Experimental Cell Research **248,** 97–109 (1999)
Article ID excr.1999.4392, available online at http://www.idealibrary.com on IDEAL®

Molecular Characterization of the Tight Junction Protein ZO-1 in MDCK Cells

L. González-Mariscal,*[,1] S. Islas,* R. G. Contreras,* M. R. García-Villegas,* A. Betanzos,* J. Vega,† A. Diaz-Quiñónez,‡ N. Martín-Orozco,‡ V. Ortiz-Navarrete,‡ M. Cereijido,* and J. Valdés*

**Department of Physiology, Biophysics, and Neuroscience and †Department of Genetics and Molecular Biology, Center for Research and Advanced Studies (CINVESTAV), Apdo. Postal 14-740, 07000 México, D.F., México; and ‡Unidad de Investigación Médica en Inmunoquímica, Centro Médico Nacional Siglo XX1, IMSS, México, D.F., México*

Most of the information on the structure and function of the tight junction (TJ) has been obtained in MDCK cells. Accordingly, we have sequenced ZO-1 in this cell type, because this protein is involved in the response of the TJ to changes in Ca^{2+}, phosphorylation, and the cytoskeleton. *ZO-1* of MDCK cells comprises 6805 bp with a predicted open reading frame of 1769 amino acids. This sequence is 92 and 87% homologous to human and mouse *ZO-1*, respectively. Two nuclear sorting signals located at the PDZ1 and GK domains and 17 SH3 putative binding sites at the proline-rich domain were detected. We found two new splicing regions at the proline-rich region: β had not been reported in human and mouse counterparts, and γ, which was previously sequenced in human and mouse *ZO-1*, is now identified as a splicing region. The expression of different β and γ isoforms varies according to the tissue tested. With the information provided by the sequence, Southern blot, and PCR experiments we can predict a single genomic copy of *MDCK-ZO-1* that is at least 13.16 kb long. *MDCK-ZO-1* mRNA is 7.4 kb long. Its expression is regulated by calcium, while the expression of MDCK-ZO-1 protein is not.

Key Words: **tight junctions; ZO-1; MAGUK; MDCK; epithelia.**

INTRODUCTION

In the past century, the tight junction (TJ) was regarded as a mere seal at the outermost end of the space between epithelial cells [10]. The scope has since drastically changed, because the paracellular permeation route is by no means impermeable, as different epithelia may present a wide range of transepithelial resistance: as low as 8 $\Omega \cdot cm^2$ in the proximal tubule of the kidney to 10,000 in the urinary bladder. Furthermore, TJ structure and properties can be changed according to physiological and pathological conditions and through challenge with hormones and pharmaceutical agents [30]. The response of TJs to such a variety of conditions is in keeping with the fact that it is composed of a large number of molecular species, which can themselves vary as a consequence of phosphorylation, the presence of Ca^{2+}, cell–cell attachment, etc. [5, 6, 12, 13, 28, 31, 35]. Three of these molecules (ZO-1, ZO-2, and ZO-3) are members of the membrane-associated guanylate kinase homologue protein family (MAGUK) [3, 17]. MAGUK proteins are composed of modular units including: (a) PDZ domains, named after the proteins in which they were first identified (PSD95, DLG, and ZO-1), consisting of an approximately 90-amino-acid domain, repeating a glycine-leucine-glycine-phenylalanine signature sequence. These domains seem to bind homo- or heterotypically to neighboring proteins and promote their clustering. (b) An SH3 region (Src homology 3) that has been described on several cytoskeletal and signaling proteins, where it mediates protein–protein binding to proline-rich regions. (c) A guanylate kinase (GK) domain whose function remains enigmatic, since neither ZO-1, ZO-2, nor ZO-3 can express this enzymatic activity because they lack specific residues required for ATP and GMP binding. (d) In addition to those domains, ZO-1 and ZO-2 present a proline-rich moiety toward the 3′ end. In ZO-3 [17] this region is instead intercalated between PDZ2 and PDZ3.

In addition to their participation in tight and septate junction structure and function, MAGUK proteins are known to be involved in biological processes as important as tumor suppression, protein clustering of ion channels and receptors, differentiation, regulation of gene expression, and binding to cytoskeleton [3]. This indicates that junctional proteins may be involved in a much larger network of biological functions than regulation of the paracellular permeability.

The understanding of the structure and function of

Sequence data from this article have been deposited with the GenBank Data Library under Accession No. U55935.

[1] To whom reprint requests should be addressed. Fax: (525) 7-47-71-05. E-mail: lorenza@fisio.cinvestav.mx.

0014-4827/99 $30.00

the TJ was greatly promoted by the use of model systems prepared with epithelial cell lines, such as MDCK [4, 9, 13, 14]. In fact, most of the information available on the synthesis and assembly of the TJ was obtained through the use of these model systems and, therefore, to pursue the study of the structure and function of the TJ, the information of the sequence of *ZO-1* in MDCK cells is needed. In the present work we sequence and analyze the *ZO-1* gene and its expression in MDCK cells (*MDCK-ZO-1*). In addition to an α domain previously described in human ZO-1 [5], we find two new regions of alternative splicing in the proline-rich domain (β and γ) not previously described for human and mouse *ZO-1* (*h-* and *m-ZO-1*) and determine the size of intronic regions surrounding them. Sequence analysis revealed nuclear signals existing in tight and septate junction-associated proteins, as well as phylogenetic relationships of ZO-1 to synaptic and other proteins of this superfamily.

MATERIALS AND METHODS

Cell culture. MDCK cells were either grown in DMEM with or without calcium or grown in DMEM and subjected to a calcium switch as previously described [14].

RNA isolation and expression library construction. Total RNA from MDCK cells grown for 3 days in low-calcium medium was isolated as described by Chomczynski and Sacchi [11]. Intact mRNA was purified by filtration through poly(U)–Sephadex, and a cDNA library was constructed (10^6 pfu) and amplified using the SuperScript Lambda System (Life Technologies, Inc.; 5941SB and 8256SA, respectively) according to the manufacturer's specifications.

Library screening, DNA analysis, hybridizations, and subcloning. Plaque lifting, DNA preparation, hybridizations, and subcloning were carried out as described by Zomerdijk *et al.* [38] and Sambrook *et al.* [29]. Random-primer-labeled probes were as follows: for library screening, a human *ZO-1* cDNA fragment, clone B11 [37] kindly provided by Dr. J. M. Anderson (Yale University); for Southern analysis, the ZO-1-specific cDNA fragment from bp 2828 to 3316 in Fig. 2A.

Southern blot analysis. Genomic DNA of MDCK cells was digested with *Eco*RI, *Pvu*II, *Bam*HI, or *Xho*I restriction enzyme (10 μg/lane). DNA was size fractionated in a 0.8% agarose gel and blotted onto nitrocellulose. ^{32}P-labeled ZO-1-specific probe (nt 2828 to 3316 in Fig. 2A) was hybridized overnight at 42°C and washed two times, for 20 min in 0.2× SSC/0.1 SDS at 55°C. The blot was exposed to X-ray film for 72 h.

DNA sequence analysis. DNA sequencing was performed using the dideoxy chain termination method with the Sequenase version 2.0 enzyme as described by the manufacturer (U.S. Biochemicals; 70770).

PCR amplification of MDCK genomic DNA and RT-PCR. Regions identified as undergoing alternative splicing were amplified with the Elongase enzyme mix (Life Technologies, Inc; 10480-028) according to the manufacturer. Primer 1 encompassed from bp 5005 to 5029, primer 2 from 5485 to 5462, primer 3 from 4154 to 4178, primer 4 from 4634 to 4609, primer 5 from 5188 to 5169, primer 6 from 5203 to 5223, primer 7 from 5170 to 5190, primer 8 from 5225 to 5202, primer 9 from 2537 to 2562, and primer 10 from 3449 to 3423. To obtain the 5′ overlapping end of the *MDCK-ZO-1* cDNA, reverse transcription was primed with a 3′ oligonucleotide complementary to nt 1488–1503 in Fig. 2A.

Reverse transcription reactions for different dog tissues were primed with a downstream oligonucleotide complementary to nt 6182–6199 in Fig. 2A. Templates were either >100 ng of MDCK genomic DNA or 10% of a reverse transcription reaction prepared as described above from mRNA obtained from different canine tissues or cultured cells [11].

