Myomesin is a high molecular weight protein that is present in the M-band of all fiber types of cross-striated skeletal muscle and heart. We have isolated two cDNAs encoding tissue-specific isoforms of chicken myomesin with calculated molecular masses of 174 kDa in skeletal muscle and 182 kDa in heart. Distinct sequences are found at the 3'-end of the two cDNAs, giving rise to different C-terminal domains. Partial analysis of the gene structure has shown that in chicken, both isoforms are generated by alternative splicing of a composite exon. Amino acid sequences show that the main body of myomesin consists of five fibronectin type III (class I motifs) and seven immunoglobulin-like domains (class II motifs). An identical structure was found in M-protein and human 190K protein (the human counterpart of chicken myomesin), and a comparable domain arrangement occurs in the M-band-associated protein skelemin. We postulate that myomesin, M-protein, and skelemin belong to the same subfamily of high molecular weight M-band-associated proteins of the immunoglobulin superfamily and that they probably have the same ancestor in evolution.

The M-band is a striking structure in the middle of the sarcomere of cross-striated skeletal and heart muscle, where no myosin cross-bridges are present. Detailed electron microscopic investigations of the M-band region have shown (Knappes and Carlse, 1968; Luther and Squire, 1978, 1980; Luther et al., 1981; Sjöström and Squire, 1977) that up to 17 sublines can be detected. The five most prominent lines, M1, M4/M4, M6/M6 according to the nomenclature of Sjöström and Squire (1977), were assigned to the so-called M-bridges, which are supposed to link neighboring myosin filaments to each other, as seen in corresponding cross-sections of myofibrils (Luther et al., 1981; Luther and Squire, 1980). Interestingly, the M-band pattern depends on the muscle fiber type (Pask et al., 1994). Fiber type-specific patterns due to the arrangements of certain proteins have been observed: a five-line pattern occurs in intermediate fibers (type IIA) of chicken skeletal muscle, whereas slow fibers (type I) have a four-line pattern and fast fibers (type IIB) have only three prominent lines (Edman et al., 1988).

The M-band appears to be a stable part of the A-segment, because in this region M-bridges connect thick filaments directly to each other, providing mechanical stability to the thick filament lattice. Therefore, the M-band seems to be an ideal junction for interaction of the cytoskeleton with the center of the sarcomere, similar to the Z-disk interaction at the distal borders of sarcomeres. So-called M-cables attached to the M-band on one end and to another M-band or the sarcolemma on the other end have been visualized by electron microscopy (Pierobon-Bormioli, 1981; Street, 1983; Wang and Ramirez-Mitchell, 1983). More recently an M-band associated protein, skelemin, has been described, possibly anchoring the intermediate filament lattice to the M-band (Price, 1987).

So far, three integral components of the M-band have been identified: the homodimeric muscle isoform of creatine kinase (MM-CK) (Turner et al., 1973; Wallimann et al., 1983a, 1983b; Schäfer and Perriard, 1988), M-protein (Masaki and Takaiti, 1972, 1974; Noguchi et al., 1992; Streher et al., 1979), and myomesin (Eppenberger et al., 1981; Grove et al., 1984). With the exception of MM-CK, localized in the M4/M4' sublines (Wallimann et al., 1983a, 1983b; Streher et al., 1983; Doetschmann and Eppenberger, 1984), it is not clear how the other M-band proteins are participating in the M-band structure. Morphological studies indicate that M-protein and myomesin are either confined to the substructures M6/M6' of the M-band or may occur within the whole region between these substructures, ensheathing the thick filaments in this region (Strehler et al., 1983). Correlation between the absence of M-protein from slow muscle fibers (type I) of rat skeletal muscle (musculus tibialis anterior and musculus soleus) and the absence of subline M1 from M-bands of the same fibers indicate that M-protein might be the main component of the M-bridges in M1 (Carlsson and Thornell, 1987). This idea is supported by the observation that M-bridges in M4/M4' rather than in M1 seem to be essential for the maintenance of the thick filament lattice in both skeletal and heart muscle (Luther et al., 1981; Luther and Squire, 1980; Pask et al., 1994). If M-protein is indeed the main component of the M-bridge in M1, then myomesin could contribute to the stabilization of the distal region of the M-band where the M6/M6' sublines occur. The dimension of both high molecular M-band proteins, M-protein and myomesin, is also large enough to simultaneously contribute to M-bridges and M-filaments together with other candidates like titin, which has been shown to extend into the center of the M-band (Fürst et al., 1989; Vinkemeier et al., 1993) or some as yet unknown proteins.

Whatever the exact arrangement of myomesin and M-protein is, one of their functional roles could be the anchoring of titin to the thick filaments in the M-band. Both proteins have a strong affinity for titin (Nave et al., 1989) as well as for myosin (Mani and Kay, 1978), indicating that these proteins could make a connection between myosin and titin. Since myomesin is present in all fiber types of adult cross-striated muscle...
tissue, it seems to be better suited for such a fundamental function than M-protein, which is only present in fast fibers (type II) and heart (Grove, 1989; Grove et al., 1987, 1989). For an elucidation of the M-band structure in detail, it will be indispensable to study the molecular parameters of the M-band associated proteins. In order to investigate structural and functional aspects of myomesin, cloning of the chicken myomesin cDNAs was carried out. Here, the complete coding region of a myomesin cDNA derived from chicken heart and cDNA fragments derived from chicken skeletal muscle and rat muscle is reported. The comparison of the transcripts of chicken heart and skeletal muscle shows that there is a tissue-specific isoform present in chicken heart, which very likely arises by alternative splicing from a myomesin gene encoding both isoforms. Sequence data, results obtained from Northern blot analysis, and polyme- rase chain reaction (PCR) on chicken genomic DNA confirm this hypothesis.

