Proc. Natl. Acad. Sci. USA
Vol. 88, pp. 8024–8028, September 1991
Developmental Biology

Expression of M-cadherin, a member of the cadherin multigene family, correlates with differentiation of skeletal muscle cells

(polymerase chain reaction/cDNA isolation)

MICHAEL DONALIES*, MATTHIAS CRAMER*, MARTIN RINGWALD†, AND ANNA STARZINSKI-POWITZ*‡

*Institut für Genetik der Universität zu Köln, Zülpicher Strasse 47, D-5000 Cologne 1, and †Max-Planck-Institut für Immunbiologie, Stübeweg 51, D-7800 Freiburg, Federal Republic of Germany

Communicated by François Jacob, June 21, 1991 (received for review April 17, 1991)

ABSTRACT **Cadherins, a multigene family of transmembrane glycoproteins, mediate Ca^{2+}-dependent intercellular adhesion. They are thought to be essential for the control of morphogenetic processes, including myogenesis. Here we report the identification and characterization of the cDNA of another member of the cadherin family, M-cadherin (M for muscle), from differentiating muscle cells. The longest open reading frame of the cDNAs isolated contains almost the entire coding region of the mature M-cadherin as determined by sequence homology to the known cadherins. M-cadherin mRNA is present at low levels in myoblasts and is upregulated in myotube-forming cells. In mouse L cells (fibroblasts), M-cadherin mRNA is undetectable. This expression pattern indicates that M-cadherin is part of the myogenic program and may provide a trigger for terminal muscle differentiation.**

Cadherins are a multigene family of Ca^{2+}-dependent cell adhesion molecules. They are transmembrane glycoproteins consisting of an extracellular domain divided into subdomains EC1–EC5, a transmembrane region, and a cytoplasmic domain (1, 2). The extracellular domains mediate Ca^{2+}-dependent intercellular adhesion by homophilic interactions (1, 3). The different members of the cadherin family share a defined pattern of sequence homology. The homologies are most prominent in the cytoplasmic domain and in EC1 and EC2 and much less so in EC5 and in the transmembrane region (2). The binding properties and specificities of the adhesive function are located in the N-terminal part of the molecules (3, 4).

Three members of the family, the neuronal (N), epithelial (E), and placental (P) cadherins, have been characterized molecularly in the mouse (5–7) and in other species. Recently, two additional cadherins, truncated (T) cadherin (8) and EP-cadherin (9), have been identified in chicken and *Xenopus*, respectively. Each member of the cadherin family exhibits a tissue- and development-specific pattern of expression. The cadherins have been shown to be involved in morphogenetic processes such as guidance of nerves (10), polarization of epithelial cells (11), and myogenesis (12). Thus, cadherins contribute to cell–cell communications during development and maintenance of tissues and thereby could influence these processes.

In myogenic cells, N-cadherin has been described in myoblasts and myotubes in culture (12, 13). *In situ* it is transiently expressed in skeletal muscle cells during development (1). To date, no other cadherin has been identified in skeletal muscle cells *in situ* or in cell culture. There has been only one report suggesting that anti-N-cadherin antibodies detect a second cadherin in developing rat L6 myotubes (13).

Here we report the identification of a member of the cadherin family, M-cadherin (M for muscle), in myogenic mouse cells. M-cadherin is not found in fibroblasts and is expressed at low levels in myoblasts. It is upregulated after induction of myotube formation, indicating a specific function in skeletal muscle cell differentiation.§

MATERIAL AND METHODS

Cells. C2 cells, a myogenic mouse cell line originally isolated by Yaffe and Saxel (14), were grown in Waymouth medium containing 20% fetal bovine serum. Formation of myotubes was induced by switching cells to Dulbecco's modified Eagle's medium supplemented with insulin (10 μg/ml). For immunofluorescence studies, cells were grown on coverslips coated with calf skin collagen (Sigma) and were fixed with methanol.

Amplification, Cloning, and Analysis of cDNA. Cytoplasmic RNA was isolated from C2 myoblasts and C2 myotubes according to published procedures (15). For first-strand cDNA synthesis, RNA was hybridized to oligonucleotide primer P1 (Fig. 1) and the primer was extended with avian myeloblastosis virus reverse transcriptase (Boehringer Mannheim). The polymerase chain reaction (PCR) was performed (16) using primers P1 and P2 (Fig. 1). After analysis by electrophoresis in 1% agarose gels, PCR products were treated with Klenow enzyme in the presence of all four dNTPs to improve cloning efficiency and cloned into the *Sma* I site of plasmid vector pUC19 (Boehringer Mannheim). Nucleotide sequences of selected PCR cDNA clones were determined by the enzymatic dideoxy chain-termination method (17) on double-stranded DNA with a T7 sequencing kit (Pharmacia).

Isolation of cDNA Clones. M-cadherin cDNA clones were isolated by plaque hybridization (18) from a λgt11 cDNA library made from RNA of C2 myotubes (gift from John Merlie, Washington University, St. Louis). The probe for screening the library was a digoxygenin-labeled PCR clone that had ≈60% amino acid sequence identity to the known cadherins. cDNA clones hybridizing to this clone were recloned into plasmid vector pUC19 and subjected to nucleotide sequence analysis.

General Molecular Methods. All other RNA and DNA techniques used in this work were performed according to standard procedures (18). These include Northern blots, random nucleotide priming for labeling of hybridization probes, and size determination of cDNA fragments.