Northern blots and densitometry. Total RNA was extracted from MDCK cells using the Perfect RNA isolation kit (5 Prime → 3 Prime, Inc., Boulder, CO; Cat. No. 2-036364). Samples were taken from cells cultured in NC, LC, and Ca-switch (see Results) at times indicated in Fig. 9. Forty micrograms of the extracted RNA was mixed with 5 μl of formamide, incubated in a 60°C water bath for 5 min, chilled on ice for 10 min, and applied to a 1.2% (w/v) agarose gel containing 2.2 M formaldehyde. To ascertain that all lanes contained equal amounts of RNA, samples were prepared as described in Sambrook *et al.* [29] except that ethidium bromide was added to a final concentration of 125 μg/ml. After electrophoresis, ribosomal RNA bands were visualized by UV light and the relative levels of RNA compared. RNAs were transblotted overnight onto a Hybond-N$^+$ filter (Amersham) as described in Sambrook *et al.* [29]. The first 1483 bp of the *MDCK-ZO-1* gene and 2.1 kb of human actin were tagged with a DIG High Prime DNA Labeling and Detection kit for chemiluminescent detection with CSPD (Boehringer Mannheim, Cat. No. 1585614). Nitrocellulose membranes were prehybridized in 10 ml of Church buffer for 2 h and hybridized at 40°C for 24 h in Church buffer, with an excess of *MDCK-ZO-1* and actin DNA probes. After hybridization, filters were washed five times in 2× SSC and 0.1% SDS for 15 min each at room temperature. The hybridized probe was detected by chemiluminescence.

Autoradiograms were scanned in an Eagle Eye densitometer (Fotodyne, Hartland, WI) and densitometric analysis was performed with the Collage 3.0 program. The ratios of densitometric values of ZO-1 and actin at each experimental point were plotted.

Peptide synthesis. Multiple antigenic peptides (MAPs) were synthesized on a polylysine core resin [34] using Fmoc/tBu chemistry on a Synergy Model 432A (Applied Biosystems), according to the general procedures described by Merrifield [25]. The following peptides were synthesized: peptides 1, 2, 3, and 4 correspond to β_1 (ASMTPDG), β_2 (ALKSSDSSSG), and γ_1 segments (EKRYEPVQATPPPPPL) and (ALHTHAKGAHGE), respectively. After the synthesis peptides were cleaved and deprotected using standard protocols. Purity of the peptides was checked by reverse-phase HPLC using a Delta Pak C18 column (Millipore Corporation, Bedford, MA; 3.9 × 150 mm).

Generation of polyclonal antibodies. Polyclonal antibodies against synthesized multiple antigenic peptides were generated in male New Zealand rabbits according to standard protocols [16]. The antibodies were affinity purified with protein A–Sepharose and their specificity to the corresponding peptides was determined by ELISA.

Protein analysis. To detect protein expression by immunoblotting, MDCK cells were washed four times for 5 min with ice-cold PBS and lysed for 30 min in RIPA buffer (150 mM NaCl, 2 mM EDTA, 10% glycerol, 1% Triton X-100, 0.5% sodium deoxycholate, 0.2% SDS, 1 mM PMSF, 15 μg/ml leupeptin, 25 μg/ml pepstatin, 40 U/ml aprotinin, 0.2 mM sodium orthovanadate in 40 mM Tris–HCl, pH 7.6) under continuous and vigorous shaking. The extract was collected with a rubber policeman, sonicated for 6 s, and centrifuged for 30 min at 14,000 rpm to eliminate undissolved cellular debris. After protein content was determined [8], samples were separated in a 6% SDS–PAGE (50 μg per well) and subsequently transferred to PVDF or nitrocellulose Hybond-C membrane (Amersham Corp., Arlington Heights, IL). A rabbit polyclonal antibody against ZO-1 (Zymed 61-7300) and the antibodies described above against the various ZO-1 isoforms were used along with the ECL chemiluminescence detection system (Amersham).

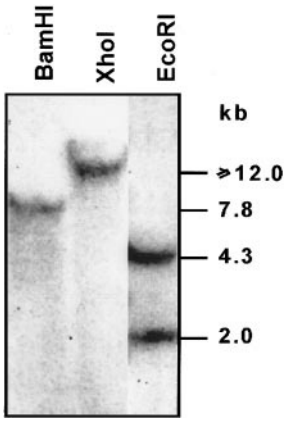


FIG. 1. Southern blot of *MDCK-ZO-1*. Genomic DNA (10 μg/lane) of MDCK cells was digested with *Bam*HI, *Xho*I, or *Eco*RI. Fractionated DNA was blotted onto nitrocellulose and hybridized with a probe that spans from nt 2828 to 3316 in Fig. 2A, i.e., most of the α region and the beginning of the proline-rich region.

Immunofluorescence. Glass coverslips containing MDCK monolayers were rinsed twice with PBS, fixed and permeabilized with −20°C methanol for 45 s, washed with PBS, incubated with 3% fetal bovine serum in PBS for 30 min, and treated with rabbit polyclonal antibodies against ZO-1 (Zymed 61-7300), ZO-1β_1, ZO-1β_2, and ZO-1γ_1 [both the 12 (γ_{1-12}) and the 16 (γ_{1-16}) segments]. Monolayers were then rinsed three times for 5 min each, with PBS, incubated with a FITC-labeled goat anti-rabbit antibody (Sigma Chemicals, St. Louis, MO) for 60 min, and rinsed as above. Occasionally, cells were incubated 5 min with propidium iodide (Sigma; P4170) and washed thoroughly, first with PBS and then with water. They were mounted with Vectashield (Vector Laboratories, Inc., Burlingame, CA) and examined with a confocal microscope (MRC-600; Bio-Rad).

Computer analysis. For sequence analysis the GCG package of programs version 8.1 (Wisconsin Genetic Computer Group) was used. The values of Table 1 were obtained using the GAP program. The multiple sequence alignments were done with PILEUP, COMPARE, and DOTPLOT programs. Nuclear consensus signals were detected using the Network interface of PSORT program version 6.4 [27]. Proteins with homologous regions were analyzed using the PAUP program version 3.1 [33].

RESULTS

ZO-1 Is Present Only Once in the Genome of MDCK Cells

The Southern blot in Fig. 1 shows that when genomic DNA is digested with *Bam*HI, only one band is found, as expected for a single copy of the *ZO-1* gene. With *Xho*I two labeled fragments of DNA could be expected (restriction site at *MDCK-ZO-1* cDNA nt 3169); however, the domain available for hybridization of one of them is so small that it remains undetectable. In the *Eco*RI digests, we detected two products of 4.3 and 2.0 kb. They arise due to an *Eco*RI site in an intron of 627 bp located between the α region and the 5′ proline-rich region (J. Valdés, unpublished observations). Therefore the above restrictions suggest that, in dog as in humans, there is a single copy of the *ZO-1* gene.

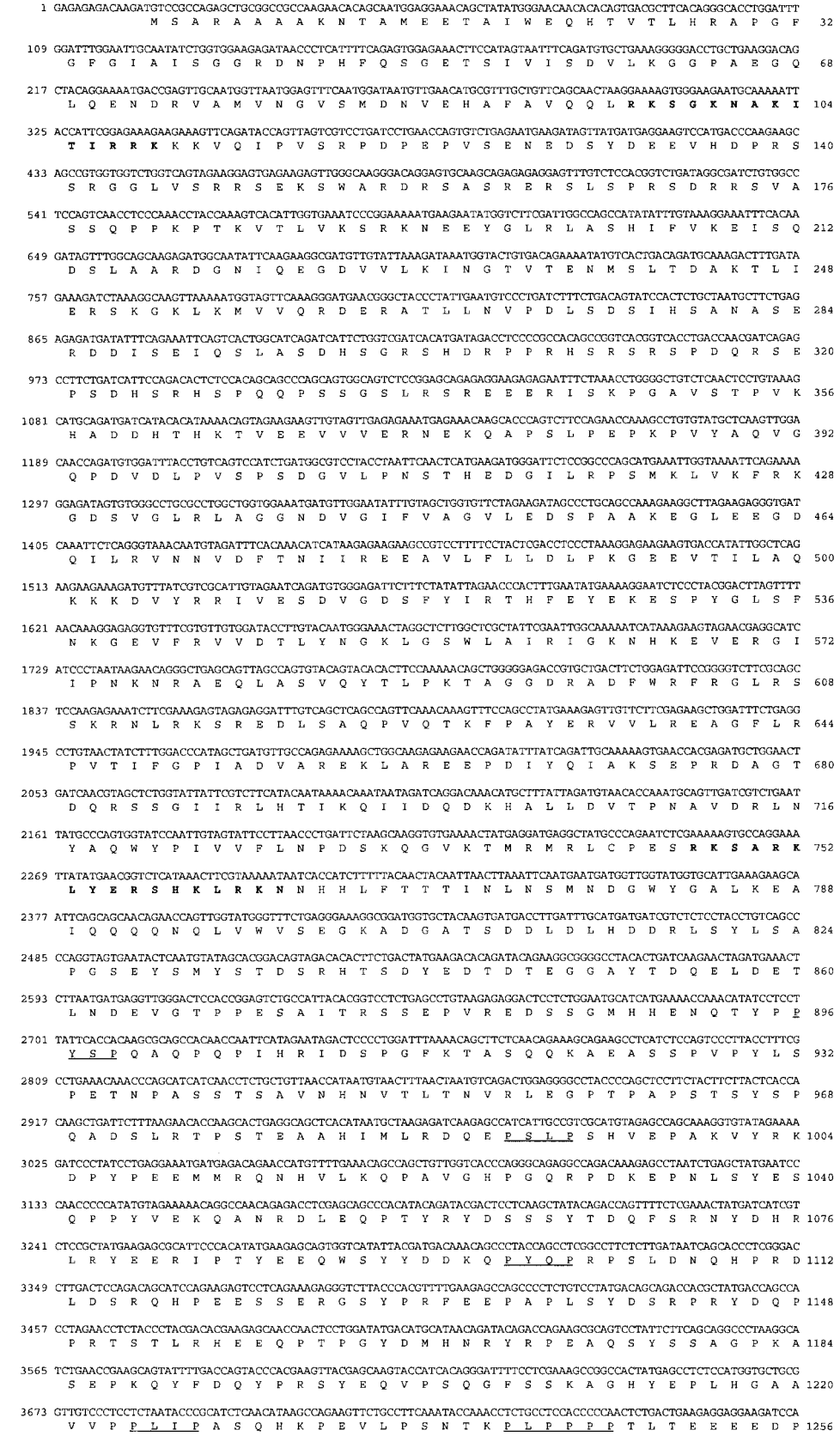
MDCK-ZO-1 cDNA Is Highly Homologous to Human and Mouse ZO-1

