# Molecular Genetic Characterization of a Developmentally Regulated Human Perinatal Myosin Heavy Chain

**Rebecca Feghali and Leslie A. Leinwand**

Department of Microbiology and Immunology, Albert Einstein College of Medicine, Bronx, New York 10461

*Abstract.* We have isolated a human cDNA which corresponds to a developmentally regulated sarcomeric myosin heavy chain. RNA hybridization and DNA sequence analysis indicate that this cDNA, called SMHCP, encodes a perinatal myosin heavy chain isoform. The nucleotide and deduced amino acid sequences of the 3.4-kb cDNA insert show strong homology with other sarcomeric myosin heavy chains. The strongest homology is to a previously described 970-bp cDNA encoding a rat perinatal isoform (Periasamy, M., D. F. Wieczorek, and B. Nadal-Ginard. 1984. *J. Biol. Chem.* 259:13573–13578). The homology between the analogous human and rat perinatal myosin heavy chain cDNAs is maintained through the highly isoform-specific final 20 carboxyl-terminal amino acids, as well as the 3′ untranslated region. Ribonuclease protection studies show that the mRNA encoding this isoform is expressed at high levels in 21-wk fetal skeletal tissue and not in fetal cardiac muscle. In contrast to the rat perinatal isoform, which was not found to be expressed in adult hind-leg tissue, the gene encoding SMHCP continues to be expressed in adult human skeletal tissue, but at lower levels relative to fetal skeletal tissue.

THE myosin molecule, through its interaction with actin, generates movement in processes as diverse as cytokinesis and muscle contraction. The hexameric myosin molecule consists of a pair of myosin heavy chains (MHC)[1] and two pairs of myosin light chains. The heavy chains contain the site of ATP hydrolysis and the sequences that comprise the thick filament. Vertebrate sarcomeric MHCs are encoded by multigene families encompassing 10–15 genes (11, 35) each encoding a distinct protein, or isozyme. Members of this gene family show both tissue-specific and developmentally regulated expression (10, 19, 34). During development, fetal, neonatal, and adult MHC isozymes are expressed in a sequential program (1, 3, 14, 32, 33). However, the functional significance of each of these isozymes, and the mechanisms by which their expression is regulated, have yet to be clarified. It may be that changes in the physiological properties of the developing muscle necessitate switches in MHC isozyme expression to accommodate new demands.

cDNA and genomic clones corresponding to several sarcomeric MHC isoforms have been isolated and found to show strong homology to each other. These include sequences from chicken (18), rat (15, 20, 27, 35), rabbit (4, 8), mouse (31), and human (23, 24) striated muscle. When sequences of analogous MHC isoforms from different species are compared, they are even more homologous than intraspecies comparisons, suggesting functional constraints on the divergence of isoform sequences. The 3′ untranslated region (UTR) sequences are also maintained across species when analogous isoforms are compared, and appear to be quite isoform specific (19, 20).

The patterns of MHC isoform expression in muscle fiber types during vertebrate skeletal muscle development have been investigated using a variety of cyto- and immunochemical techniques. Using anti-myosin monoclonal antibodies, Silberstein and Blau (25) have shown that in human skeletal muscle at 17 wk of gestation, two populations of myotubes begin to emerge; all fibers react with an anti-fetal MHC antibody, while a small subpopulation also reacts with an anti-slow MHC antibody. By 30 wk of gestation, a population of fibers which is stained by anti-fetal and anti-fast antibodies has emerged, and some slow fibers no longer express any fetal MHC; only a small percentage of the total fibers do not yet express either fast or slow MHC. By birth, few fibers still express a fetal isoform. By one year, no fibers appear to express fetal MHC, and the distribution of fast and slow fibers changes very little as compared to adult tissue (30).

To study the progression of MHC gene expression during human muscle development, we have isolated cDNA clones corresponding to four human sarcomeric myosin gene products. These include adult-fast and slow-skeletal muscle MHCs (22, 23) along with a human fetal skeletal cDNA clone (Karsch-Mizrachi, I., M. Travis, H. Blau, and L. Leinwand, manuscript submitted for publication). In this pa-

This sequence has been registered with the EMBL/GenBank, accession number Y00821.

1. *Abbreviations used in this paper*: MHC, myosin heavy chain; UTR, untranslated region.

