μ-Protocadherin, a Novel Developmentally Regulated Protocadherin with Mucin-like Domains*

Received for publication, January 12, 2000, and in revised form, May 4, 2000
Published, JBC Papers in Press, May 5, 2000, DOI 10.1074/jbc.M000234200

Michael Goldberg‡‡§§, Courtney Peshkovsky‡§, Ashkan Shifteh‡, and Qais Al-Awqati††

From the Departments of ‡Pediatrics and §§Medicine and Physiology, College of Physicians and Surgeons of Columbia University, New York, New York 10032

Branching morphogenesis is a central event during the development of kidneys, lungs, and other organs. We previously generated a monoclonal antibody, 3D2-E9, that inhibited branching morphogenesis and caused widespread apoptosis. We now report the purification of its antigen and cloning of its full-length cDNA. Its cDNA encodes an integral membrane protein that contains four cadherin-like ectodomains and a thrice tandemly repeated region enriched in threonine, serine, and proline, similar to those of mucins. We thus term this protein μ-protocadherin, reflecting the hybrid nature of its extracellular region. μ-Protocadherin is expressed in two forms that are developmentally regulated, with the shorter isoform lacking the mucin-like repeats. Expression of the long isoform in heterologous cells results in adhesion of the expressing cells, suggesting that it is a new cell adhesion molecule. μ-Protocadherin contains both N and O glycosylations. It is expressed at lateral and basal surfaces of epithelia during kidney and lung development and is located in coated pits. Colocalization of μ-protocadherin with β-catenin was noted primarily at the junction of the lateral and basal membrane. The cytoplasmic domain contains four proline-rich regions, similar to SH3 binding regions. Thus, it is likely that adhesive interactions mediated by μ-protocadherin induce signaling events that regulate branching morphogenesis.

The mature mammalian kidney develops when an outgrowth of the wolffian duct, the ureteric bud, induces the metanephric mesenchyme to differentiate into the nephron. Induction is critical for further development since the metanephric blastema is destined to die (programmed cell death) unless it is rescued by inducers (1). Induced cells are then selected for further proliferation and differentiation. The ureteric bud is, in turn, induced to branch by the mesenchyme to form the collecting system of the kidney including the cortical collecting tubules. Critical to epithelial branching are the signals provided by the mesenchyme. Localized production of glial-derived neurotrophic factor by the mesenchyme signals its receptor ret on the ureteric bud tip to branch (2). The mesenchyme forms the epithelial cells of the rest of the nephron. As contact is established between these two cell types, a complex transformation occurs whereby the mesenchymal cells undergo conversion into an epithelial type, which later differentiate and segment into the various components of the nephron (3).

Cell to cell and cell to matrix interactions are critical for almost every stage of these complex morphogenetic programs. Among the mediators of these interactions are growth factors and their receptors (4), receptors for extracellular matrix molecules such as integrins (5), and cell adhesion molecules such as the cadherins (6). Cadherins are classically defined by their calcium-dependant homophilic cell adhesion properties. In addition to responding to signaling events, cadherins transduce signals into the cell (7). The differential expression of cadherins may drive nephrogenesis by stimulating aggregation by subsets of cells expressing the same type of cadherin (8).

To identify novel cell surface molecules critical for nephrogenesis, we generated a panel of monoclonal antibodies (mAb)1 to embryonic day 15 (E15) rat kidneys. The immunization protocol ensured that all molecules identified are authentic embryonic kidney molecules. Developmentally significant mAbs were selected based on their pattern of expression and/or their ability to inhibit tubulogenesis in vitro. Several of the antibodies disrupted development in vitro, suggesting that their antigens are necessary for development, at least in the organ culture system. One mAb, 3D2-E9, inhibited tubulogenesis by producing widespread apoptosis in induced mesenchymal cells (9). Remarkably, this antibody also inhibited pulmonary branching and caused cylindrical dilatation of the proximal airways in lung organ culture, suggesting that its antigen might play a more general role in branching morphogenesis (10). The 3D2-E9 mAb recognized a disulfide-linked integral membrane glycoprotein whose molecular mass was 100 kDa under reducing conditions. We present here the purification and cloning of the cDNA of its antigen. This novel protein is an integral membrane protein whose extracellular domain contains four sequences homologous to calcium-dependent binding motifs of the cadherin superfamily. In addition, a 31-amino acid threonine, serine, and proline (TSP)-rich tandemly repeated motif similar to mucin domains is contained within its longer isoform. We have termed this novel cell adhesion molecule μ-protocadherin to reflect the mucin-type repeat contained within its longer isoform. Transfection of μ-protocadherin in L929 cells mediated aggregation, confirming its role in intercellular adhesion. Thus, μ-protocadherin

* This work was funded by National Institutes of Health Grant DK 55388-01, a March of Dimes award, and Spunk Fund, Inc (to M. G.). The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked "advertisement" in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

The nucleotide sequence(s) reported in this paper has been submitted to the GenBank™/EBI Data Bank with accession number(s) AF221952.