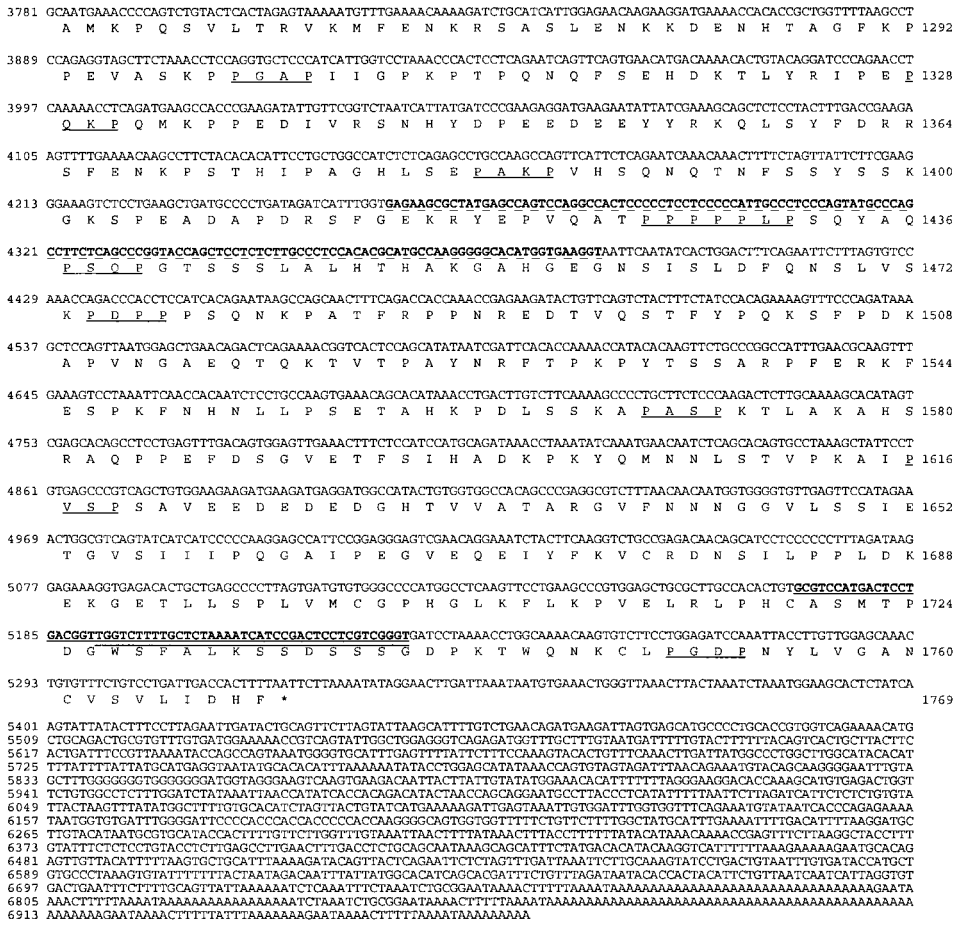
Using a human *ZO-1* probe (clone B11) we screened 10^5 plaques from an amplified MDCK cDNA library. A positive clone was selected for sequencing (bp 1360 to 6805 in Fig. 2A). By RT-PCR using the 5′ primer from *h-ZO-1* (nt 1179 to 1201 in [37]) and the 3′ primer from *MDCK-ZO-1* (nt 1455 to 1483 in Fig. 2A) we generated the upstream overlapping portion of the sequence (bp 1 to 1483 in Fig. 2A). The complete nucleotide sequence is 6805 bp long (Fig. 2A). The predicted coding region starts at position 13, followed by a 5307-nucleotide-long ORF, corresponding to 1769 amino acids that might render a protein of 212 kDa. The 3′ UTR is 1485 nucleotides long. In addition to the common MAGUK domains and α spliced domain reported for h- and m-ZO-1, it contains two previously unreported splicing regions (Fig. 2A, nucleotides in bold): β (underlined and double underlined) and γ (dashed line).

When nucleotide and predicted amino acid sequences of the *MDCK-ZO-1* ORF are compared with those of human and mouse *ZO-1*, human *ZO-2*, and *Drosophila Tamou*, the following correlations are found (Table 1): (a) A high nucleotide and amino acid identity with *h-* and *m-ZO-1*. Comparing amino acid sequences, we find that only 38 aa of *MDCK-ZO-1* are different from both human and mouse *ZO-1*, and 22 of these changes are conservative (data not shown). (b) When the complete ORF sequence of

FIG. 2. Nucleotide and deduced amino acid sequence of *MDCK-ZO-1* and comparisons with other members of the MAGUK protein family. (A) The complete nucleotide sequence is 6805 bp long. Coding region starts at position 13, followed by a 5307-nucleotide long ORF. 3′ UTR is 1485 nucleotides long. The boldface nucleotides show the β and γ splice regions. β_1 spliced form is underlined, β_2 includes the underlined and double underlined residues. γ_1 is marked with a dashed line. Two nuclear sorting signals (amino acids in boldface) are located from aa 96 to 109 (nt 298–339) and from aa 747 to 763 (nt 1251–2301). Seventeen putative PXXP motifs that potentially bind SH3 domains (underlined amino acids) are located at aa 896, 990, 1098, 1224, 1242, 1244, 1300, 1328, 1382, 1425, 1426, 1428, 1437, 1474, 1569, 1616, and 1750. Common MAGUK domains are as follows: PDZ1, aa 23 to 109 (nt 79–339); PDZ2, aa 181 to 261 (nt 553–795); PDZ3, aa 422 to 502 (nt 1276–1519); SH3 domain, aa 519 to 580 (nt 1567–1752); guanylate kinase homologous domain, aa 642 to 792 (nt 1939–2389); acidic domain, aa 817 to 894 (nt 2461–2694); α spliced domain, aa 921 to 1000 (nt 2773–3012); the four leucine residues in the leucine zipper motif are aa 753, 760, 767, and 774. (B) Nuclear sorting signals in the MAGUK family. Using PSORT version 6.4, two nuclear consensus sequences were found in MDCK-, h-, and m-ZO-1 and one for ZO-2, Dlga, and Tamou (in capitals). *1st link refers to the region located between PDZ 1 and PDZ 2. (C) A 21-aa sequence of h-ZO-1 that is not shared in MDCK- nor m-ZO-1 appears in uppercase.

