts are approximately equal in the larval stages. The 0.9-kb RNA is much less abundant than the 1.0-kb RNA in the embryonic, pupal, and adult stages.

Vigoreaux and Tobin, 1987). The relative abundance of the two mRNAs may reflect a preferential choice of one of the two polyadenylation sites during the precursor processing, a difference in the relative stability of the two mRNAs or a tissue specificity which is not detected when RNA is made from whole organisms.

Genes for several muscle proteins of *Drosophila* have now been cloned. All of the major myofibrillar proteins that have been sequenced, actin, tropomyosin, myosin heavy chain, myosin light chain 2, and myosin light chain 3, show a high level of sequence homology with their vertebrate counterparts (Fyrberg et al., 1981; Bernstein et al., 1983; Basi et al., 1984; Karlik et al., 1984; Falkenthal et al., 1985; Parker et al., 1985; Toffenetti et al., 1987). In contrast, the mp20 protein sequence shows no strong homology with any known protein but has two regions with similarity to a number of known calcium binding sites. Each of these putative calcium binding regions consists of 12 amino acids (residues 20–31 and 93–104) which show a good alignment to the E–F hand

structure proposed by Tufty and Kretsinger (1975) as the calcium binding site. The E–F hand structure was first determined based on the conformation and sequence of regions in parvalbumin and troponin C. In these two proteins the calcium binding regions contain 12 amino acids of which seven positions are critical (Coffee and Bradshaw, 1973; Kretsinger and Nockolds, 1973; Collins et al., 1973; Herzberg and James, 1985; see Fig. 11). Five of the seven critical residues are present in both sites of the mp20 sequence (Fig. 11). There are other examples of calcium binding proteins which have, at their known calcium binding sites, an E–F hand structure but do not conserve all seven of these critical residues. Fig. 11 presents some examples of such proteins. The EDTA myosin light chain from scallop (homology 5 out of 7 residues) is the regulatory subunit for the actomyosin interaction (Szent-Gyorgyi et al., 1973; Kendrick-Jones, 1974; Kendrick-Jones and Jakes, 1977) and plays a role equivalent to troponin C in these muscles. The calcium binding to trypsinogen and trypsin is supposed to play a role in the maturation and the stability of this protease. The calcium binding region has been demonstrated to correspond to the position of the E–F hand structure (Bode and Schwager, 1975; homology 5 out of 7 residues). Other examples include villin (Hesterberg and Weber, 1983), fibrinogen (Doolittle, 1984; Dang et al., 1985), and S100 protein (Baudier et al., 1984). These sequence analogies suggest that mp20 is a calcium binding protein although this must be confirmed biochemically.

A second difference between mp20 and the other cloned *Drosophila* muscle proteins lies in the restriction of mp20 to particular muscles. Two very different types of muscle are present in adult *Drosophila*; the asynchronous indirect flight muscles, which are the specific and predominant muscles of the thorax; and the synchronous muscles, present in every part of the adult and larva. All the myofibrillar proteins studied so far are common in every muscle type, although most of these proteins have specific isoforms which are restricted to one or more muscle types. In contrast, our protein localization studies show that mp20 is absent from the asynchronous oscillatory IFM, although it is present in the rest of the muscles in both larva and adult. One very notable difference between the IFM and all the other adult and larval muscles lies in the way in which the frequency of contraction is determined. In most muscles there is a 1:1 ratio between the fre-