A

```
Q K T K D E L A K S E A K R K E L E E K
CAGAAAACCAAAGATGAACTCGCCAAGTCAGAGGCAAAACGGAAGGAGCTAGAGGAAAAA
        10        20        30        40        50        60

M V T L L K E K N D L Q L Q V Q S E A D
ATGGTCACTCTCTTAAAAGAGAAAAATGACCTGCAACTCCAGGTTCAATCTGAAGCAGAT
        70        80        90       100       110       120

S L A D A E E R C E Q L I K N K I Q L E
AGCTTGGCTGATGCAGAGGAAAGGTGTGAGCAACTGATTAAAAACAAAATCCAACTTGAG
       130       140       150       160       170       180

A K I K E V T E R A E E E E E I N A E L
GCCAAAATCAAAGAGGTGACTGAAAGAGCTGAGGAGGAGGAAGAGATCAATGCTGAGCTG
       190       200       210       220       230       240

T A K K R K L E D E C S E L K K D I D D
ACAGCCAAGAAGAGAAAACTGGAGGATGAATGTTCAGAACTCAAGAAAGACATTGATGAC
       250       260       270       280       290       300

L E L T L A K V E K Q K H A T E N K V K
CTTGAGCTGACACTGGCCAAGGTTGAGAAGCAGAAACATGCCACGGAGAACAAGGTGAAA
       310       320       330       340       350       360

N L T E E M A G L D E T I A K L S K E K
AATCTTACAGAAGAGATGGCAGGCCTGGATGAAACCATTGCAAAACTGTCCAAGGAGAAG
       370       380       390       400       410       420

K A L Q E T H Q Q T L D D L Q A E E D K
AAGGCTCTCCAAGAGACCCACCAGCAGACCCTGGATGACCTGCAGGCAGAGGAGGACAAA
       430       440       450       460       470       480

V N I L T K A K T K L E Q Q V D D L E G
GTCAACATCCTGACCAAAGCTAAAACCAAGCTAGAACAGCAAGTGGATGATCTTGAAGGG
       490       500       510       520       530       540

S L E Q E K K L R M D L E R A K R K L E
TCTCTGGAACAAGAAAAGAAGCTTCGAATGGATCTAGAAAGAGCAAAGCGGAAACTGGAG
       550       560       570       580       590       600

G D L K L A Q E S T M D M E N D K Q Q L
GGTGACCTCAAATTGGCCCAAGAATCCACAATGGATATGGAAAATGACAAACAGCAACTT
       610       620       630       640       650       660

D E K L E K K E F E I S N L I S K I E D
GATGAAAAGCTTGAAAAGAAAGAATTTGAAATCAGCAATTTGATAAGCAAAATTGAAGAT
       670       680       690       700       710       720

E Q A V E I Q L Q K K I K E L Q A R I E
GAGCAAGCTGTAGAAATTCAACTACAGAAGAAGATCAAAGAGTTGCAGGCCCGCATTGAG
       730       740       750       760       770       780

E L G E E I E A E R A S R A K A E K Q R
GAGCTGGGGGAAGAAATCGAGGCAGAGAGGGCGTCCCGAGCCAAAGCGGAGAAGCAGCGC
       790       800       810       820       830       840

S D L S R E L E E I S E R L E E A G G A
TCTGACCTCTCCCGGGAACTGGAGGAGATCAGCGAGAGGCTGGAAGAAGCCGGTGGGGCA
       850       860       870       880       890       900

T S A Q V E L N K K R E A E F Q K L R R
ACTTCTGCTCAGGTGGAATTGAACAAGAAGCGGGAGGCTGAGTTTCAGAAACTGCGCAGG
       910       920       930       940       950       960

D L E E A T L Q H E A M V A A L R K K H
GACCTGGAGGAGGCCACCCTGCAGCATGAAGCTATGGTGGCTGCTCTTCGGAAGAAGCAC
       970       980       990      1000      1010      1020

A D S M A E L G E Q I D N L Q R V K Q R
GCAGACAGTATGGCTGAGCTTGGGGAGCAGATTGACAACTTGCAGCGGGTCAAACAGAAG
      1030      1040      1050      1060      1070      1080

L E K E K S E L K M E T D D L S S N A E
CTGGAGAAGGAGAAGAGTGAGCTGAAGATGGAGACTGATGACCTCAGCAGTAACGCAGAG
      1090      1100      1110      1120      1130      1140

A I S K A K G H L E K M C R S L E D Q V
GCCATTTCCAAAGCCAAGGGCCACCTTGAAAAGATGTGCCGCTCTCTAGAAGATCAAGTG
      1150      1160      1170      1180      1190      1200

S G L K T K E E E Q Q R L I N D L T A Q
AGTGGGCTTAAGACCAAGGAAGAGGAGCAGCAGCGGCTGATCAATGACCTCACAGCACAG
      1210      1220      1230      1240      1250      1260

R A R L Q T E A G E Y S R Q L D E Q D A
AGAGCGCGCCTGCAGACAGAAGCGGGTGAATATTCTCGACAATTAGATGAGCAAGATGCT
      1270      1280      1290      1300      1310      1320

L V S Q L S R S K Q A S T Q Q I E E L K
TTAGTCTCTCAGCTTTCAAGGAGCAAGCAAGCATCTACTCAGCAGATTGAAGAGCTGAAA
      1330      1340      1350      1360      1370      1380

H Q L E E E T K A K N A L A H A L Q S S
CATCAACTAGAGGAAGAAACTAAAGCCAAGAACGCCCTGGCACACGCCCTGCAGTCCTCC
      1390      1400      1410      1420      1430      1440

R H D C D L L R E Q Y E E E Q E G K A E
CGCCATGACTGCGACCTGCTGCGGGAACAGTATGAGGAAGAGCAGGAAGGCAAAGCTGAG
      1450      1460      1470      1480      1490      1500

L Q R A L S K A N S E V A Q W R T K Y E
CTGCAGAGGGCGCTGTCCAAGGCCAACAGTGAGGTTGCCCAGTGGAGAACCAAATACGAG
      1510      1520      1530      1540      1550      1560

T D A I Q R T E E L E E A K K K L A Q R
ACGGATGCCATCCAGCGCACAGAGGAGCTGGAGGAGGCCAAGAAAAAGTTGGCCCAGCGC
      1570      1580      1590      1600      1610      1620

L Q E A E E H V E A V N A K C A S L E K
CTGCAAGAAGCTGAGGAACATGTAGAAGCTGTGAACGCCAAATGTGCTTCCCTTGAGAAG
      1630      1640      1650      1660      1670      1680

T K Q R L Q N E V E D L M L D V E R S N
ACGAAGCAGCGGCTCCAGAATGAAGTTGAAGACCTCATGCTTGATGTGGAAAGGTCTAAT
      1690      1700      1710      1720      1730      1740

A A C A A L D K K Q R N F D K V L S E W
GCAGCCTGTGCAGCCCTTGATAAGAAGCAAAGGAACTTTGACAAGGTCCTATCAGAATGG
      1750      1760      1770      1780      1790      1800

K Q K Y E E T Q A E L E A S Q K E S R S
AAGCAGAAGTATGAGGAAACTCAGGCTGAACTTGAGGCCTCCCAGAAGGAGTCACGTTCT
      1810      1820      1830      1840      1850      1860

L S T E L F K V K N V Y E E S L D Q L E
CTTAGCACTGAGCTGTTCAAGGTGAAGAATGTCTATGAGGAATCCCTGGATCAACTCGAA
      1870      1880      1890      1900      1910      1920

T L R R A H K N L Q Q E I S D L T E Q I
ACGCTAAGAAGAGCACATAAGAACTTGCAACAGGAGATTTCTGACCTCACTGAGCAGATT
      1930      1940      1950      1960      1970      1980

A E G G K Q I H E L E K I K K Q V E Q E
GCAGAGGGAGGAAAGCAAATTCATGAATTGGAGAAAATAAAGAAGCAAGTAGAACAAGAG
      1990      2000      2010      2020      2030      2040

K C E I Q A A L E E A E A S L E H E E G
AAATGTGAAATTCAGGCTGCTTTAGAGGAAGCAGAGGCATCTCTTGAACATGAAGAAGGA
      2050      2060      2070      2080      2090      2100

K I L R I Q L E L N Q V K S E V D R K I
AAGATTCTGCGTATCCAGCTTGAGTTAAACCAAGTCAAGTCTGAAGTTGATAGAAAAATC
      2110      2120      2130      2140      2150      2160

A E K D E E I D Q L K R N H T R V V E T
GCAGAAAAGGATGAGGAAATTGACCAGCTGAAGAGAAACCACACTAGAGTCGTGGAGACA
      2170      2180      2190      2200      2210      2220

M Q S T L D A E I R S R N D A L R V K K
ATGCAGAGCACGGCTGGATGCAGAGATTAGAAGCAGAAATGATGCTCTGAGAGTCAAGAAG
      2230      2240      2250      2260      2270      2280

K M E G D L N E M E I Q L N H A N R L A
AAAATGGAAGGAGATCTGAATGAAATGGAAATCCAGCTGAACCATGCCAATCGCTTAGCT
      2290      2300      2310      2320      2330      2340

A E S L R N Y R N T Q G I L K E T Q L H
GCAGAGAGTTTAAGGAACTACAGGAACACCCAAGGAATCCTGAAGGAAACCCAGCTCCAC
      2350      2360      2370      2380      2390      2400

L D D A L R G Q E D L K E Q L A I V E R
CTGGATGATGCTCTCCGGGGCCAGGAGGACCTCAAGGAACAGCTGGCAATTGTGGAGCGC
      2410      2420      2430      2440      2450      2460

R A N L L Q A E I E E L W A T L E Q T E
AGAGCCAACCTGCTGCAGGCTGAGATCGAGGAGCTGTGGGCCACTCTGGAACAGACAGAG
      2470      2480      2490      2500      2510      2520

R S R K I A E Q E L L D A S E R V Q L L
AGAAGCAGGAAAATCGCCGAACAGGAGCTCCTGGATGCCAGTGAGCGTGTCCAGCTCCTC
      2530      2540      2550      2560      2570      2580

H T Q N T S L I N T K K K L E N D V S Q
CACACCCAGAATACCAGTCTCATTAACACCAAGAAGAAATTAGAAAATGACGTTTCCCAA
      2590      2600      2610      2620      2630      2640

L Q S E V E E V I Q E S R N A E E K A K
CTCCAAAGTGAAGTGGAAGAAGTAATCCAAGAATCACGCAATGCAGAAGAGAAAGCCAAG
      2650      2660      2670      2680      2690      2700

K A I T D A A M M A E E L K K E Q D T S
AAGGCCATCACTGATGCTGCCATGATGGCTGAGGAGCTGAAGAAGGAACAGGACACCAGC
      2710      2720      2730      2740      2750      2760

A H L E R M K K N L E Q T V K D L Q H R
GCCCACCTGGAGCGGATGAAGAAGAACCTGGAGCAGACGGTGAAGGACCTGCAGCATCGT
      2770      2780      2790      2800      2810      2820

L D E A E Q L A L K G G K K Q I Q K L E
CTAGATGAGGCCGAGCAGCTGGCGCTGAAGGGTGGGAAGAAGCAGATCCAGAAACTGGAG
      2830      2840      2850      2860      2870      2880

A R V R E L E G E V E N E Q K R N A E A
GCCAGGGTACGTGAGCTTGAAGGAGAGGTTGAAAATGAACAGAAACGTAATGCAGAGGCT
      2890      2900      2910      2920      2930      2940

V K G L R K H D R R V K E L T Y Q T E E
GTTAAAGGTTTACGGAAACATGACCGACGAGTAAAAGAACTCACCTACCAGACTGAAGAA
      2950      2960      2970      2980      2990      3000

D R K N V L R L Q D L V D K L Q A K V K
GATCGCAAGAATGTTCTCAGGCTGCAGGACTTGGTAGATAAATTACAGGCGAAGGTGAAA
      3010      3020      3030      3040      3050      3060

S Y K R Q A E E A E E Q S N A N L S K F
TCATACAAGAGACAAGCTGAGGAGGCTGAGGAACAATCCAATGCTAATCTATCTAAATTC
      3070      3080      3090      3100      3110      3120

R K L Q H E L E E A E E R A D I A E S Q
CGCAAACTCCAGCATGAGCTGGAGGAGGCCGAGGAACGGGCTGACATTGCTGAGTCCCAG
      3130      3140      3150      3160      3170      3180

V N K L R V K S R E V H T K I S A E *
GTCAACAAATTGCGAGTGAAGAGCCGAGAGGTTCACACAAAAATCAGTGCAGAGTAAACA
      3190      3200      3210      3220      3230      3240

CACCTGCCTGATGCTATCAAGAGGCTGAAGAAAGGCACAAAATGTGCTATTTTTGGTCAC
      3250      3260      3270      3280      3290      3300

TTGCTTTATGACGTTTATTTTCCTGTTAAAGCTGAATAAATAAAAACTACAGTAAATGTA
      3310      3320      3330      3340      3350      3360

TACATTAAAAAAAA
      3370
```