A


1 GAGAGAGACAAGATGTCCGCCAGAGCTGCGGCCGCCAAGAACACAGCAATGGAGGAAACAGCTATATGGGAACAACACACAGTGACGCTTCACAGGGCACCTGGATTT
M S A R A A A A A K N T A M E E T A I W E Q H T V T L H R A P G F 32
109 GGATTTGGAATTGCAATATCTGGTGGAAGAGATAACCCTCATTTTCAGAGTGGAGAAACTTCCATAGTAATTTCAGATGTGCTGAAAGGGGGACCTGCTGAAGGACAG
G F G I A I S G G R D N P H F Q S G E T S I V I S D V L K G G P A E G Q 68
217 CTACAGGAAAATGACCGAGTTGCAATGGTTAATGGAGTTTCAATGGATAATGTTGAACATGCGTTTGCTGTTCAGCAACTAAGGAAAAGTGGGAAGAATGCAAAAATT
L Q E N D R V A M V N G V S M D N V E H A F A V Q Q L **R K S G K N A K I** 104
325 ACCATTCGGAGAAAGAAGAAAGTTCAGATACCAGTTAGTCGTCCTGATCCTGAACCAGTGTCTGAGAATGAAGATAGTTATGATGAGGAAGTCCATGACCCAAGAAGC
T I R R K K K V Q I P V S R P D P E P V S E N E D S Y D E E V H D P R S 140
433 AGCCGTGGTGGTCTGGTCAGTAGAAGGAGTGAGAAGAGTTGGGCAAGGGACAGGAGTGCAAGCAGAGAGAGGAGTTTGTCTCCACGGTCTGATAGGCGATCTGTGGCC
S R G G L V S R R S E K S W A R D R S A S R E R S L S P R S D R R S V A 176
541 TCCAGTCAACCTCCCAAACCTACCAAAGTCACATTGGTGAAATCCCGGAAAAATGAAGAATATGGTCTTCGATTGGCCAGCCATATATTTGTAAAGGAAATTTCACAA
S S Q P P K P T K V T L V K S R K N E E Y G L R L A S H I F V K E I S Q 212
649 GATAGTTTGGCAGCAAGAGATGGCAATATTCAAGAAGGCGATGTTGTATTAAAGATAAATGGTACTGTGACAGAAAATATGTCACTGACAGATGCAAAGACTTTGATA
D S L A A R D G N I Q E G D V V L K I N G T V T E N M S L T D A K T L I 248
757 GAAAGATCTAAAGGCAAGTTAAAAATGGTAGTTCAAAGGGATGAACGGGCTACCCTATTGAATGTCCCTGATCTTTCTGACAGTATCCACTCTGCTAATGCTTCTGAG
E R S K G K L K M V V Q R D E R A T L L N V P D L S D S I H S A N A S E 284
865 AGAGATGATATTTCAGAAATTCAGTCACTGGCATCAGATCATTCTGGTCGATCACATGATAGACCTCCCCGCCACAGCCGGTCACGGTCACCTGACCAACGATCAGAG
R D D I S E I Q S L A S D H S G R S H D R P P R H S R S R S P D Q R S E 320
973 CCTTCTGATCATTCCAGACACTCTCCACAGCAGCCCAGCAGTGGCAGTCTCCGGAGCAGAGAGGAAGAGAGAATTTCTAAACCTGGGGCTGTCTCAACTCCTGTAAAG
P S D H S R H S P Q Q P S S G S L R S R E E E R I S K P G A V S T P V K 356
1081 CATGCAGATGATCATACACATAAAACAGTAGAAGAAGTTGTAGTTGAGAGAAATGAGAAACAAGCACCCAGTCTTCCAGAACCAAAGCCTGTGTATGCTCAAGTTGGA
H A D D H T H K T V E E V V V E R N E K Q A P S L P E P K P V Y A Q V G 392
1189 CAACCAGATGTGGATTTACCTGTCAGTCCATCTGATGGCGTCCTACCTAATTCAACTCATGAAGATGGGATTCTCCGGCCCAGCATGAAATTGGTAAAATTCAGAAAA
Q P D V D L P V S P S D G V L P N S T H E D G I L R P S M K L V K F R K 428
1297 GGAGATAGTGTGGGCCTGCGCCTGGCTGGTGGAAATGATGTTGGAATATTTGTAGCTGGTGTTCTAGAAGATAGCCCTGCAGCCAAAGAAGGCTTAGAAGAGGGTGAT
G D S V G L R L A G G N D V G I F V A G V L E D S P A A K E G L E E G D 464
1405 CAAATTCTCAGGGTAAACAATGTAGATTTCACAAACATCATAAGAGAAGAAGCCGTCCTTTTCCTACTCGACCTCCCTAAAGGAGAAGAAGTGACCATATTGGCTCAG
Q I L R V N N V D F T N I I R E E A V L F L L D L P K G E E V T I L A Q 500
1513 AAGAAGAAAGATGTTTATCGTCGCATTGTAGAATCAGATGTGGGAGATTCTTTCTATATTAGAACCCACTTTGAATATGAAAAGGAATCTCCCTACGGACTTAGTTTT
K K K D V Y R R I V E S D V G D S F Y I R T H F E Y E K E S P Y G L S F 536
1621 AACAAAGGAGAGGTGTTTCGTGTTGTGGATACCTTGTACAATGGGAAACTAGGCTCTTGGCTCGCTATTCGAATTGGCAAAAATCATAAAGAAGTAGAACGAGGCATC
N K G E V F R V V D T L Y N G K L G S W L A I R I G K N H K E V E R G I 572
1729 ATCCCTAATAAGAACAGGGCTGAGCAGTTAGCCAGTGTACAGTACACACTTCCAAAAACAGCTGGGGGAGACCGTGCTGACTTCTGGAGATTCCGGGGTCTTCGCAGC
I P N K N R A E Q L A S V Q Y T L P K T A G G D R A D F W R F R G L R S 608
1837 TCCAAGAGAAATCTTCGAAAGAGTAGAGAGGATTTGTCAGCTCAGCCAGTTCAAACAAAGTTTCCAGCCTATGAAAGAGTTGTTCTTCGAGAAGCTGGATTTCTGAGG
S K R N L R K S R E D L S A Q P V Q T K F P A Y E R V V L R E A G F L R 644
1945 CCTGTAACTATCTTTGGACCCATAGCTGATGTTGCCAGAGAAAAGCTGGCAAGAGAAGAACCAGATATTTATCAGATTGCAAAAAGTGAACCACGAGATGCTGGAACT
P V T I F G P I A D V A R E K L A R E E P D I Y Q I A K S E P R D A G T 680
2053 GATCAACGTAGCTCTGGTATTATTCGTCTTCATACAATAAAACAAATAATAGATCAGGACAAACATGCTTTATTAGATGTAACACCAAATGCAGTTGATCGTCTGAAT
D Q R S S G I I R L H T I K Q I I D Q D K H A L L D V T P N A V D R L N 716
2161 TATGCCCAGTGGTATCCAATTGTAGTATTCCTTAACCCTGATTCTAAGCAAGGTGTGAAAACTATGAGGATGAGGCTATGCCCAGAATCTCGAAAAAGTGCCAGGAAA
Y A Q W Y P I V V F L N P D S K Q G V K T **M R M R L C P E S R K S A R K** 752
2269 TTATATGAACGGTCTCATAAACTTCGTAAAAATAATCACCATCTTTTTACAACTACAATTAACTTAAATTCAATGAATGATGGTTGGTATGGTGCATTGAAAGAAGCA
L Y E R S H K L R K N N H H L F T T T T I N L N S M N D G W Y G A L K E A 788
2377 ATTCAGCAGCAACAGAACCAGTTGGTATGGGTTTCTGAGGGAAAGGCGGATGGTGCTACAAGTGATGACCTTGATTTGCATGATGATGATCGTCTCTCCTACCTGTCAGCC
I Q Q Q Q N Q L V W V S E G K A D G A T S D D L D L H D D R L S Y L S A 824
2485 CCAGGTAGTGAATACTCAATGTATAGCACGGACAGTAGACACACTTCTGACTATGAAGACACAGATACAGAAGGCGGGGCCTACACTGATCAAGAACTAGATGAAACT
P G S E Y S M Y S T D S R H T S D Y E D T D T E G G A Y T D Q E L D E T 860
2593 CTTAATGATGAGGTTGGGACTCCACCGGAGTCTGCCATTACACGGTCCTCTGAGCCTGTAAGAGAGGACTCCTCTGGAATGCATCATGAAAACCAAACATATCCTCCT
L N D E V G T P P E S A I T R S S E P V R E D S S G M H H E N Q T Y P P 896
2701 TATTCACCACAAGCGCAGCCACAACCAATTCATAGAATAGACTCCCCTGGATTTAAAACAGCTTCTCAACAGAAAGCAGAAGCCTCATCTCCAGTCCCTTACCTTTCG
Y S P Q A Q P Q P I H R I D S P G F K T A S Q Q K A E A S S P V P Y L S 932
2809 CCTGAAACAAACCCAGCATCATCAACCTCTGCTGTTAACCATAATGTAACTTTAACTAATGTCAGACTGGAGGGGCCTACCCCAGCTCCTTCTACTTCTTACTCACCA
P E T N P A S S T S A V N H N V T L T N V R L E G P T P A P S T S Y S P 968
2917 CAAGCTGATTCTTTAAGAACACCAAGCACTGAGGCAGCTCACATAATGCTAAGAGATCAAGAGCCATCATTGCCGTCGCATGTAGAGCCAGCAAAGGTGTATAGAAAA
Q A D S L R T P S T E A A H I M L R D Q E P S L P S H V E P A K V Y R K 1004
3025 GATCCCTATCCTGAGGAAATGATGAGACAGAACCATGTTTTGAAACAGCCAGCTGTTGGTCACCCAGGGCAGAGGCCAGACAAAGAGCCTAATCTGAGCTATGAATCC
D P Y P E E M M R Q N H V L K Q P A V G H P G Q R P D K E P N L S Y E S 1040
3133 CAACCCCCATATGTAGAAAAACAGGCCAACAGAGACCTCGAGCAGCCCACATACAGATACGACTCCTCAAGCTATACAGACCAGTTTTCTCGAAACTATGATCATCGT
Q P P Y V E K Q A N R D L E Q P T Y R Y D S S S Y T D Q F S R N Y D H R 1076
3241 CTCCGCTATGAAGAGCGCATTCCCACATATGAAGAGCAGTGGTCATATTACGATGACAAACAGCCCTACCAGCCTCGGCCTTCTCTTGATAATCAGCACCCTCGGGAC
L R Y E E R I P T Y E E Q W S Y Y D D K Q P Y Q P R P S L D N Q H P R D 1112
3349 CTTGACTCCAGACAGCATCCAGAAGAGTCCTCAGAAAGAGGGTCTTACCCACGTTTTGAAGAGCCAGCCCCTCTGTCCTATGACAGCAGACCACGCTATGACCAGCCA
L D S R Q H P E E S S E R G S Y P R F E E P A P L S Y D S R P R Y D Q P 1148
3457 CCTAGAACCTCTACCCTACGACACGAAGAGCAACCAACTCCTGGATATGACATGCATAACAGATACAGACCAGAAGCGCAGTCCTATTCTTCAGCAGGCCCTAAGGCA
P R T S T L R H E E Q P T P G Y D M H N R Y R P E A Q S Y S S A G P K A 1184
3565 TCTGAACCGAAGCAGTATTTTGACCAGTACCCACGAAGTTACGAGCAAGTACCATCACAGGGATTTTCCTCGAAAGCCGGCCACTATGAGCCTCTCCATGGTGCTGCG
S E P K Q Y F D Q Y P R S Y E Q V P S Q G F S S K A G H Y E P L H G A A 1220
3673 GTTGTCCCTCCTCTAATACCCGCATCTCAACATAAGCCAGAAGTTCTGCCTTCAAATACCAAACCTCTGCCTCCACCCCCAACTCTGACTGAAGAGGAGGAAGATCCA
V V P P L I P A S Q H K P E V L P S N T K P L P P P P T L T E E E E D P 1256