EXPERIMENTAL PROCEDURES

Isolation of cDNAs

λ-Clones containing M-band protein cDNAs were identified by conventional immunoscreening of a λgt 11 chicken skeletal muscle cDNA library (Hassler et al., 1986) with a polyclonal antibody against chicken M-protein and myomesin. The inserts of five λ-cDNAs were subcloned into the vector pBluescript II KS(+)(Clontech, Palo Alto, CA) and were further analyzed by sequencing. M-protein cDNA clones were determined and sorted out by comparison of the obtained sequence with the chicken M-protein sequence (Noguchi et al., 1992), using the application “ALIGN” (DNASTAR Inc., Madison, WI). Since the inserts of two clones, SB194 and SB198, contain sequences different from M-protein, these clones were postulated to contain myomesin cDNAs. Therefore, a fragment of SB194 was used as a probe to screen a λgt 11 chicken heart muscle cDNA library (Clontech) by DNA hybridization, and in parallel, immunoscreening with the monoclonal antibody B4 against myomesin (Grove et al., 1984) was carried out according to the manufacturer’s recommendations (Clontech). Probes for nucleic acid hybridization were prepared from the 2.0-kb EcoRI-fragment of clone SB194 (positions corresponding to the chicken cardiac myomesin sequence shown in Fig. 2, 1388–3649), the 740-bp EcoRI-fragment of clone SB275 (positions 960–1700) for detection of the 5'-UTR and the first 1245 bp of the coding region, which is identical in both the skeletal and cardiac cDNA. 2) A heart-specific probe containing 237 bp of the heart-specific coding region (bases 4653–4890) was derived from clone SB289 by PCR with the primers Seq1 (positions 4653–4672) and An3 (positions 4869–4890), yielding subclone SB286 in the vector pDirect™. 3) A 237-bp probe containing the last 61 bp of the skeletal muscle coding region as well as 176 bp of the skeletal muscle-specific 3'-UTR was amplified from SB198 using the primers Seq3 and An1 (see Fig. 6A) and subcloned into pDirect™, resulting in clone SB294. For detection of rat myomesin transcripts, a 228-bp probe was amplified from clone SB324 containing the region of the rat cDNA that corresponds to the chicken skeletal muscle probe. The hybridization pattern of this probe was identical to the one obtained from hybridization with a 500-bp-long probe derived from the 5'-end of clone DA10.2 (Fig. 1) and corresponding to common coding sequences (see Fig. 4D).

In Situ Hybridization of Myomesin Transcripts in Chicken Embryos

Preparation of Embryos and Sections—Stage 34 embryos were removed from the eggs and treated as described by Sassoon et al. (1988), with one minor alteration. Instead of an incubation in xylene-ethyl methylated paraplast, the xylene-treated embryos were directly transferred into molten paraplast for 3 × 30 min at 60°C.

Preparation of Plasmids and cRNA Probes—Plasmids SB286 and SB294 were linearized by XbaI and SacI and were transcribed with T3 RNA polymerase (Promega), producing antisense probes of 237 nucleotides. For the negative control, a 1500-bp RNA probe derived from a full-length cDNA clone, CMD1, of chicken MyoD1 (Lin et al., 1989) was used. The construct was linearized by KpnI/BamH1-digestion and transcribed by T3 RNA polymerase, yielding a probe of 1500 bp identical to the mRNA. In vitro transcription was performed in 20 µl assays under the following conditions: 1 × transcription buffer (Promega), 10 mM dithiothreitol (Sigma Chemie, Buchs, Switzerland; nuclease-free), 400 ng each, 1 µg of RNA, in (40 Ci) of [35S]rUTP (Amersham International, Buckinghamshire, United Kingdom). The end volume was adjusted to 20 µl by diethylpyrocarbonate (Sigma)-treated distilled H2O and template to a final concentration of 40–50 ng µl−1. RNAsin (40 units) and RNA polymerase were added at the end. Another aliquot of RNA polymerase was added after 1 h of incubation at 37°C, and after another hour of incubation the reaction was stopped by digestion of the template with DNase I. The probe of the negative control was alkali-hydrolyzed to an average length of 150 bp, and all probes were purified by G50-Sephadex columns.

All steps of prehybridization, hybridization, subsequent washing, and exposure were carried out according to published methods (Sassoon et al., 1988) with three modifications. First, for the hybridization, the probes were diluted in hybridization buffer to a concentration of 20,000 cpm/µl. Second, the dithiothreitol concentration in the hybridization buffer was increased to 100 mM (Dagerlind et al., 1993). Third, more stringent washes were carried out at 55°C in 0.5 × SSC, 50% formamide, and 10 mM dithiothreitol for 2 h.

Delfield’s Hematoxylin-Eosin Stain—Staining was done after dehydrating the sections as follows: incubation for 1 h in 1% Delfield’s hematoxylin, washing with water, incubation for 10 min in 0.1% eosin, washing with water and with 80% EtOH briefly, dehydration by EtOH and xylene, and embedding.