‡ To whom correspondence should be addressed: Dept. of Pediatrics, College of Physicians and Surgeons of Columbia University, 3959 Broadway BHS Rm. 724, New York, NY 10032. Tel.: 212-305-1559; Fax: 212-305-3475; E-mail: mg81@columbia.edu.

1 The abbreviations used are: mAb, monoclonal antibody; PCR, polymerase chain reaction; SANIII, neuraminidase III; ELISA, enzyme-linked immunosorbent assay; PAGE, polyacrylamide gel electrophoresis; HPLC, high performance liquid chromatography; RT, reverse transcriptase.
represents a new receptor capable of mediating signals during branching morphogenesis.

**EXPERIMENTAL PROCEDURES**

**Materials—**Kidneys were obtained from Pel-Freez Biologicals (Rogers, AR). U. S. A. standard testing sieves and Teflon polymerase were obtained from Fisher. Deoxynucleotide primers for polymerase chain reactions (PCR) were synthesized by Gene Link (Thornden, NY). Hi-Trap-SP, HiTrap protein A, Mono-Q, and PD10 columns were obtained from Amersham Pharmacia Biotech, phenylmethylsulfonyl fluoride, leupeptin, pepstatin A, aprotinin, piperase, Tris base, glycerol, sodium chloride, and Triton X-100 were from Sigma. RNA-zol B was purchased from Tel-Test Inc. (Friendswood, TX). Purification of mRNA utilized Oligotex beads from Qiagen (Valencia, CA). Ag11 library was obtained from CLONTECH (Palo Alto, CA). O-Glycosidase, neuraminidase III (NANIII), and N-glycanase F were purchased from Glyko Inc. (Novato Ca.).

**Purification of gp100—**Rat kidneys were harvested, snap-frozen, and shipped the following day (Pel-Freeze). One hundred and fifty thawed kidneys were finely macerated by a razor and passed first through a 2-mm opening and then a 1.1-mm opening sieve. They were then resuspended in a total volume of 425 ml of homogenization buffer that contained 10 mM Tris-Phosphate-Tris buffer, pH 7.5, and 300 mM sucrose. All buffers contained the protease inhibitors phenylmethylsulfonyl fluoride (20 μg/ml), leupeptin (20 μg/ml), pepstatin A (1 μg/ml), and aprotinin (50 μg/ml). The homogenate was homogenized by 20 strokes with a motor-driven Teflon pestle in a glass homogenizer at 1200 rpm. The homogenate was spun at 1,800 g for 16 min, and the resultant supernatant was spun at 16,800 × g for 20 min. The pellet consisted of two layers, a lower brown layer and a fluffy white upper layer, which was separated by gentle swirling five times. The fluffy white layer was spun at 30,000 × g for 2 h to obtain the membrane fraction and resuspended in 250 mM sucrose, 5 mM Tris, pH 8.0, 1 mM EDTA, and 10% glycerol and frozen.

The fraction was then diluted with 10 mM Tris, pH 7.4, and placed onto a 19–29% sucrose step gradient containing 10% glycerol in a Hypes/Tris pH 7.5 buffer and spun in a SW-28 rotor for 3 h. The 19–29% interface was collected and membrane vesicles were pelleted at 100,000 × g for 60 min. The pellet was resuspended in a 100 mM sodium carbonate/bicarbonate buffer, pH 11, and homogenized 5 times in a 100–mL tight-fitting glass homogenizer. After a 100,000 × g spin, the pellet was resuspended and homogenized 3 times in 10 mM sodium phosphate, pH 6.8, and selectively solubilized with Triton X-100 using a 1:1 detergent/protein ratio and a final Triton concentration of 0.4%. After 100,000 × g spin, the supernatant was discarded. The precipitation exchange column equilibrated at pH 4.75 in a 50 mM sodium acetate buffer containing 10% glycerol and 0.4% Triton. The flow-through fractions contained the majority of gp100, as measured by ELISA. These fractions were concentrated using a Centriprep-10 and exchanged over a PD-10 column with 19 mM piperase buffer, pH 5.75, containing 0.3% Triton X-100. A precipitation exchange column (Mono-Q) at pH 5.75 was performed with the eluents monitored for gp100 by ELISA. gp100 eluted in a peak between 255 and 298 mM sodium chloride concentration. Tris buffer, pH 7.4, was added immediately to each fraction as it eluted from the column. These fractions were pooled on a 5–25% sucrose gradient containing 20 mM Tris, pH 7.4, 0.3% Triton x-100, 500 mM salt, and 10% glycerol. After overnight centrifugation, fractions of the sucrose gradient were tested by ELISA for gp100. The final purified fraction was further separated by SDS-PAGE under non-reducing conditions. A Western blot was performed, and an adjacent lane containing the majority of the purified sample was used for in-gel proteolytic digestion (11) and peptide sequencing.