	X	Y	Z	-Y		-X			-Z		
Asp	*	N	*	N	Gly	*	Ile	N	*	*	Glu	
Asp	Lys	Glu	Ala	Gln	Glu	Trp	Ile	Glu	Ala	Ile	Ile	Drosophila mp-20 site I
Asp	Ile	Asp	Val	Phe	Gln	Thr	Val	Asp	Leu	Tyr	Glu	" " site II
Asp	Gln	Asp	Lys	Ser	Gly	Phe	Ile	Glu	Glu	Asp	Glu	Carp Parvalbumin site I
Asp	Ser	Asp	Gly	Asp	Gly	Lys	Ile	Gly	Val	Asp	Glu	" " site II
Asp	Ala	Asp	Gly	Gly	Gly	Asp	Ile	Ser	Val	Lys	Glu	rabbit troponin-C site I
Asp	Glu	Asp	Gly	Ser	Gly	Thr	Ile	Asp	Phe	Glu	Glu	" " site II
Asp	Arg	Asn	Ala	Asp	Gly	Tyr	Ile	Asp	Pro	Glu	Glu	" " site III
Asp	Lys	Asn	Asn	Asp	Gly	Arg	Ile	Asp	Phe	Asp	Glu	" " site IV
Asp	Arg	Asp	Gly	Asp	Gly	Cys	Ile	Thr	Thr	Met	Glu	chicken calmodulin site I
Asp	Ala	Asp	Gly	Ser	Gly	Thr	Ile	Asp	Phe	Pro	Glu	" " site II
Asp	Lys	Asp	Gly	Asn	Gly	Tyr	Ile	Ser	Ala	Ala	Glu	" " site III
Asp	Cys	Asn	Asn	Asp	Gly	Gln	Val	Asn	Tyr	Glu	Glu	" " site IV
Asp	Val	Asp	Arg	Asp	Gly	Phe	Val	Ser	Lys	Asp	Asp	scallop EDTA myosin light chain
Leu	Gly	Glu	Asp	Asn	Ile	Asn	Val	Val	Glu	Gly	Asn	Glu	trypsinogen
Asp	Asn	Asp	Asn	Asp	Lys	Phe	Glu	Gly	Asn	Cys	Ala	Glu	fibrinogen
Gly	Val	Asp	Pro	Ser	Arg	Lys	Glu	Asn	His	Leu	Ser	villin
Gly	Arg	Glu	Gly	Asp	Lys	His	Lys	Leu	Lys	Lys	Ser	Glu	S100 site I
Asp	Asn	Asp	Gly	Asp	Gly	Glu	Cys	Asp	Phe	Gln	Glu	" site II

Figure 11. Comparison of known calcium binding sequences from various proteins with amino acids 20–31 (site I) and 93–104 (site II) of *Drosophila* mp20. The E–F hand consensus sequence is given at the top. *N* represents any of the following amino acids: Glu, Gln, Asp, Asn, Ser, Thr; the asterisks represent any amino acid. X, Y, Z, -X, -Y, -Z refer to the position of the calcium ion coordination octahedron (Tufty and Kretsinger, 1975).

quency of excitation from the motor neurons and the frequency of contraction (hence, synchronous); in the IFM, however, contraction frequency is controlled by an intrafibrillar stretch–activation mechanism, so that the frequency of contraction greatly exceeds the frequency of excitation (hence, asynchronous; Pringle, 1978, 1981). The variation in the control of contraction between synchronous and asynchronous muscle is reflected in visible structural differences in the membrane systems (transverse T system and sarcoplasmic reticulum) involved in the linking of excitation to contraction (see Smith, 1984, for review). The synchronous muscles rely on a well-developed sarcoplasmic reticulum to pump the calcium out of and into the sarcoplasm at each contraction–relaxation cycle; the asynchronous muscles do not rely entirely on this mechanism to continue contraction. In addition, the metabolism of these two muscle types, as determined by their energy requirement, must be different (Crabtree and Newsholme, 1975). It seems likely that the function of the mp20 protein might be concerned with creating and/or participating in the specific physiology of the synchronous muscles.

The two potential calcium binding sites in the sequence and lack of tight association with the myofibrillar apparatus suggest the possibility that mp20 might be the *Drosophila* functional equivalent of one of the calcium binding proteins that are found in the sarcoplasm surrounding the myofibrils or in the sarcoplasmic reticulum: vertebrate parvalbumin (Nockolds et al., 1972), invertebrate sarcoplasmic calcium binding proteins (Wnuk et al., 1982), or calsequestrin (Cozens and Reithmeier, 1984). It seems possible that mp20 is involved in the system linking the nerve impulse with the contraction or in the relaxation process. Both of these systems are different in synchronous and asynchronous muscles and therefore may involve different specific proteins, one of them being mp20 in the synchronous muscle.