Analysis of the 5'-Region of the Skeletal Muscle Transcript by RT-PCR

In order to confirm the presence of identical sequences in the 5'-region of the skeletal muscle message as in the cardiac cDNA, RT-PCR

1 The abbreviations used are: PCR, polymerase chain reaction; kb, kilobase pair(s); bp, base pair(s); UTR, untranslated region; RT, reverse transcription.
was performed on total RNA isolated from leg muscle tissue of 17-day-old chicken embryos, which was also used for Northern blot analysis. About 1 μg of RNA was reverse-transcribed in a 20-μl assay using 400 units of reverse transcriptase of Moloney murine leukemia virus in the presence of 40 units of RNase inhibitor (RNAsin), 500 μM dNTP (dATP, dCTP, dGTP, and dTTP), 1 × RT buffer (Life Technologies, Inc.), and 1 μg of antisense primer. After 1 h of incubation at 37°C, the reaction was stopped by incubation at 65°C for 10 min and the addition of 30 μl of Taq polymerase (Promega) in the presence of 100 ng/μl of the RT assay buffer (Promega), 1 μg of cDNA, 1 × PCR buffer (Promega), 1 μg of antisense primer, and 3 × antisense primer. 40 cycles of amplification were performed under standard conditions: 15 s at 95°C, 30 s at 58°C, and 30 s at 72°C. The length of the PCR products was analyzed on 1% agarose, EtBr gels.

Three sets of primers, containing sequences derived from the chicken cardiac myomesin cDNA, were used for different assays: 1) RT primer, 5′-CAGGACTGTGACAAC-3′ (positions corresponding to the chicken cardiac myomesin sequence shown in Fig. 2, 1222–1238); 5′-PCR primer, 5′-AGTAAGACTGATGACAACC-3′ (positions 3777–3796); 3′-PCR primer, 5′-CCAGCCACTGTACAGTGAA-3′ (positions 3232–3251); 3′-PCR primer (same as RT primer; positions 3777–3796) (size of the product, 564 bp); and 3) RT-primer, 5′-TCGATCAGACCAGTGACAG-3′ (positions 4869–4890); 5′-PCR-primer, 5′-AGTAAGAGCTGCTGGCACT-3′ (positions 4232–4251); 3′-PCR primer, 5′-CAGCCACTGTACAGTGAA-3′ (positions 4825–4843); and the primers An1, An2, An3, and An4 contain a 12-nucleotide-long extension at their 5′-end of the coding region including intron sequence derived from clone SB371; An2, 5′-GGCACGGAGACCAGCGA-3′, 4628–4644; Se5, 5′-ACAACTGCGATTGACAGG-3′, 4356–4374; Se3, 5′-CTCTTTTCTTCCAAAGAGGCA-3′, intron sequence derived from clone SB371; Se4, 5′-GCCACGGAACGACCCGA-3′, 4628–4644; Se5, 5′-AAGTTGGTTCTCATACAGGAGG-3′, skeletal muscle coding region; Se6, 5′-GATCGTGGATTCTGTTGCAC-3′, 4653–4672; Se7, 5′-TGCCACAGAAGAGACACCT-3′, 4825–4843; An1, 5′-AACAGAATGGGGACACAGT-3′, intron sequence derived from clone SB371; An2, 5′-CTCTTTCTGGTGGACA-3′, 4412–4430; An3, 5′-ATATTAGGGAAGGACATTTTTT-3′, skeletal muscle-specific 3′-UTR; An4, 5′-GCTTACATCGTGCTTC-3′, 4672–4689; An5, 5′-GCTTTTTGCTTTTTCTTTT-3′, 4869–4890; An6, 5′-AGACATAACTGCAACACATTT-3′, 5173–5193. The primers Se1, Se3, and Se4 contain a 12-nucleotide-long extension at their 5′-end (5′-CTGTTTCCGCGCCA-3′), and the primers An1, An2, An3, and An4 contain a 12-nucleotide-long extension at their 5′-end (5′-CTGCTTGCCCCA-3′), which does not hybridize to the templates but facilitates subcloning into pDirect™.

**RESULTS**

Isolation of Chicken Myomesin cDNA Clones—Immunoblot analysis of tissue extracts showed a slightly higher band for myomesin in chicken heart (190–195 kDa) than in chicken skeletal muscle (185 kDa); therefore, the existence of different myomesin isoforms in skeletal muscle and heart had been postulated (Grove et al., 1985). In order to investigate this isoform diversity, cDNAs of myomesin were isolated from two different libraries. Initial immunoscreening of a λgt11 chicken skeletal muscle library (Hossle et al., 1986) with a polyclonal
antibody recognizing both myomesin and M-protein (Eppenberger et al., 1981) yielded five positive l-clones, which were partially sequenced. These sequences were compared with the cDNA sequences of chicken M-protein (Noguchi et al., 1992) to sort out the clones containing M-protein sequences. The remaining clones SB194 and SB198 (Fig. 1B) contained sequences different from the M-protein cDNA. Since the polyclonal antibody used for the screening recognizes both M-protein and myomesin, we postulated that the insertsof the non-M-protein clones contain part of the myomesin cDNA.

A probe generated from the putative skeletal muscle cDNA insert of clone SB194 was used to screen the lgt11 chicken

FIG. 2. Nucleotide sequence of the chicken heart myomesin cDNA and the derived full-length amino acid sequence. The complete coding region of the chicken heart myomesin cDNA consists of 4893 nucleotides in one open reading frame. The derived amino acid sequence is shown in single letter code below the nucleotide sequence. An in frame TAA stop codon is present 39 bp upstream from the ATG start codon (position 58), and the real TAA stop codon (position 4948) is followed by alternative stop codons in all possible frames (i.e., TAG at position 4955; TGA at positions 5004 and 5092).