Antibodies and Immunofluorescence Assays. Rabbit polyclonal antibodies were raised against the synthetic pentadecapeptide FSIDKFTGRVYLNAT (purchased from Multiple Peptide Systems, San Diego) deduced from subdomain EC1. For immunization this peptide, Mcad1, was coupled to keyhole limpet hemocyanin. The antibody titer was 1:24,000 when tested against Mcad1–bovine serum albumin conjugates in an ELISA. Specificity of these reactions was verified by com-

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. §1734 solely to indicate this fact.

‡To whom reprint requests should be addressed.
§The sequence reported in this paper has been deposited in the GenBank data base (accession no. M74541).

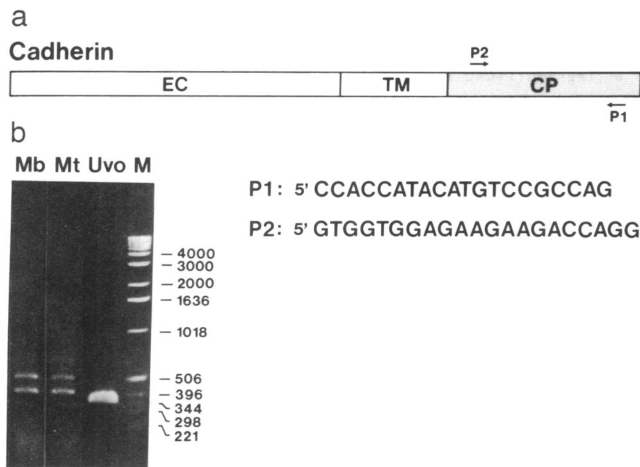

FIG. 1. (*a*) Schematic representation of cadherin protein structure and location and sequence of synthetic oligonucleotide primers (P1 and P2) used for amplification of cadherin sequences from mouse C2 myoblasts and myotubes. EC, extracellular domain; TM, transmembrane domain; CP, cytoplasmic domain. (*b*) Ethidium bromide-stained agarose gel showing P1/P2-amplified PCR products from C2 myoblasts (lane Mb) and C2 myotubes (lane Mt) as well as from uvomorulin (E-cadherin) cDNA (lane Uvo) used as positive control. Sizes in marker lane M are given in base pairs.

paring inhibition by the homologous peptide with inhibition by a peptide of unrelated sequence. Heat-inactivated antiserum was used for immunofluorescence staining (19). Second antibodies were affinity-purified fluorescein-labeled goat anti-rabbit antibodies (Dianova, Hamburg, F.R.G.) at 25–30 μg/ml. Negative control reactions were done by incubation of the cells with preimmune serum. An Axiophot microscope (Zeiss) equipped for immunofluorescence was used for photographs.

RESULTS

Isolation and Sequence Determination of M-cadherin cDNA. The basic assumption for the experimental approach to isolate M-cadherin by PCR was that regions of homology between N-, E-, and P-cadherin should be shared by a novel cadherin. The best candidate domain for such a homology search was the cytoplasmic domain of the cadherins, since they are 60–80% identical at the amino acid sequence level. Two regions of nucleotide homology were identified by computer analysis in the cytoplasmic domains of the three known mouse cadherins. Consensus sequences were derived from these conserved nucleotide sequence elements and used for the synthesis of oligonucleotide primers P1 and P2 (Fig. 1) for PCR. First-strand cDNAs from RNA of mouse C2 myoblasts and developing myotubes were amplified and PCR products were cloned into plasmid vector pUC19. In the known cadherin sequences, the distances between the PCR primers are 379 (N-cadherin), 358 (E-cadherin), and 361 (P-cadherin) nucleotides. Thus, the fragment of a novel cadherin sequence was expected to be about this length. Nucleotide sequences of 39 clones that had about the expected size were determined and their amino acid sequences were deduced. Subsequently, PCR clones were grouped according to their amino acid homologies to the known cadherins. Of the clones sequenced, one was 100% identical to N-cadherin, which has been shown to be expressed in myogenic cells (12, 13). A second clone, showing 60% identity and thus being very similar to the known cadherins, most likely represented a different and independent cadherin, which we named M-cadherin (M for muscle). The rest of the clones showed no significant homologies to proteins of the cadherin family.

From a λgt11 cDNA library prepared from RNA of C2 myotubes (gift from John Merlie), we isolated additional, overlapping M-cadherin cDNAs and determined their nucleotide sequences. Amino acid sequences were deduced from the nucleotide sequence (Fig. 2) obtained and the open reading frames were analyzed by comparison to the peptide sequences of murine N-cadherin (5), E-cadherin (6), and P-cadherin (7). This analysis showed that an open reading frame of 730 amino acids exhibits the highest homologies to the known cadherins (Fig. 3). In addition, almost all of the 3′ untranslated region of M-cadherin mRNA was identified as determined by PCR analysis (data not shown).

Analysis of the M-cadherin protein sequence revealed a hydrophobic region between amino acids 536 and 573. This putative transmembrane region is followed by a stretch of positively charged amino acids, a composition typical for transmembrane regions. Furthermore, in subdomain EC5 four cysteine residues are conserved in N-, P-, E-, and M-cadherin, suggesting an important role for these residues in the cadherin structure. Putative glycosylation sites in M-cadherin differ from those in the known cadherins (Fig. 4 and ref. 2). Finally, overall comparison of M-cadherin protein sequence to the known cadherin proteins indicated that M-cadherin has a domain structure that is similar to those of N-, E-, and P-cadherin (Figs. 3 and 4).