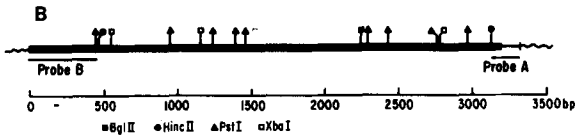


*Figure 1. A* shows the DNA sequence and predicted encoded amino acids of the cDNA, SMHCP. The coding and 3′ UTR nucleotide sequence and derived amino acid sequence are shown. The termination codon is shown as *. *B* shows the restriction map of pSMHCP. Restriction endonuclease sites are indicated for the cloned insert. The solid bar corresponds to protein coding sequences. The solid line corresponds to the 3′ UTR. Wavy lines correspond to vector sequence. Regions used to generate probes A and B are indicated.

per we report the isolation and characterization of a 3.4-kb cDNA clone, SMHCP, encoding the rod of a developmentally regulated human skeletal muscle myosin isoform. The complete nucleotide sequence of this cDNA has been determined and compared to MHC sequences from several organisms. Although all sarcomeric MHCs are highly homologous, it appears that the final 20 amino acids encoded by MHC mRNAs, as well as the 3′ UTRs, contain isoform-specific sequences which are maintained across divergent species. DNA sequence comparisons in these regions show that SMHCP is the human equivalent of a previously described rat perinatal myosin isoform cDNA (20). We show that the human perinatal MHC gene is expressed predominantly in fetal skeletal muscle after 21 wk of gestation, and, in contrast to expression patterns seen in the rat, continues to be expressed in adult skeletal tissue.

## Materials and Methods

### Construction and Isolation of Recombinant cDNA Clones

A cDNA library was constructed in the λ gt10 vector system following standard procedures (12) using 5 μg poly $(A)^+$ mRNA isolated from 21-wk human fetal heart tissue (provided by H. Blau, Stanford University School of Medicine). cDNA clones containing MHC sequences were identified by plaque filter hybridization (5) using a nick translated (21) 2.0-kb human cDNA clone previously shown to encode an adult human β/slow MHC isoform (23).

### Restriction Endonuclease Mapping and DNA Sequence Analysis

The cDNA insert was subcloned into the plasmid vector pTZ19R (Pharmacia Fine Chemicals, Piscataway, NJ). Restriction endonuclease digestions were carried out following the conditions suggested by Maniatis (16).

Nested deletions, for sequencing the insert of pSMHCP, were generated using the Erase-A-Base system (Promega Biotec, Madison, WI) (6). Double-stranded plasmid DNA was sequenced using the Sequenase DNA sequencing kit (U.S. Biochemical Corp., Cleveland, OH) following the method of Kraft et al. (9). Because of the overlapping nature of the deletions each nucleotide was sequenced an average of two times.