Isolation and Characterization of Myomesin cDNAs

antibody recognizing both myomesin and M-protein (Eppenberger et al., 1981) yielded five positive l-clones, which were partially sequenced. These sequences were compared with the cDNA sequences of chicken M-protein (Noguchi et al., 1992) to sort out the clones containing M-protein sequences. The remaining clones SB194 and SB198 (Fig. 1B) contained sequences different from the M-protein cDNA. Since the polyclonal antibody used for the screening recognizes both M-protein and myomesin, we postulated that the inserts of the non-M-protein clones contain part of the myomesin cDNA. A probe generated from the putative skeletal muscle cDNA insert of clone SB194 was used to screen the lgt11 chicken
Isolation and Characterization of Myomesin cDNAs

heart muscle library to get the complete coding region of the heart myomesin cDNA. From 10^5 screened plaques, 40 positive λ-clones were found. The fusion proteins of five of these clones reacted with the monoclonal antibody B4 against myomesin (Grove, 1984), indicating that the cDNA inserts encode part of myomesin. In Fig. 1A, some of the isolated cDNA-fragments are schematically represented in alignment with the rimer given by the coding region of the chicken heart cDNA (Fig. 1A, SB13, SB146, SB153, SB221, SB240, and SB275). Two of the fragments, SB13 and SB240, originate from λ-clones that produce a fusion protein reactive with the antibody B4. In the screening with the first probe derived from the skeletal myomesin clone SB194, we did not detect either terminal part of the coding region. Therefore, a second screening was carried out with two probes derived from the 5′-clone SB275 and the 3′-clone SB146, respectively. In 5 × 10^5 screened plaques, four positive clones were found with the 5′-probe and 17 clones were found with the 3′-probe. Two clones shown in Fig. 1, SB233 and SB280, represent this round of screening. Altogether, we have isolated 5204 bp of the chicken heart cDNA containing the complete coding region of 4890 bp, 57 bp of the 5′-UTR, and 257 bp of the 3′-UTR (Fig. 2).

The entire coding region of the chicken heart myomesin cDNA has been sequenced at least once on both strands. Some of the sequenced cDNA clones, which cover the whole coding region, are shown in Fig. 1A. The determined open reading frames of these clones have been merged at their overlapping parts, and critical links have been verified and completed by sequencing the corresponding regions of additional clones.

The determined nucleotide sequence as well as the predicted amino acid sequence are shown in Fig. 2, where the ATG start codon of translation is located in position 58. Since it is the first possible initiation site and an in frame stop codon occurs 39 bp upstream from the ATG, we presume that this is the real start codon. The TAA stop codon at position 4948 seems to be the real termination site because it is followed by alternative stop codons in all possible reading frames (i.e. TAG at position 4955 and TGA at positions 5004 and 5092). The resulting open reading frame of 4893 bp encodes a protein of 1630 amino acids, which corresponds to myomesin by several criteria. First, the monoclonal antibody B4 against myomesin (Grove et al., 1984) reacts with the fusion proteins of the two λ-clones, which are represented by SB13 and SB240 in Fig. 1 but does not react with fusion proteins produced by λ-clones containing M-protein cDNAs. Second, the calculated molecular mass of 182 kDa corresponds to the molecular mass determined by SDS-polyacrylamide gel electrophoresis (skeletal muscle myomesin, 184 kDa (Grove et al., 1984); cardiac myomesin, 190–195 kDa (Grove et al., 1985)). Third, sequence data base searches (GenBank™, Swiss-Prot 31, NBRF-PIR™ 44) showed that the sequence is not yet published and that myomesin is related to human 190K protein (Vinkemeier et al., 1993), mouse skellem (Price and Gomer, 1993), and human as well as chicken M-protein (Noguchi et al., 1992; Vinkemeier et al., 1993). M-protein and skellem sequences are distinct from myomesin, although all three share quite high homology. Chicken myomesin and human 190K protein share an overall homology of about 76% sequence identity on amino acid and nucleotide level. Meanwhile, it has been demonstrated that the monoclonal antibody B4 against myomesin (Grove et al., 1984) recognizes the human and bovine 190K proteins on immunoblots (Vinkemeier et al., 1993 (note added in proof)), indicating that the 190K protein is the human counterpart of chicken myomesin. The 190K sequence confirms the correct position of the ATG start codon and the resulting open reading frame in the presented sequence because the predicted amino acid sequences of both myomesin and the human 190K cDNA are highly homologous.

Different methods were used to analyze the putative skeletal muscle-specific cDNA clones in order to find possible isofor-
specific differences (Grove et al., 1985). Sequences derived from the putative skeletal muscle-specific clones SB194 and SB198 were aligned with the heart myomesin cDNA according to the Wilbur-Lipman algorithm. All inspected skeletal muscle sequences are identical to the corresponding region of the heart cDNA, except one piece derived from clone SB198, which contains the 3'-end of the skeletal muscle coding region and part of the 3'-UTR. In Fig. 3A, the 3'-end of the skeletal muscle coding region (line II) including the TGA stop codon, is aligned with the corresponding cardiac sequence (line I). The divergence between the two cDNAs starts after position 4652 (the 4595th base of the coding region) where a putative 5'-splice site (GTAAG) is present in the skeletal muscle sequence. The 3'-end of the coding region downstream from this position is 61 bp in length, whereas the particular cardiac coding sequence has a length of 295 bp. Both nucleotide sequences give rise to different C-terminal amino acid segments of 20 and 98 amino acids, respectively, leading to different calculated molecular masses of 182 kDa for cardiac myomesin and 174 kDa for skeletal muscle myomesin. Comparison of the chicken skeletal muscle-specific C terminus with the corresponding amino acid sequence deduced from the rat myomesin cDNA (Fig. 3B, line I) and the C termini of the related proteins, human 190K protein (Fig. 3B, line IV) and mouse skelem in (Fig. 3, line V), shows high homology between these proteins, confirming that the particular sequence of clone SB198 is part of the chicken skeletal muscle cDNA and not a cloning artifact.