References

Ayme, A., and A. Tissieres. 1985. Locus 67B of Drosophila melanogaster contains seven, not four, heat shock genes. *EMBO (Eur. Mol. Biol. Organ.) J.* 4:2949-2954.

Basi, G. S., M. Boardman, and R. V. Storti. 1984. Alternative splicing of a Drosophila tropomyosin gene generates muscle tropomyosin isoforms with different carboxy-terminal ends. *Mol. Cell. Biol.* 4:2828-2836.

Baudier, J., N. Glasser, K. Haglid, and D. Gerard. 1984. Purification, characterization and ion binding properties of human brain S100b protein. *Biochim. Biophys. Acta.* 790:164-173.

Bautch, V. L., R. V. Storti, D. Mischke, and M. L. Pardue. 1982. Organization and expression of Drosophila tropomyosin genes. *J. Mol. Biol.* 162:231-250.

Bernstein, S. I., C. J. Hansen, K. D. Becker, D. R. Wassenberg II, E. S. Roche, J. J. Donady, and C. P. Emerson, Jr. 1986. Alternative RNA splicing generates transcripts encoding a thorax-specific isoform of Drosophila myosin heavy chain. *Mol. Cell. Biol.* 6:2511-2519.

Bernstein, S. I., K. Mogami, J. J. Donady, and C. P. Emerson. 1983. Drosophila muscle myosin heavy chain encoded by a single gene in a cluster of muscle mutations. *Nature (Lond.).* 302:393-397.

Bode, W., and P. Schwager. 1975. The refined crystal structure of bovine beta trypsin at 1.8 Å resolution. II. Crystallographic refinement, calcium binding site, benzamide binding site and active site at pH 7.0. *J. Mol. Biol.* 98:693-717.

Bond, B. J., and N. Davidson. 1986. The Drosophila melanogaster actin 5C gene uses two transcription initiation sites and three polyadenylation sites to express multiple mRNA species. *Mol. Cell. Biol.* 6:2080-2088.

Coffee, C. J., and R. A. Bradshaw. 1973. Carp muscle calcium-binding protein. I. Characterization of the trypsin peptides and the complete amino acid sequence. *J. Biol. Chem.* 248:3305-3312.

Collins, J. H., J. D. Potter, M. J. Horn, G. Wilshire, and N. Jockman. 1973. The amino acid sequence of rabbit skeletal muscle troponin C: gene replication and homology with calcium-binding proteins of carp and hake muscle. *FEBS (Fed. Eur. Biochem. Soc.) Lett.* 36:268-272.

Cozens, B., and R. A. F. Reithmeier. 1984. Size and shape of rabbit skeletal muscle calsequestrin. *J. Biol. Chem.* 259:6248-6254.

Crabtree, B., and E. A. Newsholme. 1975. Comparative aspects of fuel utilization and metabolism by muscle. *In* Insect Muscle. P. N. R. Usherwood, editor. Academic Press, Inc., NY. 405-500.

Crossley, A. C. 1978. Morphology and development of the Drosophila muscular system. *In* The Genetics and Biology of Drosophila. Vol. IIb. M. Ashburner and T. R. F. Wright, editors. Academic Press, Inc., NY. 499-560.

Dang, C. V., R. F. Ebert, and W. R. Bell. 1985. Localization of a fibrinogen calcium binding site between gamma-subunit positions 311 and 336 by terbium fluorescence. *J. Biol. Chem.* 260:9713-9719.

Devereux, J., P. Haekerli, and O. Smithies. 1984. A comprehensive set of sequence programs for the VAX. *Nucleic Acids Res.* 12:387-395.

Doolittle, R. F. 1984. Fibrinogen and fibrin. *Annu. Rev. Biochem.* 53:195-229.

Elder, H. Y. 1975. Muscle Structure. *In* Insect Muscle. P. N. R. Usherwood, editor. Academic Press, Inc., London. 1-14.

Eveleth, D. D., R. D. Gietz, C. A. Spencer, F. E. Nargang, R. B. Hodgetts, and J. L. Marsh. 1986. Sequence and structure of the dopa decarboxylase gene of Drosophila: evidence for novel RNA splicing variants. *EMBO (Eur. Mol. Biol. Organ.) J.* 5:2663-2672.