### Preparation of Probes

To generate the gene-specific radiolabeled RNA probe of pSMHCP, probe A (Fig. 1 *B*), a 194-bp Hinc II/Eco RI fragment (which includes 54 nucleotides encoding the final 18 amino acids and the entire 3′ UTR) was subcloned into pGEM-1 (Promega Biotec). Following the manufacturer's suggested protocol, 1 μg of linearized plasmid DNA was used to generate radiolabeled anti-sense cRNA using uridine 5′-[α-$^{32}$P] triphosphate (10 mCi/ml) (Amersham Corp., Arlington Heights, IL) and T7 RNA polymerase (Promega Biotec). When used in the RNase protection studies, the probe was further purified by digesting the DNA template with RNase-free DNase for 15 min at 37°C.

Probe B (Fig. 1 *B*), a 481-bp fragment, was generated from an Eco RI/Hinc II digest of pSMHCP and radioactively labeled with [α-$^{32}$P] dCTP using a random-primer oligolabeling kit (Pharmacia Fine Chemicals).

The β/slow MHC specific probe was generated from a subclone of a cDNA previously characterized (23). A 238-bp fragment, containing 54 bp encoding the final 18 amino acids of the MHC, 117 bp of its 3′ UTR, a poly (A) tail of 20 nucleotides, and 47 bp derived from Okayama-Berg plasmid was subcloned into pGEM-1. Linearized DNA from this subclone was used to generate an anti-sense radiolabeled cRNA in a manner identical to that described above for probe A.

### RNA Isolation

Total cellular RNA was isolated from autopsy samples of fetal and adult cardiac and skeletal tissues (gifts from Dr. S. Kohtz, Mt. Sinai College of Medicine and from Dr. M. Thompson, University of Pittsburgh, School of Medicine) using a guanidine isothiocyanate extraction, cesium chloride pelleting protocol (16).

### Genomic DNA Analysis and Hybridization Conditions

Human genomic DNA was digested with restriction endonucleases, and 30 μg per lane was electrophoresed in 1% agarose-TAE (40 mM Tris-acetate, 2 mM EDTA) gels. DNA was transferred to filters according to Southern (26) except that GeneScreen (New England Nuclear, Boston, MA) was used in place of nitrocellulose. The DNA was fixed to the filters by a 5-min exposure to shortwave UV light followed by baking for 1 h at 80°C.

For random-primed probe hybridizations, the filters were prehybridized for at least 2 h in 5× SSC (20× is 3 M NaCl, 300 mM sodium citrate), 1× Denhardt's solution, 50 mM $NaH_2PO_4 \cdot H_2O$ and 150 μg/ml heat denatured salmon sperm DNA. Filters were hybridized overnight at 65°C in the same solution with the addition of dextran sulfate to a final concentration of 10%, and 2 × $10^5$ cpm/ml radiolabeled probe. After hybridization, filters were washed in 2× SSC, 0.2% SDS for 1 h at the hybridization temperature.

For hybridizations with labeled cRNA probes, filters were prehybridized for at least 2 h in 50% formamide, 5× SSPE (20× is 3.6 M NaCl, 200 mM $NaH_2PO_4 \cdot H_2O$, 20 mM EDTA, pH 7.4), 2× Denhardt's solution, and 0.2% SDS. Filters were hybridized overnight at 60°C in the same solution with the addition of 1 × $10^6$ cpm/ml of $^{32}$P-radiolabeled cRNA. Filters were washed in 1× SSPE for 1 h at 65°C. All filters were exposed to X-Omat AR film (Eastman Kodak Co., Rochester, NY) with intensifying screens (Cronex, DuPont Co., Wilmington, DE) at −70°C.

### RNase Protection Studies

RNase protection studies were carried out following the protocol of Melton et al. (17). 30 μg total RNA were used for each hybridization. 1 × $10^6$ cpm/ml of probe A or the β/slow MHC probe was ethanol precipitated with the various RNAs, resuspended in 10 μl of 80% formamide hybridization buffer, and placed at 60°C for 3 h. Hybridization was followed by RNase treatment for 30 min at 30°C using 96 μg/ml Ribonuclease A and 19 U/ml Ribonuclease T1. RNase-resistant products were run on 8 M urea, 6% bisacrylamide sequencing gels.

## Results

### Isolation and DNA Sequence Analysis of a Perinatal MHC cDNA Clone

To isolate and characterize cDNAs representing developmentally regulated human MHC genes, we constructed a cDNA library from fetal cardiac tissue. Due to the highly conserved nature of sarcomeric MHCs, it was possible to isolate MHC cDNA clones from this library using a previously isolated sarcomeric MHC cDNA as a probe. The cDNA used here as a probe encodes the light meromyosin region of an adult β cardiac/slow skeletal MHC and cross-hybridizes to multiple MHC genes (22). We screened 4 ×

$10^5$ clones from the fetal cardiac cDNA library, and selected a 3.4-kb cDNA, SMHCP, for further study due to its unique restriction pattern when compared to those of previously isolated MHC cDNAs.

The entire cloned insert of SMHCP was subjected to DNA sequence analysis through the generation of nested deletions. The nucleotide sequence, and its derived amino acid sequence are shown in Fig. 1 *A*. When compared with an entire sarcomeric MHC sequence (27), SMHCP was determined to extend from codon 812 through codon 1,894 (out of 1,894), and to include the complete 3′ UTR of the MHC mRNA and the poly (A) tail. SMHCP therefore encodes a polypeptide which would form the rod portion of the molecule, ending 18 amino acids before the beginning of the globular subfragment one head region. A partial restriction map of SMHCP is shown in Fig. 1 *B*.

To examine the level of homology between SMHCP and other MHCs, its sequence was compared with all mammalian DNA sequences available in published reports. This analysis revealed that SMHCP is most homologous to the cDNA encoding a perinatal isoform of myosin identified in rat hind-leg muscle (20). Table I shows a comparison between the final 257 amino acids (the length of the previously described rat perinatal cDNA) encoded by SMHCP, and the equivalent region of five rat and three human striated muscle MHC cDNAs. It is apparent from this comparison that the amino acids encoded by SMHCP are most homologous to the rat perinatal MHC isoform with differences in only 12 of the 257 compared amino acids (4.7%), eight of which are conservative. Comparison of the amino acids encoded by SMHCP with the human adult fast MHC isoform, shows 23 differences out of 258 amino acids (8.9%); almost twice as many changes as when the comparison is made to the rat perinatal MHC sequence. The perinatal MHC isoform appears to be least similar to embryonic MHC isoforms; more than 18% of the amino acids are different in both the rat and human isoforms.