B

				Location	
MDCK-ZO-1	92	vqqlRKSGKNAKITIRRKKKVqipv	116	PDZ 1 and 1st link*	
h-ZO-1	79	vqqlRKSGKNAKITIRRKKKVqipv	104	PDZ 1 and 1st link	
m-ZO-1	91	vqqlRKSGKNAKITIRRKKKVqipv	115	PDZ 1 and 1st link	
h-ZO-2	101	vqqlRKSGKVAAIVVKRPRKVqvaa	125	PDZ 1 and 1st link	
MDCK-ZO-2	79	vqqlRKSGKIAAIVVKRPRKVqlap	103	PDZ 1 and 1st link	
MDCK-ZO-3	135	dkrsRRARTGRRNQAGSRGRRspgg	159	1st link	
Dlga	108	vdalKKAGNVVKLHVKRKRGTattp	132	PDZ 1 and 1st link	
Tamou	569	fnatKKEMNANESRGNFFRRRrsth	593	Between SH3 and GK	
MDCK-ZO-1	743	cpesRKSARKLYERSHKLRKNnhhl	767	GK	
h-ZO-1	732	cpesRKSARKLYERSHKLAKNnhhl	756	GK	
m-ZO-1	743	cpesRKSARKLYERSHKLRKNnhhl	767	GK	
MDCK-ZO-3	702	apasRRSARRLYAQAQQLRQHsehl	726	GK	

C

MDCK-ZO-1	388	yaqvgqp.....dvdlpvspsdgvlpnsthedgilrps	420
h-ZO-1	377	yaqvgNQMWIYLSVHLMDSYLIQLM.....KMGFlrps	409
m-ZO-1	389	yaqvgqp.....dvdlpvspsdgalpnsahedgilrps	421

FIG. 2—*Continued*

MDCK-ZO-1 is compared to the complete ORF of *h-ZO-2*, a lower nucleotide identity is found (38%), which nevertheless codes for an amino acid sequence of high similarity (71%). If the proline-rich domain is excluded from this comparison, the nucleotide identity increases from 38 to 61%. (c) When *MDCK-ZO-1*

TABLE 1

Comparison of MDCK-ZO-1 with Different Related Proteins or Protein Segments

	bp % identity	aa % similarity	aa % identity
h-ZO-1	92	96	93
h-ZO-1 γ	39	91	89
m-ZO-1	87	95	91
m-ZO-1 γ	78	82	75
h-ZO-2	38	71	54
h-ZO-2 3′Δ	61	76	59
Tamou	44	54	32
Tamou 3′Δ	48	60	39
Tamou 5′Δ	38	47	23

Note. Nucleotide and amino acid sequences of the *MDCK-ZO-1* ORF were compared with the respective ORFs of the human, mouse, and *Drosophila* counterparts using the GAP program of GCG. In addition exclusive domains were compared by GAP: γ corresponds to the γ insertions in the respective molecules, 3′Δ comprises the ORF start to the last aa in the GK domain in the different proteins, and 5′Δ comprises the first aa after the GK domain to the last aa of the proteins.

is compared with *Tamou*, a MAGUK protein associated to the septate junction of *Drosophila*, much lower nucleotide and amino acid identities are found, even if this comparison excludes the proline-rich domain.

MDCK-ZO-1 cDNA Has Two Nuclear Sorting Signals and 17 SH3 Binding Motifs

Using the PSORT program (v 6.4) [27] two nuclear consensus signals were detected in MDCK-ZO-1 between amino acids 92 and 116 and between 743 and 767 (Fig. 2B, amino acids in bold). Applying this program we also detected these sequences in h- and m-ZO-1. The first of these sequences is also present in MAGUK proteins ZO-2 and *Drosophila* Dlg. Tamou presents another nuclear consensus signal between amino acids 569 and 593.

While MDCK-ZO-1 has 17 PXXP motifs in the proline-rich region, which may potentially bind to SH3 domains (Fig. 2A, underlined amino acids), m- and h-ZO-1 exhibit 16 and 15, respectively. Interestingly, 15 of such motifs are also found in the same proline-rich domain of Tamou, despite the fact that this protein has a low basepair identity in the 3′ region (48%), compared to *MDCK-ZO-1*. In addition, in the very short proline-rich region of ZO-2, we identify two of these PXXP motifs (data not shown).

MCDK-ZO-1 Has Three Alternative Splicing Sites

Upon comparing the predicted amino acid sequence of MDCK-ZO-1 with those of h- and m-ZO-1 using the DOTPLOT program (data not shown), we noticed three main differences: (a) h-ZO-1 contains a 21-aa sequence toward the amino-terminal portion of the protein that is not shared with MDCK- nor m-ZO-1 (Fig. 2C). A search through the Blast program provided no clue of homology between this segment and other proteins listed in GenBank. (b) MDCK-ZO-1 contains an insertion at the carboxyl-terminal of the protein, which we name β and which was not previously reported in its h- and m- counterparts (Fig. 2A, bold, underlined and double underlined nucleotides) and (c) MDCK-ZO-1 contains an insertion we call γ, previously sequenced in human and mouse ZO-1, but not identified as a splice region (Fig. 2A, bold nucleotides, dash-underlined). To explore the possibility that these insertions could have originated through alternative splicing, we performed RT-PCR using MDCK mRNA and flanking primers 1 and 2 for the β region and 3 and 4 for the γ region. To further define the nature of the insertions, PCR products were cloned and sequenced. For the β region three isoforms were generated (Fig. 3A): β_2 (480 bp), which contains a 60-nt insertion (bp 5170 to 5229 in Fig. 2A); β_1 (441 bp), which contains the first 21 nt of the β_2 insert (i.e., from bp 5170 to 5190); and β_0 (420 bp), which lacks the complete insert. In turn, the γ region originates three isoforms (Fig. 3B): γ_2 (900 bp), yet to be characterized; γ_1 (480 bp), with an insertion of 135 nt (from bp 4255 to 4389 in Fig. 2A); and γ_0 (345 bp), which lacks the insertion.