Falkenthal, S., V. P. Parker, and N. Davidson. 1985. Developmental variations in the splicing pattern of transcripts from the Drosophila gene encoding myosin alkali light chain result in different carboxyl-terminal amino acid sequences. *Proc. Natl. Acad. Sci. USA.* 82:449-453.

Falkenthal, S., V. P. Parker, W. W. Mattox, and N. Davidson. 1984. Drosophila melanogaster has only one myosin alkali light-chain gene which encodes a protein with considerable amino acid homology to chicken myosin light chains. *Mol. Cell. Biol.* 4:956-965.

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

Fyrberg, E. A., J. W. Mahaffey, B. J. Bond, and N. Davidson. 1983. Transcripts of the six Drosophila actin genes accumulate in a stage- and tissue-specific manner. *Cell.* 33:115-123.

Fyrberg, E. A., B. J. Bond, N. D. Hershey, K. S. Mixter, and N. Davidson. 1981. The actin genes of Drosophila: protein encoding regions are highly conserved but intron positions are not. *Cell.* 24:107-116.

Fyrberg, E. A., K. L. Kindle, N. Davidson, and A. Sodja. 1980. The actin genes of Drosophila: a dispersed multigene family. *Cell.* 19:365-378.

Herzberg, O., and M. N. G. James. 1985. Structure of the calcium regulatory muscle protein troponin C at 2.8 Å resolution. *Nature (Lond.).* 313:653-659.

Hesterberg, L. K., and K. Weber. 1983. Demonstration of three distinct calcium-binding sites in villin, a modulator of actin assembly. *J. Biol. Chem.* 258:365-369.

Huynh, T. V., R. A. Young, and R. W. Davis. 1985. Constructing and screening cDNA libraries in lambda gt10 and lambda gt11. *In* DNA Cloning: A Practical Approach. D. M. Glover, editor. IRL Press, Oxford, England. 49-78.

Karlik, C. C., J. W. Mahaffey, M. D. Coutu, and E. A. Fyrberg. 1984. Organization of contractile protein genes within the 88F subdivision of the D. melanogaster third chromosome. *Cell.* 37:469-481.

Kendrick-Jones, J. 1974. Role of myosin light chains in calcium regulation. *Nature (Lond.).* 249:631-634.

Kendrick-Jones, J., and R. Jakes. 1977. Myosin-linked regulation: a chemical approach. *In* International Symposium on Myocardial Failure. G. Reicker, A. Weber, and J. Goodwin, editors. Springer-Verlag, New York. 28-40.

Kretsinger, R. H., and C. E. Nockolds. 1973. Carp muscle calcium-binding protein. II. Structure determination and general description. *J. Biol. Chem.* 248:3313-3326.

Laemmli, U. K. 1970. Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature (Lond.).* 270:680-685.

Lasko, P. F., and M. L. Pardue. 1988. Studies of the genetic organization of the 49D-49F region of D. melanogaster. *Genetics.* 120:495-502.

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

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 eukaryotic DNA. *Cell.* 15:687-701.

Mischke, D., and M. L. Pardue. 1982. Organization and expression of alpha-

tubulin genes in Drosophila melanogaster. *J. Mol. Biol.* 156:449–466.

Mogami, K., S. C. Fujita, and Y. Hotta. 1982. Identification of Drosophila indirect flight muscle myofibrillar proteins by means of two-dimensional electrophoresis. *J. Biochem.* 91:643–650.

Mount, S. 1982. A catalogue of splice junction sequences. *Nucleic Acids Res.* 10:459–472.

Nockolds, C. E., R. H. Kretsinger, C. J. Coffee, and R. A. Bradshaw. 1972. Structure of a calcium-binding carp myogen. *Proc. Natl. Acad. Sci. USA.* 69:581–584.

O'Farrell, P. H. 1975. High resolution two-dimensional electrophoresis of proteins. *J. Biol. Chem.* 250:4007–4021.