The identification of SMHCP as encoding a human perinatal MHC isoform is even more apparent when 3′ UTR sequences are compared. When the 3′ UTRs of MHC cDNAs from human and rat are compared, a high degree of homology is maintained only between sequences of analogous isoforms (23). Table I shows that, in this region, the nucleotide differences between the human clone, SMHCP, and the rat perinatal isoform cDNA are the lowest (38.5%). The other comparisons show that the 3′ UTR of SMHCP is also similar to the rat fast skeletal IIb isoform cDNA, varying at 44.1% of the nucleotides. Comparisons of SMHCP to slow MHC isoform cDNAs show differences as high as 74.2%.

An interesting feature of MHC isoform sequence is seen when the carboxyl-terminal 20 amino acids encoded by four human MHC cDNAs and their rat equivalents are examined (Table II). Using the final 20 amino acids of SMHCP as the standard sequence, differences in this region among the other isoforms are indicated. It is evident that the amino acids encoded by SMHCP are most similar to the rat perinatal isoform, showing 100% homology in this region. The other human and rat sequences show similar conservation when homologous isoforms are examined. It appears that this region of the myosin molecule is isoform specific, and is maintained across species, suggesting that it may have a functional role.

The gene-specificity of this region of SMHCP is demonstrated when a radiolabeled probe, probe A (shown in Fig. 1 *B*), generated from the sequences encoding the final 18 amino acids of SMHCP, as well as its entire untranslated region (140 bp), is hybridized to a blot of total human genomic DNA (Fig. 2 *A*). This probe binds to a single band in Eco RI, Hind III, and Pst I digests of human genomic DNA. A 500-bp radiolabeled fragment, probe B (Fig. 1 *B*), from the

*Table I. Comparison of the Human Perinatal MHC cDNA and Amino Acid Sequences to Human and Rat Skeletal and Cardiac MHCs*

| MHC cDNA clone | Difference in coding region nucleotides | Difference in coding region amino acids | Difference in 3′ UTR nucleotides |
|---|---|---|---|
| | | | % |
| Rat | | | |
| Perinatal | 101/771 (13.1%) | 12/257 (4.7%) | 38.5 |
| Embryonic | 165/786 (21.0%) | 51/262 (19.5%) | 56.8 |
| Fast skeletal IIa | | | 61.7 |
| Fast skeletal IIb | | | 44.1 |
| β/slow skeletal | 203/774 (26.2%) | 50/258 (19.3%) | 74.2 |
| Human | | | |
| Embryonic | 191/786 (24.3%) | 49/262 (18.7%) | 58.1 |
| Fast skeletal | 107/774 (13.8%) | 23/258 (8.9%) | 58.2 |
| β/slow skeletal | 197/774 (25.4%) | 46/258 (17.8%) | 73.4 |

The nucleotide and derived amino acid sequences encoded by the final 771 bp of SMHCP and its 3′ UTR, were compared with analogous regions from seven other MHC cDNAs. These include the rat: perinatal (pFOD5) (20), embryonic (pMHC25) (27), adult fast oxidative IIa (pMHC40), and fast glycolytic IIb (pMHC62) (3′ UTR only) (19), and β/slow skeletal (pCMHC5) (15) isoforms; and the human: embryonic (pSMHCE) (Karsch-Mizrachi, I., M. Travis, H. Blau, and L. Leinwand, manuscript submitted for publication), adult fast skeletal (pSMHCA) (23), and β/slow skeletal (pSMHCZ) (23) isoforms.

*Table II. Comparison between the Final 20 Amino Acids of SMHCP and Homologous Regions from Five Rat MHC Sequences and Their Human Equivalents*

| Isoform | Sequence |
|---|---|
| Human perinatal | SQVNKLRVKSREVHTKISAE· |
| Rat perinatal | ********************· |
| Human embryonic | *******A*T*DFTSSRMVHESEE· |
| Rat embryonic | *******A*T*DFTSSRMVHESEE· |
| Human fast skeletal | ******************ISEE· |
| Rat fast skeletal (IIa/IIb) | ******************VISEE· |
| Human slow skeletal | *******A***DIG**GLNEE· |
| Rat slow skeletal | *******A***DIGA*GLNEE· |

The derived amino acid sequence of the human perinatal cDNA, SMHCP, encoding the final 20 carboxy terminal residues, was compared to an equivalent region from seven other MHC cDNA clones. The amino acid sequence of SMHCP is printed in full across the first line of sequence. Differences between SMHCP and the other sequences are highlighted at the appropriate residues, identical residues are denoted by *. There are no differences in this region between SMHCP and sequences from the rat perinatal MHC isoform pFOD5 (20). Other sequences include rat: embryonic (pMHC25) (27), adult fast IIa and IIb skeletal (pMHC40/pMHC62) (19), β/slow skeletal (pCMHC5) (15); and human: embryonic (pSMHCE) (Karsch-Mizrachi, I., M. Travis, H. Blau, and L. Leinwand, manuscript submitted for publication), adult fast skeletal (pSMHCA) (23), and β/slow skeletal (pSMHCZ) (23) cDNAs. The termination codon is shown as ·.