**In-gel Digestion of Proteins and Sequence Analysis—**The SDS-PAGE gel was first stained in 0.05% Coomassie Brilliant Blue-G dissolved in 0.5% acetic acid and 20% methanol for 1 h. After destaining in 30% methanol, the appropriate band was excised in a displaced volume of less than 100 μl. The protein was reduced by submerging the gel in 100 μl of 0.1 M dithiothreitol, 0.1 M Tris, pH 8.5, and shaking gently for 1 h at 55 °C. The liquid was then removed from the tube and replaced with 100 μl of 0.015 M N-iodopropiolecitoc acid, 0.1 M Tris, pH 8.5. After reacting for 30 min in the dark, the supernatant containing the alkylating reagent was removed and discarded, and the gel was washed with 500 μl of 0.05 M Tris, pH 8.5, 50% acetonitrile for 20 min with shaking. The supernatant was discarded, and the was washed repeated 2×. The washed gel pieces were dried completely in a Speed-Vac concentrator. To the tube containing the dried gel pieces, 40 μl of digestion buffer (0.025 M Tris, pH 8.5) containing 0.05 μg trypsin was added, and the tube was incubated for 20 h at 37 °C. When digestion was complete, the gel pieces were extracted three times with 100 μl of 50% acetonitrile, 0.1% trifluoroacetic acid, shaking for 30 min each time, and the supernatants were transferred to a Hewlett-Packard HPLC injection vial. The combined washes in the injection vial were dried in a Speed-Vac concentrator and redissolved in 200 μl of 0.1% trifluoroacetic acid for injection onto an HPLC column. 7.5% of the total gel was on a spectrophotometry analysis. Matrix-assisted laser desorption ionization mass spectroscopic analysis was performed on the digest and on 10% of selected fractions using a PerSeptive Voyager PE-RI mass spectrometer. Sequence analysis was performed on a Perkin-Elmer Bioskin mass spectrometer. Samples were then analyzed by SDS-PAGE and immunoblotting with the 3D2-E9 mAb or rabbit polyclonal antibody generated against the carboxyterminal tail of gp100.

**Generation of Antibodies—**The generation of the 3D2-E9 mAb has been previously described (1). Rabbit polyclonal antibodies were generated against the extracellular domain of μ-protocadherin. A sequence encoding amino acids 126–412 was ligated into pGEX (Amersham Pharmacia Biotech) to create a glutathione S-transferase fusion protein and transformed into Bl21 Eschericia coli cells. The fusion protein was purified on an immobilized glutathione S-transferase column (Phire), and the amino acid sequence was verified by mass spectroscopy following trypsin digestion. Soluble fusion protein was injected into rabbits for generation of polyclonal antisera (Pocono Rabbit Farm and Laboratory, Inc., Canadensis, PA). Monoclonal antibody to β-catenin was purchased from Transduction Laboratories (San Diego, CA).

**Cloning of cDNA for gp100—**Total RNA was isolated from adult rat kidney with RNA-zol B, and mRNA was purified using Oligoex beads following the manufacturer's guidelines. mRNA was reverse-transcribed using a “hybrid” primer that consists of 17 nucleotides of oligo(dT) followed by a unique 35-base oligonucleotide sequence cctagtgcagggaggaggagtgcaggtcatagccactcggctgtcag (“anchor” PCR) (12). Amplification was then performed using a primer containing part of the 3’ anchor sequence and bound to its 3’ end, and an additional primer derived from the peptide sequences. A second set of amplification cycles using “nested” primers was done to quench the amplification of nonspecific products. Using this strategy, we obtained an 850-base pair PCR product representing the putative 3’ end of the gene. An additional 750 base pairs were obtained by degenerate PCR using the “touchdown” PCR method (13). The 850- and 750-base pair DNA product was sequenced. Analysis and cloning of cDNA sequences not used in the construction of the primers, demonstrating that these PCR products coded for gp100. The PCR products were then used to screen an adult rat kidney Ag11 library constructed by priming with both oligo(dT) and random hexamers. Hybridization with a digoxigenin-labeled DNA probe and detection was performed following the manufacturer's directions (Roche Molecular Biochemicals). Six clones including three full-length clones were sequenced and identified. A consensus sequence was constructed from a minimum of three independent sequence reactions and was placed in GenBank. Scanning the GenBank data base revealed a similar mouse EST clone (m66e11.11). All four of the phage spanning the extracellular region of μ-protocadherin closest to the transmembrane domain contained a 179-base pair region that was not present within three independent RT-PCR reactions from anchor and touchdown PCR. That this region was alternatively spliced was confirmed by the presence of two RT-PCR products derived from embryonic day 19 kidney generated from a pair of primers flanking this area: 5’-cctagcgtgttgctac-3’; 5’-gggtgtcatgactcttg-3’). Interestingly, the longer isoform contains a 31-amino acid region that is repeated 21 times. The predicted molecular mass based on this sequence was 91 kDa, consistent with the observed molecular mass by Western blotting (9). Amino acid sequences derived from the nucleotide consensus sequence are shown in Fig. 1.