The homology of M-cadherin to the other cadherins is most prominent in the putative cytoplasmic domain and in subdomains EC1 and EC2, whereas it is much less so in the subdomains EC3–EC5 and the transmembrane region (Figs. 3 and 4; Table 1). This pattern of homology is very similar to the one that has been identified between the known cadherins. Overall, M-cadherin is less homologous to N-, E-, and P-cadherin than N-, E-, and P-cadherin are to one another (Fig. 4).

```
1     GAGAACCACA AACGCCTCCC CTACCCACTT GTGCAGATCA AGTCTGACAA ACAGCAGCTA
61    GGCAGTGTCA TCTACAGCAT CCAGGGTCCC GGAGTGGATG AGGAGCCCCG AAATGTCTTC
121   TCTATCGACA AGTTCACTGG GAGGGTGTAC CTCAACGCCA CGCTGGACCG AGAGAAGACG
181   GACCGCTTCA GGCTAAGGGC CTTTGCCTTG GACTTGGGTG GCTCTACCCT GGAGGACCCC
241   ACGGACCTGG AGATCGTTGT GGTGGATCAA AATGACAACC GGCCAGCCTT CCTACAGGAT
301   GTGTTCAGAG GCCACATCCT GGAGGGTGCC ATCCCAGGCA CCTTCGTCAC CAGGGCTGAG
361   GCCACAGATG CCGACGACCC AGAGACAGAC AATGCAGCCC TCAGGTTCTC CATCCTGGAG
421   CAGGGCAGCC CTGAGTTCTT CAGCATCGAC GAGCACACGG GAGAGATCCG CACGGTGCAA
481   GTGGGGCTGG ACCGTGAGGT GGTGGCTGTG TATAACCTGA CCTTGCAAGT GGCAGACATG
541   TCGGGAGACG GACTCACTGC CACAGCCTCG GCCATCATCT CCATAGATGA TATCAACGAC
601   AATGCCCCTG AGTTCACCAA GGATGAGTTC TTTATGGAGG CTGCAGAGGC TGTCAGTGGA
661   GTGGACGTGG GACGGCTCGA GGTGGAAGAC AAGGACCTGC CTGGTTCCCC CAACTGGGTG
721   GCCAGGTTCA CCATCCTTGA AGGTGATCCT GATGGGCAGT TCAAGATCTA CACAGACCCT
781   AAGACCAATG AGGGTGTGCT GTCCGTGGTC AAGCCCCTGG ACTATGAGAG CCGTGAGCAG
841   TATGAGCTCA GAGTGTCTGT ACAAAACGAG GCCCCGCTGC AGGCAGCTGC CCCTCGGGCT
901   CGGCGGGGCC AGACCAGGGT CAGCGTGTGG GTTCAGGACA CCAACGAAGC TCCAGTGTTT
961   CCAGAGAACC CACTGAGGAC GAGCATAGCT GAAGGAGCCC CCCCAGGCAC CTCTGTGGCC
1021  ACCTTCTCTG CCAGAGACCC TGACACAGAA CAGCTGCAGA GAATCAGCTA CTCCAAAGAC
1081  TACGACCCAG AAGACTGGCT GCAAGTGGAC GGGGCCACAG GCAGGATCCA GACCCAGCGA
1141  GTGCTGAGCC CTGCTTCACC CTTTTTGAAG GACGGCTGGT ACAGAGCCAT CATCCTAGCC
1201  CTGGACAATG CCATTCCTCC TAGCACAGCC ACAGGCACCC TGTCCATCGA GATCCTAGAA
1261  GTCAACGACC ATGCCCCTGC ACTGGCTCTT CCTCCATCTG GCAGCCTGTG CAGTGAACCA
1321  GACCAGGGCC CTGGCCTCCT CTTGGGTGCC ACGGATGAGG ACCTGCCCCC ACACGGGGCC
1381  CCTTTCCACT TCCAGCTGAA CCCCAGAGTA CCAGATCTGG GCCGGAACTG GAGCGTCAGC
1441  CAGATTAACG TGAGCCATGC ACGCTTGCGG CTCCGACATC AGGTCTCCGA GGGCCTGCAT
1501  CGCCTGAGCC TGCTACTCCA GGACTCTGGG GAGCCACCCC AGCAGCGAGA GCAAACGCTG
1561  AACGTCACTG TGTGTCGCTG TGGGTCAGAT GGCACTTGCC TGCCCGGGGC TGCCGCGCTT
1621  CGAGGAGGAG GTGTAGGCGT CAGCTTGGGG GCACTGGTCA TTGTGCTGGC CAGCACCGTG
1681  GTCTTACTAG TTCTCATCCT GTTTGCCGCG CTCCGCACAC GGTTCCGGGG GCATTCCCGG
1741  GGCAAGAGTC TGTTGCATGG CCTACAAGAG GACCTTCGGG ACAACATCCT TAACTATGAT
1801  GAGCAAGGAG GCGGGGAGGA GGACCAGGAT GCATACGACA TAAACCAGCT GCGCCACCCA
1861  GTGGAGCCGA GGGCCACAAG CCGCTCTTTG GGCAGGCCAC CCCTGCGCAG GGATGCCCCC
1921  TTCAGCTATG TGCCACAGCC ACATCGAGTG CTTCCCACCA GCCCATCTGA CATTGCCAAC
1981  TTCATCAGTG ATGGCTTGGA GGCTGCGGAC AGCGACCCCA GCGTGCCTCC CTACGACACA
2041  GCTCTCATCT ATGACTACGA GGGAGATGGC TCTGTGGCAG GGACGCTGAG CTCCATTCTG
2101  TCCAGCCTGG GAGATGAAGA CCAGGACTAT GACTATCTCC GGGACTGGGG ACCCCGCTTT
2161  GCTCGGCTGG CGGACATGTA TGGACATCAG TGAGAGCCAG GGCCAAGGGC AGATGTGCTG
2221  TGTGGATACG CCCACTTGGC CCAACTAAGG GGCTCTCTCT TGGGATGTGC ACCCAGAAAT
2281  CCTATGAGGG TCAACAGCAT GGCCCATCCT TGGCTCCATG GCAGATAAAC TCAAGGAGGT
2341  CATCTCTGTG AGGTCCAAGA GAGGACTGAG TCAGGAGTGG TGGGAGAATC TGGAGCATGG
2401  ACTAGGCAGC TAGAGGGAGC ACTGTCCTGG CAGAGTGCAG AAGCCACCCT TAGTGCCCTG
2461  CTAGGGCTCA TCCCATCTTT GATTCCCAGT TGTGACTCTT GCCTCTGTAT GAAAAGCAGG
2521  CGTCTAAAGA GTGGATTTCA ATTAAAAGCA TACTATTTGG TGGAACCC
```