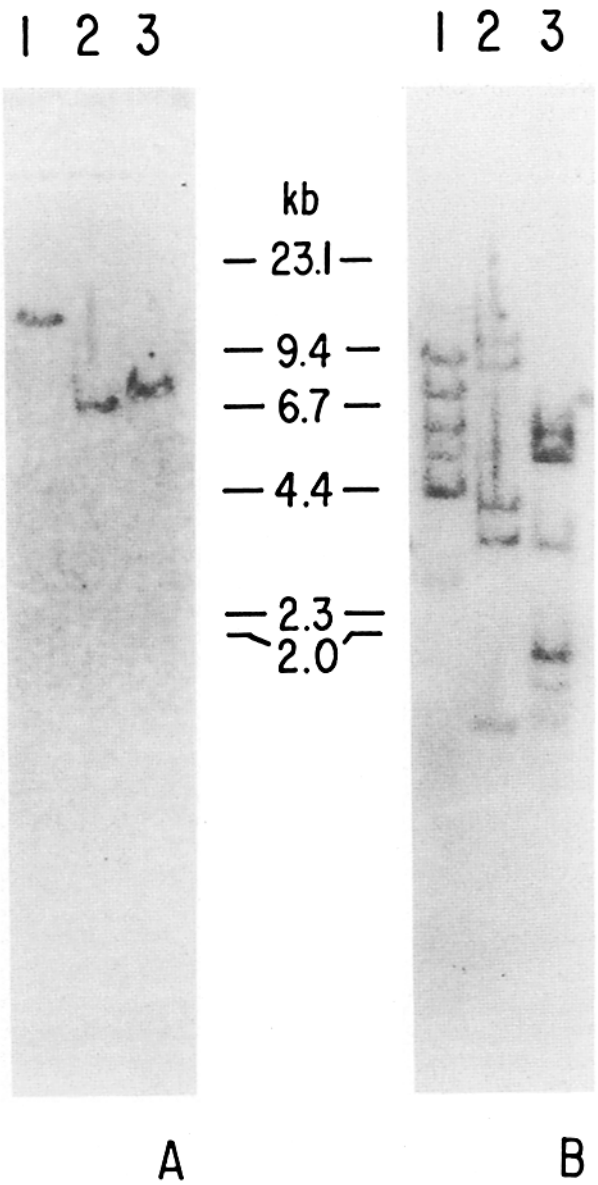


*Figure 2.* Hybridization analysis of human genomic DNA with two probes derived from pSMHCP. Total genomic DNA was digested with endonucleases Eco RI (lane *1*), Hind III (lane *2*), or Pst I (lane *3*), electrophoresed, and transferred to GeneScreen. In *A*, the filter was hybridized to radiolabeled probe A, indicated in Fig. 1 *B*. In *B*, the filter was hybridized to radiolabeled probe B, indicated in Fig. 1 *B*. The size markers (in kilobases) are from λ phage DNA digested with Hind III.

5′ end of SMHCP, encoding sequences from the subfragment 2 region, hybridizes to numerous bands (Fig. 2 *B*).

### *Analysis of Expression of the Human Perinatal MHC Gene*

The pattern of expression of the gene encoding SMHCP was determined through hybridization of the isoform-specific radiolabeled RNA probe, probe A, to RNA samples isolated from adult and fetal skeletal and cardiac tissue. RNA hybridizations, under conditions where 100 pg of homologous RNA can be detected, show that the probe hybridizes strongly to RNA in fetal skeletal tissue, weakly to adult skeletal RNA, and does not react with RNA isolated from fetal or adult cardiac tissue, or from rat liver RNA (data not shown).

To more precisely define and quantitate the expression of this gene in fetal and adult skeletal tissue, RNA was examined using more sensitive ribonuclease protection assays. Anti-sense radiolabeled RNA transcribed from the gene-specific subclone of pSMHCP, probe A, was hybridized to fetal and adult skeletal, fetal cardiac, and adult liver RNAs. Fig. 3 shows protection of a full length 194-bp band in 21-wk fetal skeletal tissue, but not in cardiac tissue from the same period. Full-length protection in two adult skeletal RNA samples was also seen, but at significantly lower levels than in fetal skeletal tissue. For example, lane *2* is a 1-h exposure of the protected product from fetal skeletal tissue. However, a 63-h exposure (lane *4*) for the first adult sample, and a 336-h exposure (lane *5*) for the second are required to obtain signals of similar intensity. The adult skeletal muscle was obtained from two individuals, ages 78 and 57, respectively. The additional lower molecular weight bands seen in the skeletal lanes are most likely the result of the somewhat degraded nature of the RNA used in these studies. The bands which are the result of probe background can be seen in lanes *7* and *8* of Fig. 3, which show the result of RNase incubation with the radiolabeled probe hybridized in the presence of liver and yeast tRNA alone. Despite the isolation of SMHCP from a fetal cardiac library, we were unable to detect the presence of the perinatal myosin isoform in fetal cardiac tissue by RNase protection (lane *6*, Fig. 3) or Northern blot analysis (data not shown). The presence of MHC sequences in this RNA sample is confirmed by the protection of a full-length β cardiac/slow skeletal MHC probe shown in lane *11*. These results lead us to conclude that the fetal cardiac tissue used to make the original cDNA library was expressing a myosin isoform not usually found in this tissue. In summary, the human perinatal MHC isoform is expressed in 21-wk fetal skeletal tissue at very high levels, and its expression is down-regulated in adult skeletal tissue.

## *Discussion*

The results shown here demonstrate that SMHCP encodes a developmentally regulated human MHC isoform, the primary site of expression of which is mid-gestation fetal skeletal muscle. DNA sequence analysis indicates that SMHCP is the human equivalent of a previously described rat perinatal MHC isoform which is present in rat skeletal tissue at 3 wk of gestation, reaching maximum levels of expression at 7 d postbirth, and which is not expressed in the hind-leg tissue of the adult rat (20). By RNA analysis, we have found that SMHCP is not expressed in fetal tissue before 13 wk of gestation (data not shown). RNase protection studies demonstrated that SMHCP is expressed in 21-wk human fetal skeletal and human adult skeletal tissue, but is not expressed in 21-wk fetal cardiac tissue. In contrast to our results, Periasamy et al. (20) did not see expression of the rat perinatal isoform in adult skeletal tissue, although the perinatal isoform mRNA has been seen in adult masseter muscle at low levels using S1 nuclease analysis (7). The difference in the pattern of expression between the rat and human isoforms in adult skeletal tissue may be due to species differences in the