**Northern Analysis of gp100—**A whole mouse embryo mRNA blot obtained from CLONTECH (Palo Alto, CA) was used to study the developmental expression of μ-protocadherin. This was hybridized with a digoxigenin-labeled homologous mouse probe derived from the cyto
plasmic region of a mouse EST clone (mbl68e11.r1). A single rat tissue Northern blot containing mRNA from heart, brain, spleen, lung, liver, skeletal muscle kidney, and testis (CLONTECH) was hybridized with a 420-base pair rat probe that spanned the transmembrane region and two of the mucin-type repeats of \( \mu \)-protocadherin. The probe was digoxigenin-labeled by a PCR reaction and hybridized with ExpressHyb as per the manufacturer’s protocol (CLONTECH).

**Immunocytochemistry**—Isolation of kidneys, fixation, cryostat sectioning, and staining protocol were performed as described previously (9) with some minor modifications. Tissue sections (E18) were blocked with 5% albumin in PBS before the addition of primary antibody. Immunocytochemistry was performed using a monoclonal antibody to \( \beta \)-catenin and a rhodamine-conjugated secondary followed by staining with a directly conjugated fluorescein isothiocyanate-labeled \( \mu \)-protocadherin in transfected L cells was assayed by immunofluorescence with rabbit polyclonal antibodies generated against its extracellular domain. Rabbit polyclonal antibody from a high titer serum was purified on a HiTrap protein A column and used at 6–10 \( \mu \)g/ml.

For immuno-electron microscopy, embryonic day 14 rat lungs were isolated and stained with the 3D-E9 mAb and a 10-nm size gold-conjugated goat anti-mouse secondary at 4 °C before embedding. They were then fixed with 4.25% glutaraldehyde in 0.1 \( \text{M} \) sodium phosphate buffer. Tissues were embedded in Epon, and sections were examined under the microscope. The area boxed seen in Fig. 1E is enlarged and reveals a cluster of gold particles within a coated pit.

**Isolation of Transfectants—** \( \mu \)-Protocadherin cDNA was cloned in frame with both the 5’ and 3’ epite tags of the pHM6 expression vector (Roche Molecular Biochemicals). L929 cells were transfected with 2 \( \mu \)g of pHM6-\( \mu \)-protocadherin DNA/35-mm well using LipofectAMINE (Life Technologies, Inc.). Selection with G418 was initiated 48 h after transformation. Resistant cell clones were isolated 10–12 days after plating using cloning rings (Fisher). They were tested for expression of \( \mu \)-protocadherin by immunofluorescence staining with rabbit polyclonal antibodies generated against its extracellular domain.

**Aggregation Assay**—The classical aggregation assay developed for the cadherins was utilized (14). A key feature is the presence of calcium during the treatment to preserve function of cadherin-like proteins. Parental L929 cells and transfected cell clones in late logarithmic growth were used for experiments. Cells were washed twice with Hanks’ balanced salt solution containing 1 mM CaCl\(_2\) and 1 mM magnesium sulfate. 0.025% trypsin was then added for 50 min at 37 °C in the presence of 5% CO\(_2\). Cells were then collected and washed twice with soybean trypsin inhibitor and resuspended at 1.0 \( \times \)10\(^6\) cells/ml in Hanks’ balanced salt solution containing either 1 mM calcium chloride or 1 mM EDTA. 1.5 ml of this suspension was placed into a 12.5-mm diameter well coated with 1.0 ml of 1.5% agarose in Hanks’ balanced salt solution and 0.45 g/100 ml glucose. Cells were incubated at 37 °C in a rotary shaker at 80 rpm for 20 and 60 min. Three individual wells per experimental condition were counted manually in a hemacytometer. Results in the figure are the average cumulative data obtained from four independent experiments. To analyze the data, the proportion (%) of aggregated cells in each individual well was first transformed by the arcsine transform (15). Then, a factorial analysis of variance was applied, with the presence and absence of \( \mu \)-protocadherin in the presence of calcium or EDTA as the main effect variables. P values were determined using the above analysis.