Parker, V. P., S. Falkenthal, and N. Davidson. 1985. Characterization of the myosin light chain-2 gene of Drosophila melanogaster. *Mol. Cell. Biol.* 5:3058–3068.

Pelham, H. R. B., and R. J. Jackson. 1976. mRNA dependent translation system from reticulocytes. *Eur. J. Biochem.* 67:247–256.

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

Pringle, J. W. S. 1981. The Bidder Lecture: The evolution of fibrillar muscle in insects. *J. Exp. Biol.* 94:1–14.

Pringle, J. W. S. 1978. The Croonian Lecture, 1977: Stretch activation of muscle: function and mechanism. Proc. R. Soc. Lond. B. 201:107–130.

Proudfoot, N. 1982. The end of the message. *Nature.* 298:516–517.

Ricciardi, R. P., J. S. Miller, and B. E. Roberts. 1979. Purification and mapping of specific RNAs by hybridization-selection and cell-free translation. *Proc. Natl. Acad. Sci. USA.* 76:4927–4931.

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

Sanger, F., S. Nicklen, and A. R. Coulson. 1977. DNA sequencing with chain-terminating inhibitors. *Proc. Natl. Acad. Sci. USA.* 74:5463–5467.

Smith, D. S. 1984. The structure of insect muscles. *In* Insect Ultrastructure. King, R. C., and H. Akai, editors. Plenum Publishing Corp., NY. 111–150.

Spindler, K. R., D. S. F. Rosser, and A. J. Berk. 1984. Analysis of adenovirus transforming proteins from early regions 1A and 1B with antisera to inducible fusion antigens produced in Escherichia coli. *J. Virol.* 49:132–141.

Storti, R. V., M. P. Scott, A. Rich, and M. L. Pardue. 1980. Translational control of protein synthesis in response to heat shock in D. melanogaster cells. *Cell.* 22:825–834.

Storti, R. V., S. J. Horovitch, M. P. Scott, A. Rich, and M. L. Pardue. 1978. Myogenesis in primary cell cultures from Drosophila melanogaster: protein synthesis and actin heterogeneity during development. *Cell.* 13:589–598.

Szent-Gyorgyi, A. G., E. M. Szentkiralyi, and J. Kendrick-Jones. 1973. The light chains of scallop myosin as regulatory subunits. *J. Mol. Biol.* 74: 179–203.

Thomas, P. S. 1980. Hybridization of denatured RNA and small DNA fragments transferred to nitrocellulose. *Proc. Natl. Acad. Sci. USA.* 77:5201–5204.

Tobin, S. L., E. Zulauf, F. Sanchez, E. A. Craig, and B. J. McCarthy. 1980. Multiple actin-related sequences in the Drosophila melanogaster genome. *Cell.* 19:121–131.

Toffenetti, J. L., D. Mischke, and M. L. Pardue. 1987. Isolation and characterization of the gene for myosin light chain two of Drosophila melanogaster. *J. Cell Biol.* 104:19–28.

Towbin, H., T. Staehelin, and J. Gordon. 1979. Electrophoretic transfer of proteins from polyacrylamide gels to nitrocellulose sheets: procedure and some applications. *Proc. Natl. Acad. Sci. USA.* 76:4350–4354.

Tufty, R. M., and R. H. Kretsinger. 1975. Troponin and parvalbumin calcium binding regions predicted in myosin light chain and T4 lysozyme. *Science (Wash. DC).* 187:167–169.

Vigoreaux, J. O., and S. L. Tobin. 1987. Stage-specific selection of alternative transcriptional initiation sites from the 5C actin gene of Drosophila melanogaster. *Genes and Development.* 1:1161–1171.

Wnuk, W., J. A. Cox, and E. A. Stein. 1982. Parvalbumins and other soluble high affinity calcium-binding proteins from muscle. *In* Calcium and Cell Function. Vol. II. W. Y. Cheung, editor. Academic Press Inc., NY. 243–278.

Young, R. A., and R. W. Davis. 1983. Efficient isolation of genes by using antibody probes. *Proc. Natl. Acad. Sci. USA.* 80:1194–1198.