FIG. 2. Nucleotide sequence of M-cadherin cDNA. The codon underlined (nucleotides 2191–2193) indicates the putative stop codon of the M-cadherin reading frame.

```
EC1 M  ENHK-RLPYPLVQIKSDKQQLGSVIYSIQGPGVDEEPRNVFSIDKFTGRVYLNATLDREKT
    E  ENEKGEFPKNLVQIKSNRDKETKVFYSITGQGADKPPVGVFIIERETGWLKVTQPLDREAI
    N  ENSRGPFPQELVRIRSDRDKNLSLRYSVTGPGADQPPTGIFIINPISGQLSVTKPLDRELI
    P  ENGKGPFPQRLNQLKSNKDRGTKIFYSITGPGADSPPEGVFTIEKESGWLLLHMPLDREKI
       **    *   *    *         **  *  *  *  *   * *    *        ****

    M  DRFRLRAFALDLGGSTLEDPTDLEIVVVDQNDNRPAF
    E  AKYILYSHAVSSNGEAVEDPMEIVITVTDQNDNRPEF
    N  ARFHLRAHAVDINGNQVENPIDIVINVIDMNDNRPEF
    P  VKYELYGHAVSENGASVEEPMNISIIVTDQNDNKPKF
           *   *   *   * *   * * * *** * *

EC2 M  LQDVFRGHILEGAIPGTFVTRAEATDADDP-ETDNAALRFSILEQGS--P--EFFSID
    E  TQEVFEGSVAEGAVPGTSVMKVSATDADDDVNTYNAAIAYTIVSQDPELPHKNMFTVN
    N  LHQVWNGSVPEGSKPGTYVMTVTAIDADDP-NALNGMLRYRILSQAPSTPSPNMFTIN
    P  TQDTFRGSVLEGVMPGTSVMQVTATDEDDAVNTYNGVVAYSIHSQEPKEPHDLMFTIH
         *  **  *** *     * *  **     *     *   *    *    *

    M  EHTGEIRTVQVGLDREVVAVYNLTLQVADMSGD---GLTATASAIISIDDINDNAPEF
    E  RDTGVISVLTSGLDRESYPTYTLVVQAADLQGE---GLSTTAKAVITVKDINDNAPVF
    N  NETGDIITVAAGLDREKVQQYTLIIQATDMEGNPTYGLSNTATAVITVTDVNDNPPEF
    P  KSTGTISVISSGLDREKVPEYRLTVQATDMDGE---GSTTTAEAVVQILDANDNAPEF
        ** *     *****   *  *   *   *   *     *   **  *     * *** * *

EC3 M  TKDEFFMEAAEAVSGVDVGRLEVEDKDLPGSPNWVARFTILEGDPDGQFKIYTDPKT
    E  NPSTYQGQVPENEVNARIATLKVTDDDAPNTPAWKAVYTVVNDPD-QQFVVVTDPTT
    N  TAMTFYGEVPENRVDVIVANLTVTDKDQPHTPAWNAAYRISGGDPTGRFAILTDPNS
    P  EPQKYEAWVPENEVGHEVQRLTVTDLDVPNWPAWRATYHIVGGDDGDHFTITTHPET
              *         *  *  *  *  *   *  *  *           *   * *

    M  NEGVLSVVKPLDYESREQYELRVSVQNEAPLQAAAPRARRGQTRVSVVVQDTNEAPVF
    E  NDGILKTAKGLDFEAKQQYILHVRVENEEPFEGSL---VPSTATVTVDVVDVNEAPIF
    N  NDGLVTVVKPIDFETNRMFVLTVAAENQVPLAKGIQHPPQSTATVSVTVIDVNENPYF
    P  NQGVLTTKKGLDFEAQDQHTLYVEVTNEAPFAVKL---PTATATVVVHVKDVNEAPVF
       *  *      *   *  *        *  *    *    *                 *  *  *  *  ** *  *

EC4 M  PENPLRTSIAEGAPPGTSVATFSARDPDT-EQLQRISYSKDYDPEDWLQVDGATGRIQ
    E  MPAERRVEVPEDFGVGQEITSYTAREPDTFMDQK-ITYRIWRDTANWLEINPETGAIF
    N  APNPKIIRQEEGLHAGTMLTTLTAQDPDRYMQQN-IRYTKLSDPANWLKIDPVNGQIT
    P  VPPSKVIEAQEGISIGELVCIYTAQDPDK-EDQK-ISYTISRDPANWLAVDPDSGQIT
                *    *       *   **         * *    *    **       * *

    M  TQRVLSP-ASPFLKDGWYRAIILALDNAIPPSTATGTLSIEILEVNDHAP
    E  TRAEMDREDAEHVKNSTYVALIIATDDGSPIATGTGTLLLVLLDVNDNAP
    N  TIAVLDRESP-YVQNNIYNATFLASDNGIPPMSGTGTLQIYLLDINDNAP
    P  AAGILDREDEQFVKNNVYEVMVLATDSGNPPTTGTGTLLLTLTDINDHGP
                        *    * *   *    ****    ** *

EC5 M  ALALPPSGSLCSEPDQGPGLLLGATDEDLPPHGAPFHFQLNPRVPDLGRNWSVSQINV
    E  IPE-PRNMQFCQR-NPQP-HIITILDPDLPPNTSPFTAELTHGASV---NWTIEYNDA
    N  QVL-PQEAETCETPEPNS-INIAALDYDIDPNAGPFAFDLPLSPVTIKRNWTINRLNG
    P  IPE-PRQIIICNQ-SPVP-QVLNITDKDLSPNSSPFQAQLTHDSDI---YWMAEVSEK
          *     *              *  *  *   **   *             *

    M  SHAR--LRLRHQVSEGLHRLSLLLQDSGEPPQQREQTLNVTVCRCGSD--GTCLP---
    E  AQESLILQPRKDLEIGEYKIHLKLADNQNKDQ--VTTLDVHVCDCEGTV-NNCMK---
    N  DFAQLNLKIK-FLEAGIYEVPIIITDSGNPPKSNISILRVKVCQCD--SNGDCTDVDR
    P  G-DTVALSLKKFLKQDTYDLHLSLSDHGNREQ--LTMIRATVCDCHGQVFNDCPRPWK
            *                *                ** *         *

TM  M  GAAALRGGGVGVSLGALVIVLASTVVLLVLILFAALRT
    E  ---AG-IVAAGLQVPAILGILGGILALLILILLLLLFL
    N  ------IVGAGLGTGAIIAILLCIIILLILVLMFVVWM
    P  -------------GGFILPILGAVLALLTLLLALLLLV
                       *     ** * *

CP  M  RFRGHSR-GKSLLHGLQEDLRDNILNYDEQGGGEEDQDAYDINQLRHPVEPRATSRSLG
    E  RRR---TVVKEPLLPPDDDTRDNVYYYDEEGGGEEDQD-FDLSQLHRGLDARPE-----
    N  KRRDKERQAKQLLIDPEDDVRDNILKYDEEGGGEEDQD-YDLSQLQQPDTVEPDA--IK
    P  RKK---RKVKEPLLLPEDDTRDNVFYYGEEGGGEEDQD-YDITQLHRGLEARPE-----
               *  *     *  ***   *  *  ********   *   **

    M  RPPLRRDAPFSYVPQPH---RVLPTSPSDIANFISDGLEAADSDPSVPPYDTALIYDYE
    E  -VT-RNDVAPTLMSVPQYRPRPA--NPDEIGNFIDENLKAADSDPTAPPYDSLLVFDYE
    N  PVGIRRLDERPIHAEPQYPVRSAAPHPGDIGDFINEGLKAADNDPTAPPYDSLLVFDYE
    P  -VVLRNDVVPTFIPTPMYRPRPA--NPDEIGNFIIENLKPANTDPTAPPYDSLMVFDYE
           *         *     *      *   *  **   *   *  **  ****    ***

    M  GDGSVAGTLSSILSSLGDEDQDYDYLRDWGPRFARLADMYG---HQ
    E  GSGSEAASLSSLNSSESDQDQDYDYLNEWGNRFKKLADMYGGGEDD
    N  GSGSTAGSLSSLNSSSSGGDQDYDYLNDWGPRFKKLADMYGGG-DD
    P  GSGSDAASLSSLTTSASDQDQDYNYLNEWGSRFKKLADMYGGGEDD
       * ** *  ***   *    **** **  ** **  ******
```