**RESULTS**

**Purification of gp100**—To follow the enrichment of the 3D2-E9 antigen, we used an ELISA and/or Western blotting assay. Immunoblots under reducing conditions revealed a 100-kDa band, and pending its identification, we termed this protein gp100. By immunofluorescence, gp100 was present in adult rat kidneys in a vesicular staining pattern in the S3 segment of the proximal convoluted tubule (9). Hence, we first purified microsomal vesicles that were enriched in gp100 using sucrose gradient centrifugation. The purification scheme is outlined in Table I. Because gp100 remained in the pellet following extraction of the vesicles with 100 mM sodium carbonate (at pH 11), we concluded that it was an integral membrane protein. Selective solubilization conditions in Triton X-100 yielded a 2-fold purification. Preliminary experiments showed that gp100 was an acidic protein; hence, the solubilized material was adsorbed onto a cation exchange column at pH 4.75. The flow-through fractions containing the majority of gp100 were then applied to an anion exchange column (Mono-Q) at pH 5.75. Gp100 eluted in a peak between 255–298 mM sodium chloride, and these fractions were pooled, concentrated, and placed on a 5–25% sucrose equilibrium density gradient in the presence of Triton X-100 and 500 mM sodium chloride to minimize protein-protein interactions. The final purified fraction was further separated by SDS-PAGE under non-reducing conditions. In the unreduced state, a 1200-kDa form was recognized by the 3D-E9 mAb by Western blotting, and an adjacent lane containing the majority of the purified sample was submitted for sequencing. Proteolysis of gp100 and subsequent partial sequence analysis of the peptides (underlined in Fig. 1) yielded six sequences from which degenerate oligonucleotides were constructed and used in a PCR reaction. We also obtained a seventh peptide, which we have not been able to find in the full-length cDNA and may have been a contaminant.

**Molecular Cloning**—Degenerate primers were constructed using two of the peptide sequences underlined in Fig. 1. The RT-PCR products obtained contained a third independently sequenced proteolytic fragment derived from gp100, showing conclusively that the cDNA indeed encodes gp100. PCR products were then used to screen an adult rat Agr11 kidney library, and three full-length clones of gp100 were obtained and sequenced. Two additional transcripts, a long and a short isoform, were isolated by RT-PCR using RNA from adult kidney as the template. However, the three Agr11 clones sequenced encoded the long isoform. GP100 is a novel protein whose amino acid sequence is presented in Fig. 1, and a schematic of its functional domains is presented in Fig. 2. The deduced amino acid sequence gives a minimum expected molecular mass of 91 kDa for the long isoform and 70 kDa for the shorter isoform. A 24-amino acid hydrophobic region immediately following the initiating methionine encodes the signal peptide. Both isoforms were sequenced from a RT-PCR reaction from embryonic day 19 rat kidney RNA with primers that flanked the putative spliced region. Analysis using BLAST Dbest (non-redundant data base of GenBank\(^a\) + EMBL + DDBJ EST divisions) confirmed the presence of these two isoforms over this region. Both isoforms are integral membrane proteins with a transmembrane region containing a 23-amino acid hydrophobic region (amino acid 670–692) as determined by a Kyte-Doolittle hydrophobicity plot (Fig. 1). In the extracellular domain, this glycoprotein contains four copies of a calcium binding fold homologous to that present in the cadherin superfamily (16). The cytoplasmic domain of gp100, however, bears no homology to the classical cadherins. The first three domains (EC1-EC3) were detected by Prosite Profile search (17). The fourth domain was found by BLAST analysis. Fig. 3 compares the four cadherin ectodomains using Clustal W alignment and presents their consensus sequence (18). The most closely related extracellular cadherin-like domains to those of gp100 based on BLAST analysis are from a tumor suppressor in Drosophila called fat, a member of the protocadherin family.

**TABLE I**

| Purification of \( \mu \)-protocadherin | Fold purification | % total activity |
|----------------------------------------|------------------|----------------|
| Kidney homogenate                      | 100              |                |
| Microsomal membrane                    | 10               | 41.0           |
| Sucrose gradient microsomes            | 40               | 16.8           |
| pH 11 extraction                       | 62               | 14.3           |
| Triton solubilization                  | 111              | 12.9           |
| Cation exchange                        | 223              | 8.9            |
| Anion exchange                         | 716              | 5.2            |
| Equilibrium density gradient           | 22,903           | 2.6            |

---

\(^a\) GenBank: The GenBank database is a repository of nucleotide and protein sequences.
Interestingly, gp100 contains an RGD peptide located in its first extracellular cadherin similar to the first repeat of the human protocadherin-\(\alpha\) family (20). Amino acids 464–655 were absent in the short isoform. This region contains a 3.5 tandem repeated 31-amino acid sequence rich in serine, threonine, and proline residues, characteristic of a mucin-type repeat structure (highlighted in yellow in Fig. 1; amino acids 545–575, 576–606, 607–636, and 637–648). A comparison of the mucin-like repeat region with the C-terminal TSP region of rat intestinal mucin peptide (rMuc2) (21) is presented in Fig. 4. Sixty-nine percent of the amino acids are serine, threonine, or prolines, providing for many sites of O-linked glycosylation. We have termed this novel cell adhesion molecule \(\mu\)-protocadherin to reflect the mucin-type repeat contained within its longer isoform.

The cytoplasmic region of both isoforms terminate with NTYY, a sequence that conforms to a type I PDZ binding motif X-(Thr/Ser)-X-Val-COOH (22). The mammalian fat protocadherin, but not the Drosophila homologue, also contains this motif (23). In Caenorhabditis elegans, this motif, by binding to a PDZ domain, is responsible for preferential localization of the epidermal growth factor receptor to the basolateral membrane (24). It is interesting to note that \(\mu\)-protocadherin is targeted to the basolateral membrane of epithelial structures during kidney and lung development (see Fig. 9).