Since the clones SB194 and SB198 do not contain the entire coding region of the skeletal muscle cDNA, RT-PCR was performed on total RNA isolated from chicken leg muscle with different sets of primers derived from the heart sequence, to confirm sequence identity in regions that were not represented in clones isolated from the library (for details see “Experimental Procedures”). The length of all PCR products corresponded exactly to the length that was expected from the heart sequence. Furthermore, one of the PCR fragments was sequenced after subcloning into the vector pDirect™ (SB367 in Fig. 1B), and again no difference between the obtained sequence and the corresponding region of the heart cDNA (bases 232–850) was found. According to these results, the first 4595 bases of the coding region of both chicken myomesin cDNAs are identical.

Isolation of Rat Myomesin cDNAs—In order to investigate whether a similar isoform diversity occurs in mammalian species, a fragment of the chicken cDNA (clone SB275) was used to screen a λgt11 rat heart cDNA library (Zhu et al., 1991). Nine positive clones were found in a first round of screening, represented by DA4.6 and DA10.2 (Fig. 1C). Since the 3'-end of the cDNA could not be found in these clones, a second screening with a probe derived from the 3'-end of clone DA10.2 was carried out. In 5 × 10⁴ plaques, seven positive clones were found, and two representative clones, SB316 and SB324, are shown in Fig. 1C. The clone SB324 was analyzed in more detail and appeared to contain the 3'-end of the coding region and some 3'-untranslated sequence. Surprisingly, comparison of this sequence with the chicken sequences (Fig. 3A) revealed higher homology of the rat 3'-end to the corresponding chicken...
transcripts, indicating the generation of tissue-specific myomesin mRNA.

Tissue-specific Isoforms of Myomesin Are Expressed in Chicken as Transcripts of Different Sizes—The length of the myomesin transcripts in different chicken and rat tissues was determined by Northern blot analysis of total RNA that was isolated from 17-day chicken embryos (prehatching) and adult chicken or rat tissues. In chicken, three different myomesin mRNAs were detected by a probe containing the entire cDNA insert of clone SB194, corresponding to coding sequences that are identical in both isoforms. In skeletal muscle, a single myomesin mRNA of about 5.5 kb occurs that is absent from heart have the same mobility of 5.5 kb in denaturing agarose gels, they could be identical. Alternative splicing, as proposed in chicken, however, cannot be completely excluded in the tested mammalian tissues, since a possible GT splice donor was also found in the rat sequence (Fig. 3A, line IIII), at a position corresponding to the GTAAG splice site found in the chicken skeletal muscle myomesin cDNA (Fig. 3A, line III). In order to determine whether a tissue-specific portion is also present at the 3'-end of the rat cDNA, a probe containing 228 bp downstream from the possible GT splice donor was prepared from the cDNA clone SB324 (Fig. 1C). Containing the last 73 bp of the coding region and 155 bp of the 3'-UTR, this probe shares rather high homology with the 3'-end of chicken skeletal muscle cDNA (Fig. 3A, line II) but differs significantly from the chicken heart-specific 3'-end (Fig. 3A, line I). Like the common 5'-probe, the 3'-probe of rat myomesin hybridized to the 5.5-kb transcript found in skeletal muscle and heart. Unlike the chicken transcripts, the rat myomesin transcripts seem to share identical sequences in skeletal muscle and heart, and therefore only one myomesin isoform may be present in rats.

In order to investigate the expression of the two isoforms restricted in tissue-specific manner, in situ hybridization of sections from paraffin-embedded chicken embryos was performed. RNA probes containing sequences analogous to the isoform-specific probes used for Northern blot analysis (Fig. 4, A and C), were used for in situ localization of myomesin RNAs in 8-day-old chicken embryos (stage HH) as described under “Experimental Procedures.” The heart-specific probe hybridized to the developing heart of an 8-day-old chicken embryo.
exclusive (Fig. 5A). The skeletal muscle probe, however, hybridized preferentially to the muscle anlagen in the wing bud and in the trunk, and faint hybridization was also found in the heart (Fig. 5B). These findings have been confirmed by RT-PCR, which allows us to amplify a skeletal muscle-specific product from total RNA isolated from heart tissue of 17-day chicken embryos, whereas no cardiac product can be amplified from total RNA isolated from leg muscle (not shown). On the Northern blot shown in Fig. 4A, no significant hybridization of the skeletal muscle probe with a 5.5-kb skeletal transcript occurring in the heart (lane 3) can be seen, possibly due to the lower sensitivity of this method in comparison with RT-PCR and in situ hybridization. We concluded that the cardiac isoform of myomesin is exclusively expressed in the embryonic chicken heart already, whereas the second isoform is mainly expressed in skeletal muscle tissues and to a lesser extent in heart (Fig. 5B).

The Two Chicken Isoforms Are Generated by Alternative Splicing—The first 4595 coding bases of the heart and skeletal chicken myomesin cDNAs are identical; therefore, it is possible that the two types of transcripts are generated from the same gene, because even if the identical amino acid sequences were derived from two or more genes, many silent mutations would be expected in such long pieces of coding sequences. Although the generation of the heterogenous sizes of the heart and skeletal transcripts cannot be fully explained at present, the data indicate that at least two types of transcripts are generated by alternative splicing in the 3'-part of the gene leading to isoform-specific C termini. This was confirmed using different sets of primers derived from the isolated cDNAs (see "Experimental Procedures") in PCR amplifications with chicken genomic DNA, and the results are shown in Fig. 6.