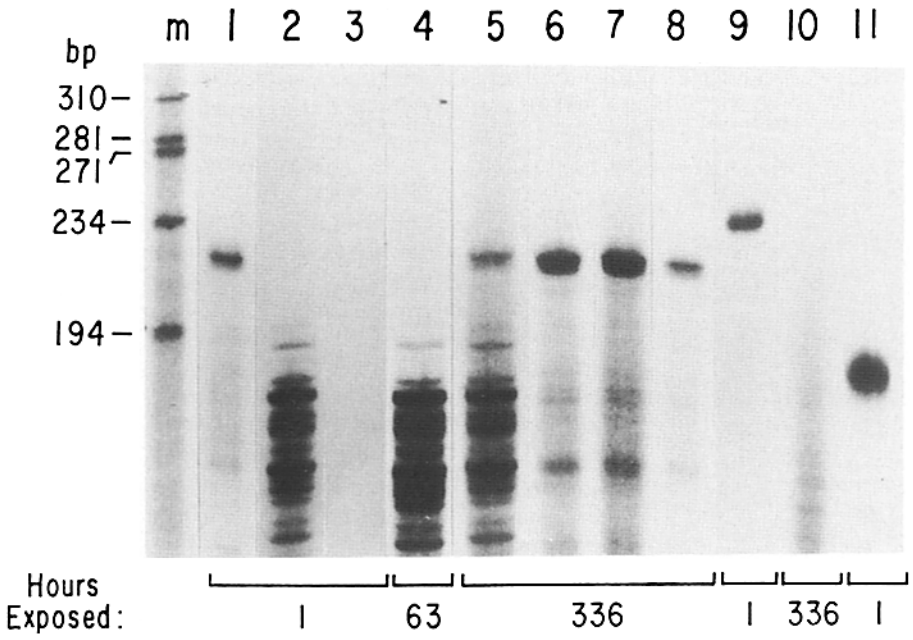


*Figure 3.* Expression of SMHCP mRNAs in fetal and adult skeletal tissue by RNase protection studies. Adult skeletal and fetal cardiac and skeletal RNA samples as well as controls were hybridized with radiolabeled anti-sense cRNA generated from the isoform-specific subclone of pSMHCP, probe A, and an isoform-specific subclone of pSMHCZ, the β/slow MHC cDNA (23). After hybridization, the samples were digested with ribonucleases A and T1, then run on a denaturing gel to identify all RNase-resistant fragments. Lane *m* shows a Hae III digest of ΦX174 DNA radiolabeled with [$\gamma$-$^{32}$P] ATP. The sizes of the bands generated, in base pairs, are listed on the left side of the figure. Lane *1* is the radiolabeled cRNA generated from probe A, which was not digested with RNases. Lanes *2* and *3* are 1-h exposures of the bands generated using total RNA from a 21-wk fetal skeletal sample and from a 57-yr-old adult skeletal sample, respectively. Lane *4* is a 63-h exposure of a 78-yr-old adult skeletal sample hybridized to probe A. Lane *5* is a 336-h exposure of the same skeletal sample seen in lane *3* as a 1-h exposure. Lane *6* was generated using a 21-wk fetal cardiac RNA sample. Lane *7* was generated using adult rat liver total RNA. The RNase-resistant bands resulting from the RNase treatment of the probe alone are seen in lane *8*. Lane *9* is the radiolabeled cRNA probe from the β/slow sequence (pSMHCZ) which was not digested with RNase. Lane *10* is the β/slow probe incubated with liver RNA. Lane *11* is the expected full-length protected fragment obtained when the β/slow probe is hybridized with the 21-wk fetal cardiac RNA sample. The cRNA generated from probe A is 208 bp long (194 bp from pSMHCP and 14 bp from the pTZ19R vector). Full-length protection of a 194-bp fragment is seen only in lanes *2*, *4*, and *5*.

mechanisms which control the developmental regulation of the expression of the isoform.

The presence of SMHCP in the fetal cardiac cDNA library from which it was isolated appears to have been due to an atypical expression of this isoform in the cardiac tissue used to make the library. The fetal cardiac RNA used to make the cDNA library was not available for further analysis. Therefore, to examine further the MHC composition of this library, we screened the fetal cardiac library with other MHC-specific probes. From this analysis, only two types of MHC were found in the library: cDNAs corresponding to the isoform encoded by SMHCP, and cDNAs encoding the α cardiac MHC isoform. α MHC gene expression is cardiac-specific (13). Therefore, the contents of the cDNA library represent cardiac tissue expression. In addition, no fetal skeletal MHC clones were detected. We were unable, however, to detect β cardiac MHC cDNAs, which are normally present at this stage of cardiac development. This suggests that the distribution of MHC cDNA clones in this library is not representative of MHC mRNA distribution in normal, healthy fetal cardiac muscle.

Tsuchimochi et al. (28) have shown, through antibody reactivity and peptide mapping, that an MHC isoform distinct from the previously identified cardiac α and β isoforms is abundantly expressed in fetal ventricular myofibers. The percentage of fibers expressing this isoform decreases significantly after birth, and is up-regulated in patients with dilated cardiomyopathies. Expression of this isoform in fetal or adult skeletal tissue was not addressed. It seems unlikely that we have isolated the cDNA corresponding to the isoform they have identified since we have not been able to reproducibly detect expression of this gene in fetal cardiac muscle, but it does suggest the presence of more than two MHC isoforms in cardiac tissue, and the ability of an isoform to be abnormally expressed in a tissue under pathological conditions.

The maintenance, through evolution, of an MHC gene family capable of undergoing tissue-specific and developmentally regulated expression, suggests a physiologic function for the many isoforms. Regions of amino acid identity in the myosin molecule, maintained among the isoforms and across species most likely encode properties intrinsic to the myosin molecule such as heavy chain association, ATP binding, and filament formation (see reference 29). In contrast, isoform-specific sequences may encode regions which confer functional diversity (2). For example, the isoform-specific sequences found at the extreme 3′ ends of the myosin rods may play a role in regulating the association of homologous heavy chains. A complete understanding of the physiological role of the various MHC isoforms will most likely come only after a detailed evaluation of MHC sequences and their relation to the functional properties of the molecules.

This research was supported by National Institutes of Health grant GM29090 to L. A. Leinwand. R. Feghali is a trainee supported by National Institutes of Health T32GM07128. L. A. Leinwand is an American Heart Association Established Investigator.

Received for publication 16 August 1988 and in revised form 2 December 1988.

