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 isofrom-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) 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 iso-zyame. Members of this gene family show both tissue-specific and developmentally regulated expression (10, 19, 34). During development, fetal, neonatal, and adult MHC isozyymes are expressed in a sequential program (1, 3, 14, 32, 33). However, the functional significance of each of these isozyymes, 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 isozyyme expression to accommodate new demands.

cDNA and genomic clones corresponding to several sarcomeric MHC isozyymes 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 isozyymes from different species are compared, they are even more homologous than intraspecies comparisons, suggesting functional constraints on the divergence of isozyyme sequences. The 3' untranslated region (UTR) sequences are also maintained across species when analogous isozyymes are compared, and appear to be quite isozyyme specific (19, 20).

The patterns of MHC isozyyme 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 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 isozyyme. 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-

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

This sequence has been registered with the EMBL/GenBank, accession number Y00821.
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 sequence. The solid line corresponds to the 3' UTR. Wavy lines correspond to vector sequence. Regions used to generate probes A and B are indicated.

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 cardiac MHC 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 Okazaya-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.

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 NaH₂PO₄-H₂O 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 × 10⁶ 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 NaH₂PO₄-H₂O, 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 × 10⁶ cpm/ml of 3²P-labeled 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 × 10⁶ 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% bis-acrylamide 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 ×
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. 1A. 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. 1B.

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. 1B), generated from the sequences encoding the final 20 amino acids of SMHCP, as well as its entire untranslated region (140 bp), is hybridized to a blot of total human genomic DNA (Fig. 2A). 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. 1B), from the

---

Table 1. Comparison of the Human Perinatal MHC cDNA and Amino Acid Sequences to Human and Rat Skeletal and Cardiac MHCs

<table>
<thead>
<tr>
<th>MHC cDNA clone</th>
<th>Difference in coding region nucleotides</th>
<th>Difference in coding region amino acids</th>
<th>Difference in 3' UTR nucleotides</th>
</tr>
</thead>
<tbody>
<tr>
<td>Rat</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Perinatal</td>
<td>101/771 (13.1%)</td>
<td>12/257 (4.7%)</td>
<td>38.5</td>
</tr>
<tr>
<td>Embryonic</td>
<td>165/786 (21.0%)</td>
<td>51/262 (19.5%)</td>
<td>56.8</td>
</tr>
<tr>
<td>Fast skeletal IIa</td>
<td>167/1074 (25.4%)</td>
<td>50/258 (19.3%)</td>
<td>74.2</td>
</tr>
<tr>
<td>Fast skeletal IIb</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>β/slow skeletal</td>
<td>203/774 (26.2%)</td>
<td>50/258 (19.3%)</td>
<td>74.2</td>
</tr>
<tr>
<td>Human</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Embryonic</td>
<td>191/786 (24.3%)</td>
<td>49/262 (18.7%)</td>
<td>58.1</td>
</tr>
<tr>
<td>Fast skeletal</td>
<td>107/174 (13.8%)</td>
<td>23/258 (8.9%)</td>
<td>58.2</td>
</tr>
<tr>
<td>β/slow skeletal</td>
<td>197/774 (25.4%)</td>
<td>46/258 (17.8%)</td>
<td>73.4</td>
</tr>
</tbody>
</table>

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 (pSMHC5) (Karsch-Mizrachi, I., M. Travis, H. Blau, and L. Leinwand, manuscript submitted for publication), adult fast skeletal (pSMHC23), 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

<table>
<thead>
<tr>
<th>Isoform</th>
<th>Sequence</th>
</tr>
</thead>
<tbody>
<tr>
<td>Human perinatal</td>
<td>SQVNLVRK SREVTAE</td>
</tr>
<tr>
<td>Rat perinatal</td>
<td></td>
</tr>
<tr>
<td>Human embryonic</td>
<td></td>
</tr>
<tr>
<td>Rat embryonic</td>
<td></td>
</tr>
<tr>
<td>Human fast skeletal</td>
<td></td>
</tr>
<tr>
<td>Rat fast skeletal (IIa/IIb)</td>
<td></td>
</tr>
<tr>
<td>Human slow skeletal</td>
<td></td>
</tr>
<tr>
<td>Rat slow skeletal</td>
<td></td>
</tr>
</tbody>
</table>

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 (pSMHC5) (Karsch-Mizrachi, I., M. Travis, H. Blau, and L. Leinwand, manuscript submitted for publication), adult fast skeletal (pSMHC23), and β/slow skeletal (pSMHCZ) (23) cDNAs. The termination codon is shown as ..
ly 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 II. 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
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 physiologic 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.

The Journal of Cell Biology, Volume 108, 1989 1796
References