The longest product of about 12 kb was obtained using the sense primer Se1 located at nucleotides 4329–4347 of the cDNA shown in Fig. 2, in a region where the tissue-specific transcripts are identical (therefore called "common region"), and using the heart-specific antisense primer An6 (Fig. 2, positions 5173–5193). This fragment containing part of a myomesin gene was subcloned to yield the clone SB 371 (Fig. 1A), which was partially sequenced (Fig. 6). Additional amplifications were carried out on both the subcloned genomic fragment Se1-An6 (SB371) and genomic chicken DNA, resulting in an identical pattern of bands shown in Fig. 6B. All amplified fragments were arranged according to the occurrence of the primers used, and the fragment lengths were estimated from agarose gels, resulting in a partial gene structure of the 3'-end of a myomesin gene encoding both isoforms (Fig. 6A).

Splice sites and coding regions were localized on clone SB371 by partial sequencing using primers as indicated by arrows in Fig. 6A. In the 5'-end of the genomic insert, sequences...
common (black) to both myomesin transcripts were found as expected. Interestingly, the last 216 bp of the common sequence are situated in a complex exon containing common as well as skeletal muscle-specific coding (open box) and nontranslated sequences. An additional splice junction was identified at the fusion point of both types of sequences. This GTAAG donor sequence (Fig. 6A, D4) matches perfectly with the consensus GTRAG for 5'-splice sites of introns (Kornblihtt et al., 1984; Petersen et al., 1983). Several kb further downstream from the 3'-end, heart-specific coding (hatched boxes) and nontranslated sequences were identified (Fig. 6B). This partial gene structure clearly indicates that the mechanism of alternative splicing must be active in chicken tissues, and the presence of splice sites in all necessary positions confirms our hypothesis. The generation of the 3'- and the 5'-end of the mRNAs, however, is not within the scope of this paper, but it is under investigation.

**DISCUSSION**

Here we report the occurrence of tissue-specific chicken myomesin isoforms generated by alternative splicing and the identification of three species of myomesin mRNA with different molecular sizes. While 94% of the coding sequences are shared in all transcripts, tissue-specific sequences are present in their 3'-ends leading to a C-terminal segment of 98 amino acids in cardiac myomesin and 20 amino acids in the skeletal muscle isoform. The faster mobility in SDS-polyacrylamide gel electrophoresis of myomesin found in chicken skeletal muscle versus myomesin isolated from heart muscle (Grove et al., 1985) can now be explained by different C-terminal domains resulting in proteins of distinguishable calculated molecular masses of 182 kDa (1680 residues) for cardiac myomesin and 174 kDa (1552 residues) for skeletal myomesin. The skeletal muscle-specific C terminus of myomesin shares high homology with the C-terminal sequence of rat myomesin derived from an embryonic heart cDNA library as well as the C-terminal sequence of human 190 kDa protein (Vinkemeier et al., 1993) (Fig. 3B). Since the cDNA of the 190 kDa protein is derived from skeletal muscle, the homology between this protein and the chicken skeletal muscle myomesin was not surprising. The homology between the C termini of rat myomesin derived from heart and chicken skeletal muscle myomesin, however, was unexpected (Fig. 3B) but, as discussed below, can be explained by the possible absence of a tissue-specific isoform in mammalian sarcomeric muscle.

Another type of myomesin isoform diversity was found by Northern blot analysis, at the level of mRNA size. In chicken heart RNA, two transcripts of 7.5 and 9.0 kb with the heart-specific C-terminal sequence occur (Fig. 4C), while in skeletal muscle a single myomesin mRNA of 5.5 kb including the skeletal muscle-specific C terminus is present exclusively (Fig. 4A). All three chicken transcripts contain the complete common coding sequences because they were detected equally well with probes derived from any region of the common part of the cardiac myomesin cDNA (Fig. 4B). Multiple tissue-specific promoters of a chicken myomesin gene or additional tissue-specific splice events might explain the observed transcript sizes. A similar heterogeneity of mRNAs appears to be absent from mammalian sarcomeric muscle indicated by Northern blot analysis of rat RNA (Fig. 4D). RNA probed with either a 3'-coding rat probe corresponding to tissue-specific sequences of the chicken transcripts or with a 5'-coding rat probe revealed a unique transcript of 5.5 kb in both rat skeletal muscle and heart. Analysis of mouse RNA (not shown) gave the same results as in rat. This is not surprising, since the available partial mouse cDNA sequences are conserved to over 90% with rat. These results are consistent with the existence of only one type of myomesin transcript in rat and mouse tissues.

The heart- and skeletal muscle-specific probes were used for in situ localization of myomesin mRNAs in 8-day-old chicken embryos. The heart specificity of the probe was confirmed by its exclusive hybridization to the heart, while the skeletal muscle-specific probe mainly hybridized to the developing skeletal muscle, and traces were detected in heart. On the other hand, the same skeletal muscle probe gave no signal on Northern blot analysis of embryonic heart RNA isolated from 17-day-old embryos (Fig. 4A). This discrepancy could be due to decreasing expression levels of skeletal myomesin in heart during late embryogenesis or to differential sensitivity of detection in the two assays. The expression of myomesin during chicken development is under study, and forthcoming results will shed more
light on its developmental regulation.