As a first step toward understanding the biological meaning of this variability in MDCK isoforms, we investigated their expression in different dog tissues. The presence of β and γ isoforms in several dog tissues, other than kidney, was analyzed by performing RT-PCR using the flanking primers described above for β and γ regions. As shown in Fig. 3A, testis expresses the same three β isoforms as MDCK cells. Aorta, ileum, brain, and lung express only β_0 and β_1 isoforms. Figure 3B shows that testis, aorta, and ileum express isoforms γ_0, γ_1, and γ_2. Lung and brain lack the γ_2 isoform. Testis has an additional band of 400, and aorta has bands of 600 and 700 bp, which represent other γ isoforms.

To determine the size of the intronic regions surrounding the β domain, we performed PCR experiments on genomic DNA. For the region ahead of β_1 we used primers 1 and 5, the latter constituting the 5′ end of β_1 exon, and found a PCR product of 1720 bp (Fig. 4, lane 1). Subtracting the basepairs of cDNA, it yields an intronic region of 1537 bp. When we used primer 6, encompassing the 3′ end of β_2 exon and primer 2, we obtained a PCR product of 2966 bp, of which 283 belong to the cDNA and 2683 to the intron (Fig. 4, lane 3). When primer 7, corresponding to the 3′ end of β_1, and primer 8, located at the 5′ end of β_2, were used, we obtained no detectable PCR product (Fig. 4, lane 2).

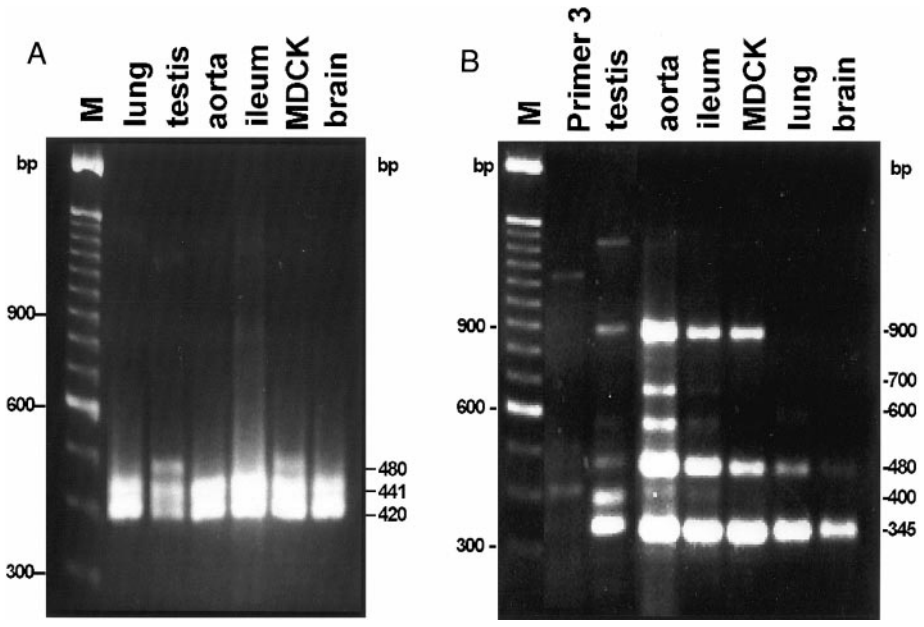


FIG. 3. Expression of β and γ isoforms in MDCK cells and several dog tissues. Poly(A)$^+$ RNA was extracted from different tissues and RT-PCR was performed. (A) For β isoforms flanking primers 1 (bp 5005–5029) and 2 (bp 5485–5462) were used. In testis and MDCK cells three bands of 420, 441, and 480 bp, corresponding to isoforms β_0, β_1, and β_2, respectively, are observed. In aorta, ileum, brain, and lung only the 420- and 441-bp bands were detected. (B) For γ isoforms, primers 3 (bp 4154–4178) and 4 (4634–4610) were used. Testis, aorta, MDCK cells, and ileum expressed three bands of 345, 480, and 900 bp, corresponding to isoforms γ_0, γ_1, and γ_2, respectively. Lung and brain lack the γ_2 isoform. Additional bands of 400 and of 600 and 700 bp are present in testis and aorta, respectively.

Finally, when primers 7 and 2 were used, we obtained a PCR product of 3208 bp (Fig. 4, lane 4), of which only 315 appear in the cDNA. The resulting nontranscribed 2893 bp compares with 2683 obtained as the intronic region obtained after β_2. Therefore, an intronic region of 210 bp is expected between β_1 and β_2 domains, even when it cannot be detected in the gel. In summary, the two detected introns plus the β exons account for the ca. 4800 bp found when genomic DNA was amplified around the β region using primers 1 and 2 (Fig. 5). In this gel three bands of 424, 441, and 480 bp can also be observed. When they were eluted from the gel and sequenced, they were found to correspond to β_0, β_1, and β_2 isoforms, respectively. As discussed below, the possibility exists that the presence of these three bands in genomic DNA is due to a processed pseudogene (see Discussion).

MDCK Cells Express β and γ Isoforms of ZO-1

The existence of several mRNAs poses the question of whether MDCK cells do express different isoforms of the ZO-1 protein. This was tested by preparing polyclonal antibodies against synthetic peptides of the β and γ predicted sequences. Figure 6 shows the bands obtained by immunoblotting with a commercial polyclonal antibody against ZO-1 and the antibodies raised against the different ZO-1 isoforms. As previously reported in MDCK cells, the pattern consists of a clear band of 212 kDa [1] and a fainter one of 190 kDa [5], which in the case of the commercial antibody can be observed only when the blot is overexposed (data not shown). Therefore these results suggest that isoforms β_1, β_2, γ_{1-12}, and γ_{1-16} are present in both α^+ and α^- ZO-1. Figure 7 shows the patterns of ZO-1 isoform distribution obtained in confluent monolayers of MDCK cells observed with immunofluorescence. Figure 7a illustrates the typical pattern of ZO-1 occupying the periphery of the cells. Figures 7b, 7c, 7d, and 7e show that all molecular species carrying the corresponding fragments are in fact expressed in the periphery of MDCK cells, as expected for a TJ-associated protein. Figures 7b to 7e were obtained with antibodies prepared against synthetic peptides. This forced us to use high concentrations, which might result in considerable background. Those containing β_1, β_2, and γ_{1-12} (Figs. 7b, 7c and 7d) show accumulation in discrete points within the cytoplasm, and that with γ_{1-16} shows a marked nuclear staining. Since both antibodies γ_{1-12} and γ_{1-16} are directed against the same isoform, the observation that the second exhibits a nuclear staining might indicate either that it crossreacts with a nuclear protein or that ZO-1 is in the nucleus in a slightly altered conformation in which the binding site for γ_{1-16} is accessible, while γ_{1-12} is not.

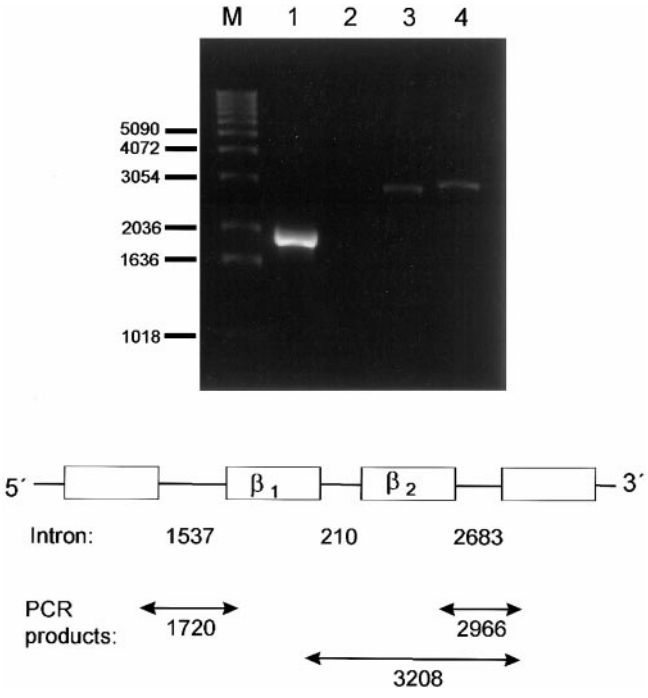


FIG. 4. PCR amplifications of genomic DNA surrounding the β region. M, molecular weight markers. Lane 1, product obtained with primers 1 and 3; lane 2, with primers 7 and 8; lane 3, with primers 6 and 2; and lane 4, with primers 7 and 2. The diagram at the bottom schematizes the genomic regions that surround the β_1 and β_2 splice regions. The size of the introns was calculated according to the PCR products obtained with the different sets of primers.