In addition, both isoforms contains four proline-rich regions (PXXP) in the cytoplasmic region predicted to interact with SH3 domains, suggesting that this protein is a receptor capable of transduction of extracellular signals (Fig. 1). Six PXXP regions are present in the cytoplasmic region of the protocadherin-\(\alpha\) gene family (20).

**Carbohydrate Analysis of \(\mu\)-Protocadherin**—Sequence analysis of \(\mu\)-protocadherin reveals five potential asparagine-linked glycosylation sites (amino acids 135, 173, 201, 311, and 408), which are noted by asterisks in Fig. 1. Treatment of partially purified adult rat kidney vesicles with \(N\)-glycanase F (a recombinant enzyme that releases asparagine-linked high mannose and hybrid and complex oligosaccharides) inhibited recognition of \(\mu\)-protocadherin by the 3D2-E9 mAb, revealing that its epitope was a carbohydrate (Fig. 5A). When the same blot was stripped and re-probed with a polyclonal antibody generated against the extracellular domain, the molecular mass of \(\mu\)-protocadherin was reduced (Fig. 5B). Thus, the findings that the 3D2-E9 disrupts branching morphogenesis and that its epitope is a carbohydrate suggest that the carbohydrate structure of \(\mu\)-protocadherin is important in its function.

Biochemical analysis of O-linked structures is more complex due to the many substitutions that may occur such as with sialic acid and fucose. Treatment of adult rat kidney vesicles with NANIII generated a new band migrating at 66 kDa, demonstrating that \(\mu\)-protocadherin is heavily sialylated (Fig. 5C). A second band with decreased mobility at approxi-
mately 105 kDa may be due to a reduction in negative charges and/or change in conformation. In addition, treatment with NANIII and O-glycosidase resulted in a decrease in intensity in the 97-kDa band after immunoblotting, suggesting that m-protocadherin is also O-glycosylated similar to the mucins.

Developmental Regulation of the m-Protocadherin Gene—m-Protocadherin is expressed in a complex manner during embryonic development. At day 7 (E7) of mouse embryonic life, both transcripts of gp100 were seen in total embryo RNA (Fig. 6), but its expression was down-regulated at E11 only to reappear at E15–17. In addition, as gestation nears term, independent regulation of each isoform was noted. An RT-PCR reaction with primers that span the repeat area is shown in Fig. 7. In E19 of gestation, both maternal (E19M) and the embryonic rat kidneys (E19) express almost exclusively the longer isoform. To account for possible differences in cDNA template, the PCR reaction was performed with more than a 100-fold increase in template without a change in results (data not shown). This expression pattern dramatically changes as one nears term. In the post-partum female (Fig. 7, birthing mother (BM)), as well as in the neonate, the expression of the shorter isoform increases. Note that Fig. 6 was obtained using whole embryo mRNA rather than kidney-specific mRNA; hence, each isoform might be derived from a different organ.

Thus, the expression of the two isoforms of m-protocadherin in the embryonic kidney is independently regulated during the perinatal period, suggesting they have distinct functions.

Northern analysis of a multiple tissue cDNA blot in the adult reveals that expression of gp100 is kidney-specific (Fig. 8). PCR reaction detected the two transcripts, however, in the adult lung and small intestine (data not shown).

Localization of m-Protocadherin during Development—Immunocytochemistry with the 3D2-E9 antibody showed that m-protocadherin was localized in E18 embryonic kidney at points of cell to cell and cell to matrix contacts (arrowhead in Fig. 9) and diffusely stained the mesenchyme. Green staining reflects m-protocadherin expression, whereas red staining is β-catenin. Some colocalization with β-catenin is noted at the junction of the basal and lateral membrane (arrows in figure). More detailed analysis was shown by immunogold electron
 Isoform Specific Expression

![Image](image_url)

**Fig. 7.** Independent regulation of α-protocadherin isoforms. RT-PCR using kidney-derived mRNA from embryonic day 19, maternal day 19 of gestation, and birthing mother (BM) is presented. The primers used (5'-ctctaactgctgctc-3'; 5'-ggttgctagacgctgg-3') span the mucin-like repeat area. bp, base pairs.

**Fig. 8.** Expression of α-protocadherin in adult rat tissues. A multiple adult rat tissue Northern blot containing mRNA from heart, brain, spleen, lung, liver, skeletal muscle kidney, and testis (CLONTECH) was hybridized with a 420-base pair rat probe that spanned the transmembrane region and two of the mucin type repeats of α-protocadherin. Two micrograms of poly(A)-rich mRNA was loaded/lane. kb, kilobases.