The occurrence of multiple chicken myomesin mRNA species and the different 3'-ends of the cDNAs can be explained by tissue-specific alternative splicing. The partial structure of the 3'-end of a myomesin gene shows (Fig. 6B) that exons encoding the C termini of both isoforms are present on the same genomic fragment separated by introns at a distance of about 4 kb. The exons encoding the heart-specific form follow downstream from the skeletal sequence, which is part of a complex exon composed of common as well as skeletal sequence. Therefore we propose that alternative splicing leads to tissue-specific transcripts in skeletal muscle and heart. During RNA processing, the cardiac transcripts have to be generated by splicing of the heart-specific exons to the common part of the complex exon, thereby deleting the skeletal muscle-specific sequence. The presence of 5'- and 3'-splice sites in all necessary positions confirms our hypothesis (Fig. 6A). However, none of the three transcripts detected in chicken can be the precursor itself, because none of them contains both the skeletal and the cardiac 3'-element (Fig. 4). It remains to be seen whether the cardiac splice variant is specific for birds or occurs in other species, too.

Data from genomic PCR have shown that fragments from the 3'-end of a chicken myomesin gene encoding both isoforms can be amplified using different combinations of a sense primer derived from common sequences and either a skeletal muscle- or a heart-specific antisense primer. The obtained patterns of bands were identical with genomic DNA and the clone SB371 (Fig. 6), suggesting the presence of only one myomesin gene in the chicken genome. However, no definitive proof can be given, because only a partially known gene structure and incomplete data of the untranslated sequences are available, and therefore the existence of a second myomesin gene cannot be excluded.

In addition to the isoform diversity in chicken, an interesting difference was found in the sequences of chicken and mammalian N-terminal domains of myomesin. Eight consecutive copies of a KQSTAS-like motif are present in the N-terminal domain of the human 190K protein (Fig. 7, line III; Vinkemeier et al., 1993) and three copies in rat myomesin (Fig. 7, line I), but no similar sequence is encoded by the chicken cDNA (Fig. 7, line I). The absence of this particular sequence has been verified in two clones from different origins, in clone SB233 isolated from a heart library and in clone SB367 generated by RT-PCR from leg RNA, where no cardiac myomesin is expressed (Fig. 5A). Recently, we also found three copies of the same motif in the N-terminal domain of mouse skelemin (Fig. 7, line IV; Price and Gomer, 1993). This hexapeptide might be characteristic for mammalian M-band-associated proteins, but it seems not to be essential for the myomesin function in general. It will be interesting to express chicken myomesin in rat cardiomyocytes in order to see whether and how it functions in mammalian cells.

Inspection of the amino acid sequences covered by the modular part of the chicken myomesin cDNA reveals a secondary structure composed of 14 domains, which have been numbered continuously from the N to the C terminus (Fig. 8). Both terminal domains have been discussed above but not the rod part of myomesin, which consists of five fibronectin type III domains (class I motif; Fig. 8, rectangles) and seven immunoglobulin-like domains (class II motif; Fig. 8, ovals) in the following arrangement: II-II-I-I-I-I-I-II-I-II-I-I-I-I. Both domain types are characteristic for all proteins belonging to the huge immunoglobulin superfamily. A similar rod has been found in human 190K protein (Vinkemeier et al., 1993), chicken and human M-protein (Noguchi et al., 1992; Vinkemeier et al., 1993), and mouse skelemin (Price and Gomer, 1993) for which an additional proline-serine-rich domain, which is inserted between the third and fourth fibronectin type III domains, but still containing the second desmin-like domain, which was postulated close to its C terminus (Price and Gomer, 1993). Homology was determined by the Lipman-Pearson algorithm for protein sequences and by the Wilbur-Lipman algorithm for DNA sequences. All values are given in percentage of amino acid (A) and nucleotide sequence identity (B). (cm, chicken myomesin; hm, 190K protein (human myomesin); Dm, chicken M-protein; hm, human M-protein; ms, mouse skelemin.)

| A | cm | hm | Dm | hm | ms |
|---|----|----|----|----|----|
| cm | 76.6 | 51.5 | 50.4 | 76.7 | 75.9 | 51.1 | 52.8 | 78.1 |
| hm | 76.6 | 49.3 | 49.3 | 89.8 | 75.9 | 54.0 | 44.0 | 85.8 |
| Dm | 51.5 | 49.3 | 74.8 | 49.3 | 51.1 | 54.0 | 70.5 | 54.0 |
| hm | 50.4 | 49.3 | 74.8 | 50.0 | 52.8 | 44.0 | 70.5 | 54.2 |

Table I. Myomesin shares higher homology with skelemin than with M-protein. Partial protein and cDNA sequences covering the modular part of chicken myomesin were aligned with the corresponding sequences of M-protein and skelemin using the DNA Star software package. All sequences encode five fibronectin type III domains (class I motifs) and seven immunoglobulin-like domains (class II motifs) in a conserved arrangement: II-II-I-I-I-II-II-I-II. For skelemin a truncated sequence was used, omitting the proline-serine-rich domain, which is inserted between the third and fourth fibronectin type III domains, but still containing the second desmin-like domain, which was postulated close to its C terminus (Price and Gomer, 1993). Homology was determined by the Lipman-Pearson algorithm for protein sequences and by the Wilbur-Lipman algorithm for DNA sequences. All values are given in percentage of amino acid (A) and nucleotide sequence identity (B). (cm, chicken myomesin; hm, 190K protein (human myomesin); Dm, chicken M-protein; hm, human M-protein; ms, mouse skelemin.)