## References


1. Bandman, E., R. Matsuda, and R. Strohman. 1986. Developmental appearance of myosin heavy chain and light chain isoforms *in vivo* and *in vitro* in chicken skeletal muscles. *Dev. Biol.* 93:508-518.
2. Emerson, C. P., and S. I. Bernstein. 1987. Molecular genetics of myosin. *Annu. Rev. Biochem.* 56:695-726.
3. Fitzsimons, R. B., and J. F. Hoh. 1981. Embryonic and foetal myosins in human skeletal muscle: the presence of foetal myosins in Duchenne muscular dystrophy and infantile spinal muscular atrophy. *J. Neur. Sci.* 52: 367-384.
4. Freidman, D. J., P. K. Umeda, A. M. Sinha, H. J. Hso, S. Jakovic, and M. Rabinowitz. 1984. Characterization of genomic clones specifying rabbit $\alpha$ and $\beta$ ventricular myosin heavy chains. *Proc. Natl. Acad. Sci. USA.* 81:3044-3048.
5. Grunstein, M., and D. Hogness. 1975. Colony hybridization: a method for the isolation of cloned DNAs that contain a specific gene. *Proc. Natl. Acad. Sci. USA.* 72:3961-3965.
6. Henikoff, S. 1984. Unidirectional digestion with exonuclease III creates targeted breakpoints for DNA sequencing. *Gene.* 28:351-359.
7. Izumo, S., B. Nadal-Ginard, and V. Mahdavi. 1986. All members of the MHC multigene family respond to thyroid hormone in a highly tissue-specific manner. *Science (Wash. DC).* 231:597-600.
8. Kavinsky, C. J., P. K. Umeda. J. E. Levin, A. M. Sinha, J. M. Nigro, J. Smilja, and M. Rabinowitz. 1984. Analysis of cloned mRNA sequences encoding subfragment 2 and part of subfragment 1 of $\alpha$- and $\beta$-myosin heavy chains of rabbit heart. *J. Biol. Chem.* 259:2775-2781.
9. Kraft, R., J. Tardiff, K. S. Krauter, and L. A. Leinwand. 1987. Using miniprep plasmid DNA for sequencing double stranded templates with Sequenase. *Biotechniques.* 6:544-547.
10. Kropp, K. E., J. Gulick, and J. Robbins. 1987. Structural and transcriptional analysis of a chicken myosin heavy chain gene subset. *J. Biol. Chem.* 262:16536-16545.
11. Leinwand, L., L. Saez, E. McNally, and B. Nadal-Ginard. 1983. Isolation and characterization of human myosin heavy chain genes. *Proc. Natl. Acad. Sci. USA.* 80:3716-3720.
12. Leonard, D. G., E. B. Ziff, and L. A. Greene. 1987. Identification and characterization of mRNAs regulated by nerve growth factor in PC12 cells. *Mol. Cell. Biol.* 7:3156-3167.
13. Lompre, A.-M., B. Nadal-Ginard, and V. Mahdavi. 1984. Expression of the cardiac ventricular $\alpha$- and $\beta$-myosin heavy chain genes is developmentally and hormonally regulated. *J. Biol. Chem.* 259:5437-5446.
14. Lowey, S., P. A. Benfield, D. D. Leblanc, and G. S. Waller. 1983. Myosin isozyme in avian skeletal muscles 1: sequential expression of myosin isozymes in developing chicken pectoralis muscles. *J. Musc. Res. Cell Motil.* 4:695-716.
15. Mahdavi, V., M. Periasamy, and B. Nadal-Ginard. 1982. Molecular characterization of two myosin heavy chain genes expressed in the adult heart. *Nature (Lond.).* 297:659-665.
16. Maniatis, T., E. F. Fritsch, and J. Sambrook. 1982. Molecular Cloning. Cold Spring Harbor Laboratory, Cold Spring Harbor, NY. 545 pp.
17. Melton, D. A., P. A. Krieg, R. E. Rebagliate, T. Maniatis, K. Zinn, and M. R. Greene. 1984. Efficient *in vitro* synthesis of biologically active RNA and RNA hybridization probes from plasmids containing bacteriophage SP6 promoter. *Nucleic Acids Res.* 12:7035-7056.
18. Molina, M. I., K. E. Kropp, J. Gulick, and J. Robbins. 1987. The sequence of an embryonic myosin heavy chain gene and isolation of its corresponding cDNA. *J. Biol. Chem.* 262:6478-6488.
19. Nadal-Ginard, B., R. M. Medford, H. T. Nguyen, M. Periasamy, R. M. Wydro, D. Hornig, R. Gubits, L. I. Garfinkel, D. Weiczorek, E. Bekesi, and V. Mahdavi. 1982. Structure and regulation of a mammalian sarcomeric myosin heavy-chain gene. *In* Muscle Development: Molecular and Cellular Control. M. L. Pearson and H. F. Epstein, editors. Cold Spring Harbor Laboratory, Cold Spring Harbor, NY. 143-168.
20. Periasamy, M., D. F. Wieczorek, and B. Nadal-Ginard. 1984. Characterization of a developmentally regulated perinatal myosin heavy-chain gene expressed in skeletal muscle. *J. Biol. Chem.* 259:13573-13578.
21. Rigby, P. W., M. Dieckmann, C. Rhodes, and P. Berg. 1977. Labeling deoxyribonucleic acid to high specific activity *in vitro* by nick translation with DNA polymerase I. *J. Mol. Biol.* 113:237-251.
22. Saez, L. J., and L. A. Leinwand. 1986. Cloning and characterization of myosin cDNAs in adult human skeletal muscle. *In* Molecular Biology of Muscle Development. C. Emerson, D. Fischman, B. Nadal-Ginard, and M. A. Q. Siddiqui, editors. Alan R. Liss, Inc., NY. 263-272.
23. Saez, L., and L. A. Leinwand. 1986. Characterization of diverse forms of myosin heavy chain expressed in adult human skeletal muscle. *Nucleic Acids Res.* 14:2951-2969.
24. Saez, L. J., K. M. Gianola, E. M. McNally, R. Feghali, R. Eddy, T. Shows, and L. A. Leinwand. 1987. Human cardiac myosin heavy chain genes and their linkage in the genome. *Nucleic Acids Res.* 15:5443-5459.
25. Silberstein, L., and H. M. Blau. 1986. Two fetal-specific fast myosin isozymes in human muscle. UCLA Symp. *Mol. Cell. Biol.* 29:253-262.
26. Southern, E. M. 1975. Detection of specific sequences among DNA fragments separated by gel electrophoresis. *J. Mol. Biol.* 98:503-517.
27. Strehler, E. E., M.-A. Strehler-Page, J.-C. Parriard, M. Periasamy, and B. Nadal-Ginard. 1986. Complete nucleotide and encoded amino acid sequence of a mammalian myosin heavy chain gene. Evidence against intron dependent evolution of the rod. *J. Mol. Biol.* 190:291-317.
28. Tsuchimochi, H., Y. Yazaki, M. Kawana, S. Kinata, and F. Takaka. 1987. The existence of a fetal type myosin heavy chain (MHC) in the human heart and its abnormal expression in the ventricles of dialated cardiomyopathy. *Circulation.* 76:1045. (Abstr.)
29. Warrick, H. M., and J. A. Spudich. 1987. Myosin structure and function in cell motility. *Annu. Rev. Cell Biol.* 3:379-421.
30. Webster, C., L. Silberstein, A. P. Hays, and H. M. Blau. 1988. Fast muscle fibers are preferentially affected in Duchenne muscular dystrophy. *Cell.* 52:503-513.
31. Weydert, A., P. Daubas, I. Lazaridis, P. Barton, I. Garner, D. P. Leader, F. Bonhomme, J. Catalan, D. Simon, J. L. Guenet, F. Gros, and M. E. Buckingham. 1985. Genes for skeletal muscle myosin heavy chains are clustered and are not located on the same mouse chromosome as a cardiac myosin heavy chain. *Proc. Natl. Acad. Sci. USA.* 82:7183-7187.
32. Weydert, A., P. Barton, J. Harris, C. Pinset, and M. Buckingham. 1987. Developmental pattern of mouse skeletal myosin heavy chain gene transcripts *in vivo* and *in vitro*. *Cell.* 49:121-129.
33. Whalen, R. G., S. M. Sell, G. S. Butler-Browne, K. Schwartz, P. Bouveret, and I. Pinset-Harstrom. 1981. Three myosin heavy-chain isozymes appear sequentially in rat muscle development. *Nature (Lond.).* 292: 805-809.
34. Whalen, R. G. 1985. Myosin isoenzymes as molecular markers for muscle physiology. *J. Exp. Biol.* 115:43-53.
35. Wydro, R., H. T. Nguyen, R. Gubits, and B. Nadal-Ginard. 1983. Characterization of sarcomeric myosin heavy chain genes. *J. Biol. Chem.* 258: 670-678.