FIG. 3. Comparison of protein sequence of M-cadherin, deduced from overlapping M-cadherin cDNA clones, to protein sequences of N-, E-, and P-cadherin (see also legend to Fig. 4). Amino acids identical in all four cadherins are marked by stars, and gaps are shown by hyphens. The M-cadherin protein sequence was divided into domains (EC, extracellular; TM, transmembrane; CP, cytoplasmic) and subdomains (EC1–EC5) according to Hatta *et al.* (2). Residues underlined have been discussed as a center of homophilic interaction in the cadherin family (3).

Analysis of M-cadherin RNA Expression. Poly(A)$^+$-enriched cytoplasmic RNA from C2 myoblasts, developing C2 myotubes, and mouse L-cell fibroblasts was analyzed by Northern blot hybridization (Fig. 5). C2 myoblasts expressed M-cadherin RNA at a low level. Upon induction of myotube formation the amount of M-cadherin RNA was increased. Reprobing this Northern blot with a skeletal muscle-specific troponin T cDNA fragment isolated from C2 cells (unpublished results; ref. 21) gave rise to a signal in RNA from C2 myotubes but not from C2 myoblasts (Fig. 5*c*). This suggests that the low level of M-cadherin mRNA in C2 myoblasts is not simply due to the presence of already-differentiated muscle cells.