**Fig. 9.** Immunocytochemistry of α-protocadherin in E18 kidney. Immunocytochemistry was performed using a monoclonal antibody to β-catenin and a rhodamine-conjugated secondary and a directly conjugated fluorescein isothiocyanate-labeled 3D2-E9 mAb. Expression of α-protocadherin is noted at points of cell to matrix contacts (arrows). Colocalization with β-catenin is noted at the junction of the lateral and basal membrane (arrows).

**Fig. 10.** Immuno-electron microscopy. Immunogold electron micrographs of embryonic day 14 rat lungs after staining with the 3D2-E9 mAb and a 10 nm size gold-conjugated goat anti-mouse secondary. The area boxed is enlarged and reveals a cluster of gold particles within a coated pit.

μ-Protocadherin, a Developmentally Regulated Adhesion Molecule

micrographs. μ-Protocadherin is concentrated in the basal surface in coated pits of E14 lung (Fig. 10). These studies suggest that μ-protocadherin likely mediates adhesive interactions to extracellular matrix or to molecules present in the lateral membrane. We had previously shown that addition of 3D2-E9 to embryonic kidney induced apoptosis, and it is now well

known that many epithelial cells receive survival signals from the extracellular matrix (25).

**μ-Protocadherin Mediates Cell Aggregation**—To examine the function of μ-protocadherin, we cloned its full-length cDNA into a vector that also contains epitope tags at both the 5’ (hemagglutinin) and 3’ (His6) ends. As expected, the 5’ tag is clipped off with the processing of the signal sequence of μ-protocadherin. L cells transfected with pHM6-μ-protocadherin were selected for stable transformants after 48 h. Immunofluorescence revealed that μ-protocadherin is expressed at points of cell to cell contacts in the stable transformants (Fig. 11A). Using the classical aggregation assay developed for the cadherins (14), we allowed parental L929 cells and transfected cell clones to aggregate at 37 °C while shaking at 80 rpm. Three individual wells per experimental condition were manually counted with a hemacytometer at 20 and 60 min and averaged. The cumulative results of four independent experiments are shown in Fig. 11B, which demonstrates that there is a highly significant aggregation induced by μ-protocadherin at 20 and 60 min (p < 0.0001). Although the aggregation was calcium-dependent (p < 0.002), μ-protocadherin induced some aggregation even in the absence of calcium. Aggregation was not inhibited by the addition of rabbit polyclonal antibody directed against μ-protocadherin (data not shown), which may reflect that the polyclonal antibody is not inhibitory or that additional molecules contribute in part to the adhesive interaction. To confirm a more direct role of μ-protocadherin in aggregation, transfectants were generated in which the cytoplasmic region was deleted, but the extracellular domain remained intact. In preliminary experiments, they were unable to mediate aggregation, highlighting that the cytoplasmic region is critical for aggregation (data not shown).

**DISCUSSION**

The cadherin superfamily contains several members that include the original family, now termed classical cadherins, that are defined by their calcium-dependent homophilic cell adhesion properties. Their cytoplasmic region binds β-catenin and is linked to the actin cytoskeleton through α-catenin and actinin (7). A related family of cadherin-like molecules, the protocadherins, is expressed in the brain and are implicated in synaptic junction adhesive interactions (20, 26). It has been proposed that a wide variety of such proteins could be generated by alternate splicing and recombination, since the genomic sequences of protocadherins are organized in three closely linked clusters that contain variable extracellular domains but a conserved cytoplasmic region (20). Analysis of
μ-Protocadherin sequence reveals several structural features present in the protocadherin superfamily. μ-Protocadherin contains in its first cadherin ectodomain an RGD peptide similar to protocadherins-α. Furthermore, the first repeat does not contain an HAV sequence felt to be responsible for the homophilic binding properties of the classical cadherins. However, both LI-cadherin and the protocadherins can mediate homotypic interactions without the HAV domain. The μ-protocadherin cytoplasmic region contains four putative SH3 binding domains, whereas that of protocadherin-α contains six. The cytoplasmic region of μ-protocadherin terminates with a type I PDZ binding domain, similar to the C-terminal end of rat fat protocadherin (23). In C. elegans, this motif at the C terminus of the epidermal growth factor receptor homologue is responsible for preferential localization to the basolateral membrane (24). Thus, this sequence may be responsible for the targeting of μ-protocadherin to the basolateral membrane during kidney development.

Genomic analysis of the protocadherin family members revealed that the extracellular cadherin repeats were encoded by a single exon (20). Preliminary genomic analysis, however, revealed that μ-protocadherin contained several exons coding for its extracellular region. A unique feature of μ-protocadherin is that it contains TSP-rich mucin-like repeats in its extracellular domain (27). Recent studies have demonstrated that high levels of mucin production were associated with the decreased adhesive properties of carcinoma cells (28). Furthermore, transfection of MUC1 into certain cells interferes with cadherin-mediated cellular aggregation (29). The length of the extracellular MUC1 repeats was found to be the dominant factor determining the extent of inhibition (22). Alternatively, mucins may compete for cytosolic regulators, since cell adhesive interactions are a necessary step for Muc1 to bind to β-catenin (30). These studies raise the possibility that the adhesive properties of μ-protocadherin may be regulated by the presence of the mucin repeats.