Acknowledgments—We thank Dr. J. P. Hossle for providing first cDNA clones and Dr. K. R. Chien (University of California), for the precious gift of the Agt 11 cDNA library of embryonic rat heart. We are also grateful to Dr. H. M. Eppenberger for continued interest and support and to Drs. M. Price (Rice University), E. Strehler (Dept. of Biochemistry and Molecular Biology Mayo Clinic, Rochester, MN), D. Fürst (Max Planck Institute for Biophysical Chemistry, Goettingen), M. Gautel (European Molecular Biology Laboratory, Heidelberg), and D. Helfman (Cold Spring Harbor Laboratory, Cold Spring Harbor, NY) for reading the manuscript.
REFERENCES

Carlsson, E., and Thornell, L.-E. (1987) Cell Tissue Res. 248, 169–180
Chomczynski, P., and Sacci, N. (1987) Anal. Biochem 162, 156–159
Degerlind, A., Friberg, K., Bean, A. J., and Hokfelt, T. (1992) Histochemistry 98, 39–49
Doetschmann, T. C., and Eppenberger, H. M. (1984) Eur. J. Cell Biol. 33, 265–274
Edman, A.-C., Squire, J. M., and Sjöström, M. (1988) J. Ultrastruct. Mol. Struct. Res. 100, 1–12
Eppenberger, H. M., Perriard, J.-C., Rosenberg, U. B., and Strehler, E. (1981) J. Cell Biol. 89, 185–193
Fürst, D. O., Nave, R., Osborn, M., and Weber, K. (1989) J. Cell Sci. 102, 769–778
Grove, B. K. (1989) Crit. Rev. Neurobiol. 4, 201–233
Hossle, J. P., Rosenberg, U., Schafer, B., Eppenberger, H. M., and Perriard, J.-C. (1986) Nucleic Acids Res. 14, 1449–1463
Knappeis, G. G., and Carlsen, F. (1968) J. Cell Biol. 38, 202–211
Kornblüth, A. R., Vibe-Pedersen, K., and Sahl, P., Sottrup-Jensen, L., and Magnusson, S. (1983) Proc. Natl. Acad. Sci. U.S.A. 80, 137–141
Labeit, S., Barlow, D. P., Gautel, M., Gibson, T., Hsieh, C.-L., Francke, U., Leonard, K., Wardale, J., Whiting, A., and Trinick, J. (1990) Nature 345, 273–276
Mani, R. S., and Kay, C. M. (1978) Biochim. Biophys. Acta 536, 134–141
Masaki, T., and Takaiti, O. (1972) J. Biochem. 71, 355–357
Masaki, T., and Takaiti, O. (1974) J. Biochem. 75, 367–380
Nave, R., Fürst, D. O., and Weber, K. (1989) J. Cell Biol. 109, 2177–2187
Noguchi, J., Yanagisawa, M., Imamura, M., Kasuya, Y., Sakurai, T., Tanaka, T., and Masaki, T. (1992) J. Biol. Chem. 267, 20302–20310
Pask, H. T., Jones, K. L., Luther, P. K., and Squire, J. M. (1994) J. Muscle Res. Cell. Motil. 15, 633–645
Petersen, T. E., Thogersen, H. C., Skorzena, K., Vibe-Pedersen, K., Sahl, P., Sottrup-Jensen, L., and Magnusson, S. (1983) Proc. Natl. Acad. Sci. U.S.A. 80, 137–141
Price, M. G., and Squire, J. M. (1977) J. Mol. Biol. 104, 1161–1170
Price, M. G., and Squire, J. M. (1980) J. Mol. Biol. 141, 409–439
Strehler, E. E., Carlsson, E., Eppenberger, H. M., and Thornell, L.-E. (1983) J. Mol. Biol. 166, 441–458
Strehler, E. E., Carlsson, E., Eppenberger, H. M., and Thornell, L.-E. (1983) Mol. Biol. 166, 441–458
Sassoon, D. A., Garner, J. A., and Buckingham, M. (1988) Development 104, 155–164
Schafer, B. W., and Perriard, J.-C. (1988) J. Cell Biol. 106, 1161–1170
Sjöström, M., and Squire, J. M. (1987) J. Mol. Biol. 199, 49–68
Street, S. F. (1983) J. Cell Phys. 114, 346–364
Strehler, E. E., Pelloni, G., Hellmann, C. W., and Eppenberger, H. M. (1979) Exp Cell Res. 124, 39–45
Strieder, E. E., Carlsson, E., Eppenberger, H. M., and Thornell, L.-E. (1983) J. Mol. Biol. 166, 441–458
Turner, D. C., Wallimann, T., and Eppenberger, H. M. (1973) Proc. Natl. Acad. Sci. U.S.A. 70, 702–705
Vinkemeier, U., Obermann, W., Weber, K., and Fürst, D. O. (1993) J. Cell Biol. 106, 329–330
Wallimann, T., Moser, H., and Eppenberger, H. M. (1983a) J. Muscle Res. Cell Motil. 4, 429–441
Wallimann, T., Doetschmann, T. C., and Eppenberger, H. M. (1983b) J. Cell Biol. 96, 1722–1729
Wang, K., and Ramirez-Mitchell, R. (1983) J. Cell Biol. 96, 562–570
Weber, F. E., Vaughan, K. T., Reinach, F. C., and Fischman, D. A. (1993) FEBS Lett. 316, 661–669
Zhu, H., Garcia, A. V., Ross, R. S., Evans, S. M., and Chien, K. R. (1991) Mol. Cell. Biol. 11, 2273–2281

Isolation and Characterization of Myomesin cDNAs