The slight size heterogeneity of M-cadherin RNA is most likely not the result of degradation, since rehybridization with mouse cytochrome *c* oxidase cDNA gave rise to a band of 1.8 kilobases as expected (ref. 20 and Fig. 5*b*). It is also unlikely that the apparent heterogeneity of M-cadherin RNA is due to cross-hybridization with N-cadherin RNA. In C2 cells, N-cadherin RNA is detectable by PCR but not in Northern blots whereas M-cadherin RNA can be easily shown by hybridization (unpublished results and Fig. 5). Possibly, size heterogeneity of M-cadherin hybridization products is due to different lengths of poly(A) tails or to hybridization of other related sequences. Colinearity of M-cadherin cDNA with M-cadherin mRNA was shown by RNA-based PCR analysis (unpublished results). L-cell fibroblasts did not express M-cadherin RNA (Fig. 5*a*), a finding that was confirmed by PCR analysis (data not shown).

Expression of M-cadherin appears to be regulated in a similar manner during C2 cell differentiation as has been shown for other muscle-specific proteins such as troponin T (Fig. 5*c* and ref. 21) or myosin heavy and light chains (22) and thus might be part of the myogenic program.

Analysis of M-cadherin Protein Expression. Polyclonal antibodies were raised against a synthetic M-cadherin peptide (Mcad1) deduced from a region of the putative protein

Table 1. Protein sequence comparison of domains and subdomains of the individual members of the cadherin family

	% amino acid sequence identity						
Comparison	EC1	EC2	EC3	EC4	EC5	TM	CP
M:N	42	51	41	45	29	28	56
M:E	44	48	39	36	23	32	51
M:P	43	51	42	41	27	40	49
E:N	58	55	46	44	30	47	64
N:P	51	58	46	51	32	32	60
E:P	66	61	53	48	41	60	80

For details see legend to Fig. 4.

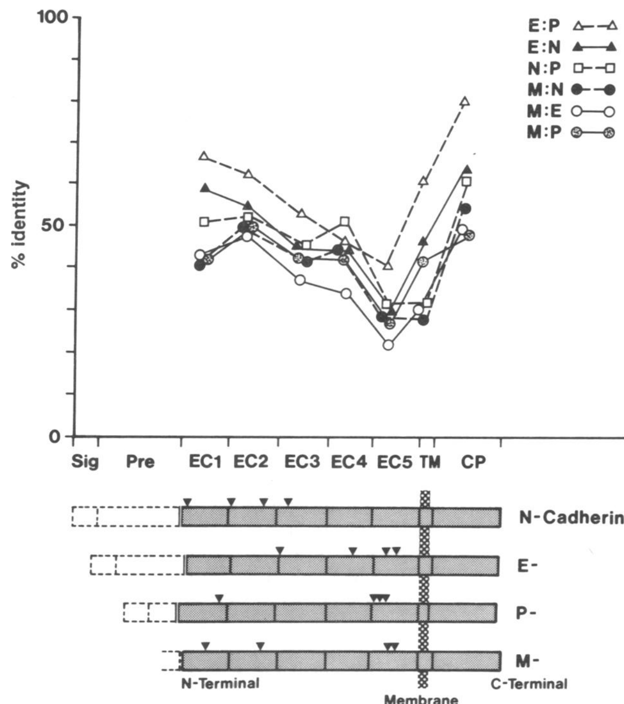

FIG. 4. (*Upper*) Homologies of peptide sequences in domains and subdomains between the different members of the cadherin family. Amino acids (aa) were compared in the following regions: E-cadherin aa 189–906; N-cadherin aa 260–996; P-cadherin aa 125–837; M-cadherin aa 1–730. For comparison, protein sequences were divided into extracellular subdomains EC1–EC5, the transmembrane domain (TM), and the cytoplasmic domain (CP) according to Hatta *et al.* (2). Precursor (Pre) and signal (Sig) peptide sequences are not available from M-cadherin and thus are not compared. (*Lower*) Schematic presentation of cadherin protein structures relative to each other. Consensus sequences of asparagine-glycosylation sites are indicated by triangles according to Hatta *et al.* (2).

sequence that exhibits low homology to the other cadherins (see Fig. 3). This peptide is located in subdomain EC1 and has the sequence FSIDKFTGRVYLNAT (for details, see *Material and Methods*). With these anti-Mcad1 antibodies, immunofluorescence studies were performed on C2 myotubes

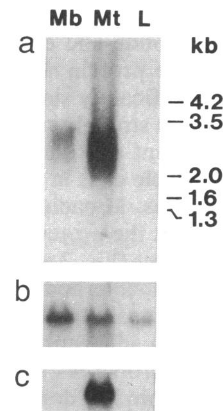

FIG. 5. Northern blot analysis of M-cadherin mRNA (*a*) in mouse C2 myoblasts (lane Mb), C2 myotubes at day 3 after induction of myotube formation (lane Mt), and mouse L-cell fibroblasts (lane L). ^{32}P-labeled probe was M-cadherin cDNA clone 37, which spans part of the cytoplasmic domain, the transmembrane region, and extracellular subdomains EC4 and EC5. kb, Kilobases. Rehybridization of the Northern blot with clone pAG82, a cDNA clone of cytochrome *c* oxidase I (20), indicates the integrity of the RNA samples (*b*) and rehybridization with a troponin T cDNA fragment isolated by our group (unpublished results) indicates muscle-specific induction of mRNA (*c*). Five micrograms of RNA was used per lane. Exposure to Kodak X-AR film was at −70°C for 30 hr (*a*) or overnight (*b* and *c*).

at day 3 after switching to differentiation medium and on C2 myoblasts (Fig. 6 *a* and *b*). Negative control experiments were done using preimmune serum instead of specific antiserum (Fig. 6 *c* and *d*). These studies indicate that a protein containing the peptide sequence deduced from M-cadherin cDNA is readily demonstrated in differentiating myotubes (Fig. 6*b*) but is not detectable in C2 myoblasts (Fig. 6*a*). It is not clear whether these antibodies fail to detect M-cadherin in C2 myoblasts, which do contain some M-cadherin mRNA (Fig. 5*a*), or whether M-cadherin mRNA is translated only after induction of differentiation.