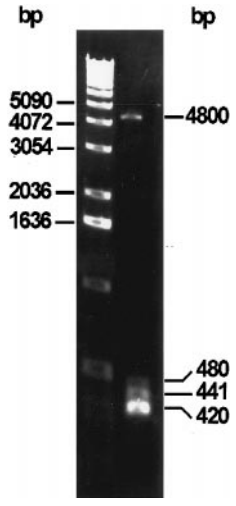


FIG. 5. Expression of β region on genomic DNA. PCR experiments were performed on genomic DNA using primers 1 and 2 flanking the β domain. Four bands of 420, 441, 480, and 4800 bp were observed. The first three correspond to β_0 to β_2, and the 4.8-kb band corresponds to the genomic β environment.

MDCK-ZO-1 Phylogenetic Relationships

Multiple sequence alignments of MDCK-ZO-1 and proteins with homologous regions, done with the PILEUP program, were used as input to the PAUP exhaustive parsimony algorithm [33]. The unrooted tree in Fig. 8A shows that this group of proteins can be divided into four main branches: (1) A branch of tight and septate junction-associated proteins (MDCK-ZO-1, m-ZO-1, h-ZO-1, h-ZO-2, MDCK-ZO-2, MDCK-ZO-3, and Tamou). Surprisingly, these proteins are more closely related to hPTP1e and rnNOS, which possess PDZ domains but do not belong to the MAGUK family, than to those proteins in the other branches that do belong to this family. (2) A branch of synaptic proteins (Chapsyn-110, SAP97, PSD95, SAP90, Dlg). Interestingly, Dlg, a protein associated with septate junctions, the insect homologues of tight junctions, appears to be more related to proteins of the synapses than to those of the tight junction. A couple of independent branches group the MAGUK proteins LIN2 (3) and p55 (4). These proteins are known to function as clustering centers for other protein species, such as neurexin [18], protein 4.1 [7], and glycophorin C [23].

The group of proteins listed in Fig. 8A was also analyzed for the three splicing domains (I1–3) present in hDlg [22]. According to our PILEUP analysis, these splicing domains are present only in proteins belonging to the synaptic branch. SAP97 presents domains similar to I1 (an amino-terminal insertion, at the proline-rich region) and to I2 (an insertion located between SH3 and GK). Chapsyn-110 and dDlg present an I3-

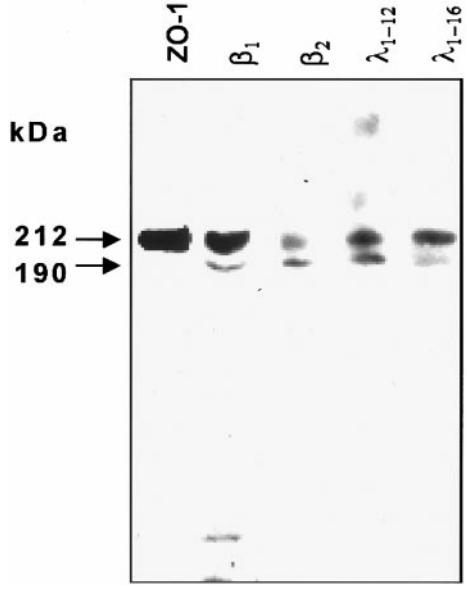


FIG. 6. Expression of ZO-1 and isoforms in MDCK cells. ZO-1 was detected with Zymed 7300 polyclonal antibody, and β_1, β_2, γ_{1-12}, and γ_{1-16} were detected with polyclonal antibodies generated against their respective synthetic peptides. Arrows indicate the ZO-1 bands found in MDCK cells. These images are representative examples of five independent experiments.

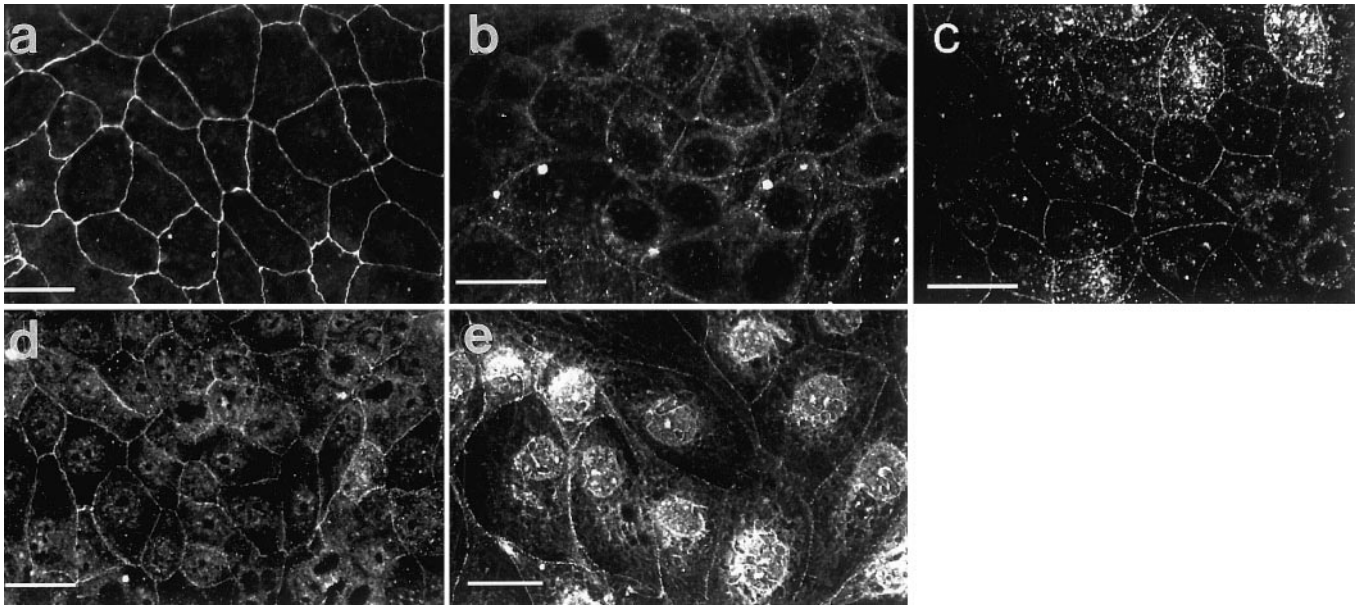


FIG. 7. Subcellular distribution of ZO-1 and isoforms in confluent MDCK monolayers. The first antibody used in each case was prepared against: (a) ZO-1, (b) ZO-1β_1, (c) ZO-1β_2, (d) ZO-1γ_{1-12}, and (e) ZO-1γ_{1-16}. Bars, 20 μm.

like insertion, which has been reported in hDlg as a protein 4.1 binding region [24] (Fig. 8B).

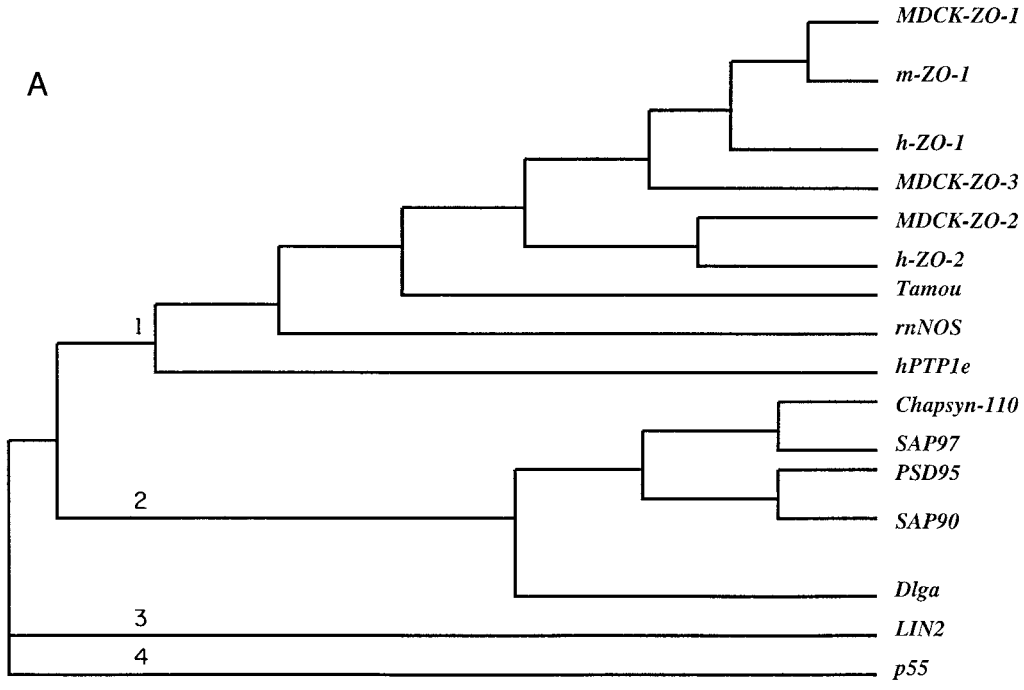
While the Expression of MDCK-ZO-1 mRNA Is Ca^{2+} Regulated, That of Its Protein Is Not