The longer isoform of μ-protocadherin containing the mucin repeat is differentially expressed during the perinatal period. This highlights the probable different function of each isoform. An instructive comparison can be made to another adhesive molecule, the mucosal addressin cell adhesion molecule-1 (MAdCAM-1), which contains immunoglobulin (Ig)-like as well as mucin domains. Two isoforms were described; in one, the mucin domains were alternatively spliced out (31). MAdCAM-1 purified from mesenteric lymph nodes can support the rolling of lymphocytes under shear through its binding to L-selectin (32). Hence, the selectin binding carbohydrate determinants of the mucin domains were postulated to play an important role in allowing movement and plasticity of cells while still permitting adhesion to occur. An analogous situation may be the requirement of epithelial cells to migrate during branching morphogenesis, yet still remain adherent to one another. This might explain how the 3D2-E9 mAb, which binds to a carbohydrate determinant, can disrupt branching morphogenesis.

Dynamic changes in cell adhesion participate in the generation of three-dimensional structures during development. These adhesion mechanisms are highly regulated and are integrated with the processes of changes in cell shape and cell migration. Differential expression of the cadherins can drive morphogenesis by stimulating aggregation of subsets of cells expressing the same type of cadherin. During nephrogenesis, for example, cadherin-11 is expressed primarily in mesenchymal cells, whereas the newly formed epithelium of the renal vesicle expresses E-cadherin near the ureteric bud tips and cadherin-6 elsewhere. As the S-shaped tubules develop, proximal tubule progenitors express cadherin-6, whereas distal tubule cells express E-cadherin and glomeruli express P-cadherin (8). It would therefore be important to examine whether the expression of the two forms of μ-protocadherin change during kidney and lung development.

In addition to responding to signaling events, cell adhesion molecules themselves transduce signals into the cell. Adding to the complexity, these adhesive interactions often must be in a dynamic state to achieve the movements of cells with respect to the matrix and to one another. Cell adhesion molecules such as the cadherins also participate in signaling pathways that control developmental patterning in embryos (for review see Ref. 6). As an example, β-catenin is an essential component of the WNT signaling pathway. In Xenopus embryos, β-catenin mediates a signaling pathway that controls dorsal-ventral and anterior posterior body axes patterning. μ-Protocadherin was shown to colocalize by immunofluorescence to β-catenin, primarily at the junction of the basal and lateral membrane, although isoform-specific expression may be more revealing.

That modulation of μ-protocadherin signaling by the 3D2-E9 antibody led to apoptosis, and disrupted tubulogenesis in both kidney and lung suggests that this protein is critically involved in these processes. Disruption of cell to cell and cell to matrix interactions induces apoptosis (25). In human bronchiolar epithelial cells grown in vitro, an antibody to the αv integrin induced apoptosis (33). Interestingly, apoptosis was also prevented by allowing the epithelial cells to aggregate with each other. Thus, the apoptotic pathway may regulate three-dimensional structure formation by ensuring that aberrant structural and functional units are removed (1, 4). μ-Protocadherin
could mediate such morphometric information into the apoptotic pathway.

Acknowledgments—We are grateful to Mary Ann Gawinowicz of the protein core facility of Columbia University for performing the proteinase and peptide sequencing of m-protocadherin. The advice of Richard Friedman in the sequence analysis of m-protocadherin is appreciated. We also thank Moshe Levinson for help in statistical analysis of the data. Finally, we are grateful to Zvi Eckstein for isolating the m-protocadherin transfectants and to Jonathan Barasch for help in electron microscopy.

REFERENCES

1. Koseki, C., Herzlinger, D., and Al-Awqati, Q. (1992) J. Cell Biol. 119, 1327–1333
2. Durbec, P., Marcos-Gutierrez, C. V., Kikenny, C., Grigoriou, M., Wartiowaara, K., Suvanto, P., Smith, D., Ponder, B., Costantini, F., Saarma, M., Sariola, H., and Pachnis, V. (1996) Nature 381, 789–793
3. Herzlinger, D., Koseki, C., Mikawa, T., and Al-Awqati, Q. (1992) Development 114, 565–572
4. Coles, S. R., Burne, J. F., and Raff, M. C. (1993) Development 118, 777–784
5. Goldberg, M. R., and Peshkovsky, C. (1998) Am. J. Respir. Crit. Care Med. 157, 12492–12494
6. Leung, E., Greene, J., Ni, J., Raymond, L. G., Lehnert, K., Langley, R., and Kristiansen, G. W. (1996) ImmunoL Cell Biol. 74, 490–496
7. Berg, E. L., McEvoy, L. M., Berlin, C., Bargatze, R. F., and Butcher, E. C. (1993) Nature 366, 695–698
8. Aoshiba, K., Renard, S. I., and Spurzem, J. R. (1997) Am. J. Physiol. 272, L28–37