DISCUSSION

In this paper we report the isolation and characterization of M-cadherin cDNA, another member of the cadherin multi-

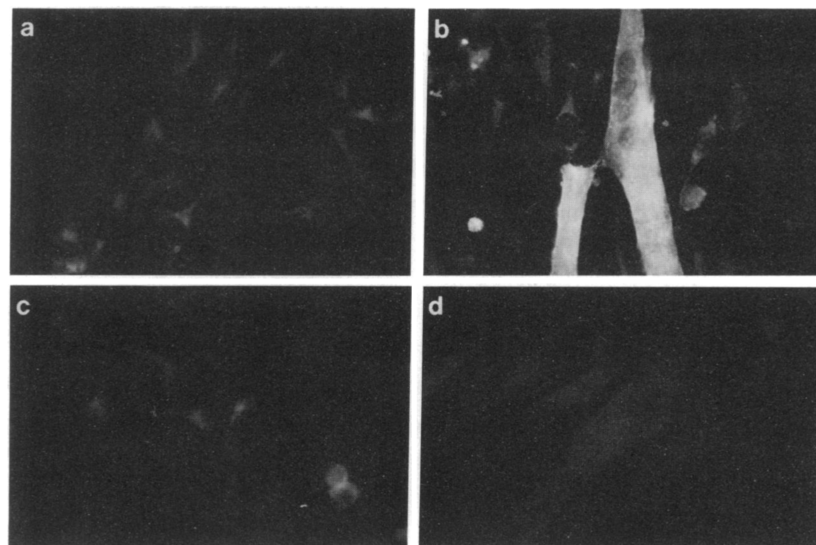

FIG. 6. Analysis of M-cadherin protein in C2 myoblasts (*a* and *c*) and developing C2 myotubes (*b* and *d*) by immunofluorescence. Staining with antiserum (*a* and *b*) or with preimmune serum (*c* and *d*). Exposure times of myoblasts were approximately 10 times longer than for myotubes. (Original magnifications were ×630.)

gene family, from differentiating mouse muscle cells. Analysis of M-cadherin nucleotide and deduced amino acid sequences revealed that the protein sequence as well as the domain structure are significantly similar to those of N-, E-, and P-cadherin. This result strongly suggests that M-cadherin is indeed a Ca^{2+}-dependent cell adhesion molecule that is expressed in skeletal muscle cells in addition to N-cadherin.

Despite these similarities, M-cadherin is less homologous to the known cadherins at the protein level than the known cadherins are to one another (Fig. 3 and Table 1; ref. 2). This applies to several of the domains and subdomains compared. Earlier data (3) have demonstrated that the N-terminal 113 amino acids of E- and P-cadherin carry the specificity for their homophilic binding properties and have suggested that amino acids 78–83 of this binding domain may be a center for the homophilic interaction. Comparison of these residues from M-cadherin with those in E-, N-, and P-cadherin revealed that this motif (underlined in Fig. 3) again diverges more in M-cadherin than in the other members of the gene family. Especially the amino acid trimer HAV is conserved in the known cadherins (3) but is different in M-cadherin. These observations suggest that M-cadherin separated early in evolution from a precursor of the other cadherins. In the context of evolution of cadherin genes, we note that the M-cadherin gene, like the E-cadherin gene, is located on mouse chromosome 8 whereas the N-cadherin gene is not (K. Kaupmann, A.S.-P. and H. Jockusch, unpublished work).

Our data show that M-cadherin is upregulated only after induction of myotube formation. If cadherins indeed play important roles in morphogenesis such as cell layer separation, cell recognition, and cell rearrangements (1), then the regulated expression of M-cadherin during terminal differentiation of skeletal muscle cells should help to ensure the proper temporal and spatial developmental pattern of the muscle phenotype *in situ*. Moreover, expression of M-cadherin seems to be regulated similarly to that of muscle-specific genes such as those for troponin T, myosins, or muscle creatine kinase (Fig. 5*c*; refs. 21–23). This raises the question of whether the same or overlapping factors that regulate expression of these genes, including myogenic factors, are also operative in the control of M-cadherin expression.

Little is known about the biological role of cadherins during muscle development. Earlier experiments (12), however, have indicated that N-cadherin mediates interaction of myoblasts, possibly in concert with the Ca^{2+}-independent cell adhesion molecule NCAM (24, 25). NCAM is expressed on myoblasts as well as on myotubes in culture and on regenerating fibers *in situ* but is undetectable on mature muscle fibers except at the motor endplate (26). Some of these studies (12, 24, 25) have also suggested that N-cadherin and NCAM are not sufficient to trigger induction of myotube formation. In fact, this might mean that an additional signal is required for myoblasts to fuse. The results presented in this report suggest that M-cadherin is a candidate for this signaling role. This, however, does not exclude the possibility that M-cadherin is also involved in the subsequent morphogenesis of muscle fibers.

Generally, it has been suspected that expression of more than one of the known cadherins might be necessary for the specificity of cadherin-mediated intercellular communication during morphogenetic events (1, 5). With regard to myogenesis, such an idea is in line with the following observations: (*i*) N-cadherin is apparently not sufficient to trigger myotube formation (12), (*ii*) study of the regulation of cadherins during development of rat L6 myotubes has suggested a second cadherin in these cells (13), and (*iii*) the identification of M-cadherin (this paper) proves molecularly that muscle cells contain at least two cadherins.

What could be the molecular action of a molecule like M-cadherin and what might be its functional relationship to N-cadherin in muscle cells? Skeletal muscle cell differentiation in the embryo as well as in adult regenerating muscle tissue certainly requires (*a*) communication of myoblasts prior to fusion, in order to distinguish between unrelated cell types and self, and (*b*) conversion of this recognition signal into the triggering signal for fusion and maturation of muscle fibers. In the first part of this process, of the Ca^{2+}-dependent adhesion molecules known, only N-cadherin might be involved, as has been indicated (12). Upon induction of development, M-cadherin is upregulated and might provide a trigger for the fusion process as well as the subsequent morphological steps towards maturation of muscle fibers.

We thank K. Rajewsky and R. S. Jack for critical reading of the manuscript; R. Kemler, D. Vestweber, and K. Zimmermann for helpful discussions; and C. Siewert for expert technical assistance. This work was supported by grants from the Bundesministerium für Forschung und Technologie to A.S.-P. and the Deutsche Forschungsgemeinschaft (Cr 70/2-1) to M.C. A Heisenberg Fellowship of the Deutsche Forschungsgemeinschaft to A.S.-P. is gratefully acknowledged.

1. Takeishi, M. (1988) *Development* **102,** 639–655.
2. Hatta, K., Nose, A., Nagafuchi, A. & Takeichi, M. (1988) *J. Cell Biol.* **106,** 873–881.
3. Nose, A., Tsuji, K. & Takeichi, M. (1990) *Cell* **61,** 147–155.
4. Ozawa, M., Engel, J. & Kemler, R. (1990) *Cell* **63,** 1033–1038.
5. Miyatani, S., Shimamura, K., Hatta, M., Nagafuchi, A., Nose, A., Matsanuga, M., Hatta, K. & Takeichi, M. (1989) *Science* **245,** 631–635.
6. Ringwald, M., Schuh, R., Vestweber, D., Eistetter, H., Lottspeich, F., Engel, J., Dölz, R., Jähnig, F., Epplen, J., Mayer, S., Müller, C. & Kemler, R. (1987) *EMBO J.* **6,** 3647–3653.
7. Nose, A., Nagafuchi, A. & Takeichi, M. (1987) *EMBO J.* **6,** 3655–3661.
8. Ranscht, B. & Bronner-Fraser, M. (1991) *Development* **111,** 15–22.
9. Ginsberg, D., DeSimone, D. & Geiger, B. (1991) *Development* **111,** 315–325.
10. Matsunaga, M., Hatta, K., Nagafuchi, A. & Takeichi, M. (1988) *Nature (London)* **334,** 62–64.
11. McNeill, H., Ozawa, M., Kemler, R. & Nelson, W. J. (1990) *Cell* **62,** 309–316.
12. Knudsen, K. A., Myers, L. & McElwee, S. A. (1990) *Exp. Cell Res.* **188,** 175–184.
13. Pouiliot, Y., Holland, P. C. & Blaschnik, O. W. (1990) *Dev. Biol.* **141,** 292–298.
14. Yaffe, D. & Saxel, O. (1977) *Nature (London)* **270,** 725–727.
15. Herget, T., Reich, M., Stüber, K. & Starzinski-Powitz, A. (1986) *EMBO J.* **5,** 659–664.
16. Zimmermann, K., Kautz, S., Hajdu, G., Winter, C., Whalen, R. G. & Starzinski-Powitz, A. (1990) *J. Mol. Biol.* **211,** 505–513.
17. Sanger, F., Nicklen, S. & Coulson, A. R. (1977) *Proc. Natl. Acad. Sci. USA* **74,** 5463–5467.
18. Sambrook, J., Fritsch, E. F. & Maniatis, T. (1989) *Molecular Cloning: A Laboratory Manual* (Cold Spring Harbor Lab., Cold Spring Harbor, NY).
19. Zimmermann, K., Herget, T., Salbaum, J. M., Schubert, W., Hilbich, C., Cramer, M., Masters, C. L., Multhaup, G., Kang, J., Lemaire, H.-G., Beyreuther, K. & Starzinski-Powitz, A. (1988) *EMBO J.* **7,** 367–372.
20. Murphy, D., Brickell, P. M., Latchman, D. S., Willison, K. & Rigby, P. W. J. (1983) *Cell* **35,** 865–871.
21. Putney, S. D., Herliky, W. C. & Schimmel, P. (1983) *Nature (London)* **302,** 718–721.
22. Silberstein, L., Webster, S. G., Travis, M. & Blau, H. M. (1986) *Cell* **46,** 1075–1081.
23. Jaynes, J. B., Chamberlain, J. S., Buskin, J. N., Johnson, J. E. & Hauschka, S. D. (1986) *Mol. Cell. Biol.* **6,** 2855–2864.
24. Knudsen, K. A., McElwee, S. A. & Meyers, L. (1990) *Dev. Biol.* **138,** 159–168.
25. Dickinson, G., Gower, H. J., Barton, C. H., Prentice, H. M., Elsom, V. L., Moore, S. E., Cox, R. D., Quinn, C., Putt, W. & Walsh, F. S. (1987) *Cell* **50,** 1119–1130.
26. Moore, S. E. & Walsh, F. S. (1985) *EMBO J.* **4,** 623–630.