The next step was to investigate the size and level of expression of *MDCK-ZO-1* mRNA in cells incubated with Ca^{2+} (1.8 mM, NC), without this ion (5 μM, LC), and during a calcium switch (CS). A Northern blot probed with a 3′-most *Eco*RI cDNA fragment of *MDCK-ZO-1* (4413 to 6527 bp in Fig. 2A) shows a *ZO-1* messenger of 7.4 kb (Fig. 9A). Figure 9B shows the blots of *ZO-1* (probed with the first 1483 bp of *MDCK-ZO-1*) and actin (probed with 2.1 kb of human actin); the latter hybridizations were taken as transcriptional expression controls. Figure 9C displays the densitometric values of the blots as a function of time after plating the cells at confluence. Newly plated cells in LC medium have a relatively low mRNA level, which thereafter increases sharply. Cells plated in NC instead show a somewhat higher and stable level of mRNA, which is nevertheless much lower than the value reached by LC-cultured cells. Transference of monolayers in LC to NC medium causes a marked reduction of *ZO-1* mRNA to a steady value. However, these changes are not reflected in the protein level, as ZO-1 content increases in the first 12 h after plating, irrespective of the presence or absence of Ca^{2+} in the medium (Fig. 10).

DISCUSSION

The Southern blot of genomic DNA restricted with *Bam*HI, *Xho*I, and *Eco*RI indicates that there is only one copy of the *ZO-1* gene in MDCK genome, a situation similar to that in humans [37]. *ZO-1* from MDCK cells has a 92 and 87% basepair identity to the genes in human and mouse, respectively, indicating that it is a highly conserved protein. Surprisingly, while basepair identity between *MDCK-ZO-1* and *h-ZO-2*, another MAGUK protein associated to tight junctions, is only 38%, the similarity between amino acids is very high (71%).

Before maturation and during remodeling of cell–cell contacts, ZO-1 can be detected in the nucleus [15]. This might be due to the presence of two lysine-rich stretches of amino acids: 95-RRKKK and 490-KKKD-VYRR. Nakai and Kanehisa's program for the detection of nuclear consensus signals does not attribute a high probability to the second one, but detects the first between amino acids 92 and 116 of MDCK-ZO-1. Additionally, this program also finds another nuclear sorting signal between amino acids 743 and 767. When we performed this analysis in the whole MAGUK family, we found that these nuclear sorting signals are also present in TJ-associated proteins ZO-2 and ZO-3 and in the septate junction-associated Dlga and Tamou. Interestingly, these sorting signals were not detected in the synapses-associated proteins: SAP90, PSDd95, SAP97, and Chapsyn-110.



B

```
I1
  hDlg 157 ispikPTEAVLPSPPTVPVIPVLPVPAENTVILPTIPQanpp 198
 SAP97 159 ispikPTEAVPPSSPIVPVTPALPVPAESPVVLPSTPQanpp 200

I2
  hDlg 632 rdkgEIPDDMGSKGLKhvts 651
 SAP97 665 gdkgEIPDDMGSKGLKhvts 684

I3
   hDlg 632 rdkgQSFNDKRKKN·LFSRKFPFYKNKDQSEQETSDA·DQhvt 672
   dDlg 689 ldk·QSTLDRKKKNFTFSRKFPFMKSRD··EKNE·DGSDQepn 727
Chap110 623 dskg·SFNDKRKKSFIFSRKFPFYKNKEQSEQETSD····per 660
```

FIG. 8. Phylogenetic analysis of MDCK-ZO-1 and other proteins with homologous domains. (A) Multiple sequence alignments of MDCK-ZO-1 and proteins with homologous regions were analyzed using the PAUP exhaustive parsimony algorithm. The unrooted tree was obtained after 100 bootstrap replicates. The four main branches displayed correspond to: (1) tight and septate junction-associated proteins and the PDZ domains of two neural enzymes, (2) synaptic proteins, and (3 and 4) independent branches. Nomenclature and accession numbers: h-ZO-1 (human, L14837); m-ZO-1 (mouse, D14340); ZO-2 (human, L27476); MDCK-ZO-2 (canine, L27152); MDCK-ZO-3 (canine, AF023617); Tamou, ZO-1 homologue (*Drosophila*, D83477); LIN2, vulval induction protein (*Caenorhabditis elegans*, X92564); p55, erythrocyte membrane protein (human, M64925); SAP90, synaptic protein (rat, X66474); PSD95, synaptic protein (rat, M96853); SAP97, synaptic protein (rat, U14950); Chapsyn-110, synaptic protein (human, U32376); Dlga, tumor supressor protein (*Drosophila*, M73527); hPTP1e, tyrosine phosphatase (human, U12128); and rnNOS, neuron nitric oxide synthase (rat, X59949). (B) Comparison of I1–I3 splicing domains in hDlg with similar domains found in dDlg, SAP97, and Chapsyn-110. Insertions are shown in uppercase, and identical amino acids are in bold. For better alignment some spaces were introduced.

All MAGUK proteins present at least one PDZ domain, an SH3 region, and a GK domain, which are highly conserved. However, within this superfamily only proteins associated with TJ and septate junctions have a proline-rich domain. Interestingly, it is precisely this region that presents the potentially binding SH3 motifs and alternative splicing domains. The β isoforms that we find in MDCK cells have not been reported for other species. J. M. Anderson though, has found identical counterparts in *h-ZO-1* (personal communication). The γ isoforms, instead, share high similarity with both human and mouse counterparts (Table 1). It should be noted that, while *m-ZO-1*-γ has a very high nucleotide (78%) and amino acid (75%) identity,

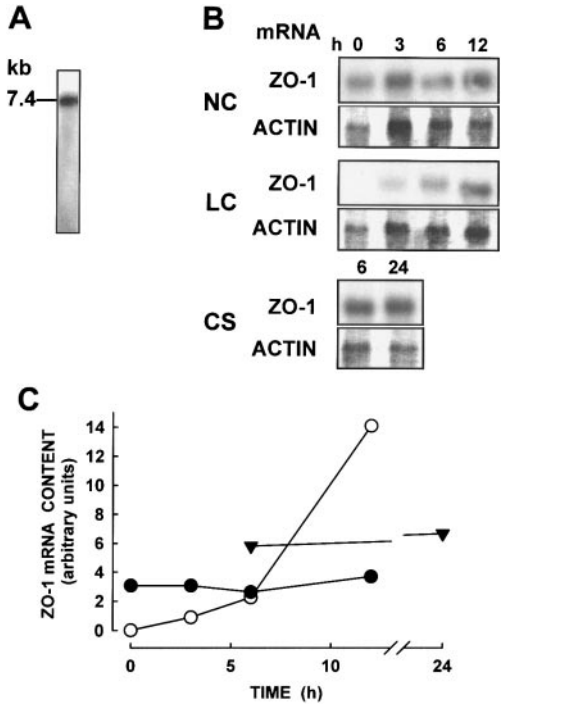


FIG. 9. Kinetics of *ZO-1* mRNA expression in the presence of absence of Ca^{2+}. (A) Northern blot of *MDCK-ZO-1* shows a 7.4-kb messenger. (B) *MDCK-ZO-1* and actin mRNA are detected at several experimental time points (h) in samples taken from monolayers cultured with calcium (NC) or without this ion (LC) or subjected to a calcium switch (CS). (C) The ratios of densitometric values found for *MDCK-ZO-1* and actin are shown in arbitrary units for each time point. Filled circles (NC), empty circles (LC), triangles (CS).

h-ZO-1-γ has a relatively low nucleotide identity (39%) but a significantly high amino acid identity (89%). The combination of the three β (for example, in testis) and five γ isoforms (for example, in aorta) indicates that in principle, the genome can offer 15 different isoforms of *ZO-1*. Taking the two α isoforms, the number would increase to 30. This suggests that whatever the role of ZO-1 in the tight junction, its function may be highly modulated. This possibility is further stressed by the fact that the tissues tested exhibit different combinations of the isoforms.

PCR studies of genomic DNA not only support the notion that the β region originated by alternative splicing, but reinforce the observation made with Southern analysis that a single copy of *ZO-1* is present in the MDCK genome, as reported for humans. As stated above, the three lower bands observed in Fig. 5, when genomic DNA is amplified with a primer flanking the β domain, might be due to processed pseudogenes, i.e., nonfunctional genomic copies of mRNA that might have been reversed-transcribed and randomly integrated into the genome [36].

The fact that antibodies prepared against discrete segments contained within β and γ isoforms detect ZO-1 in immunoblots and exhibit the immunofluorescent pattern expected for TJ-associated molecules indicates that these isoforms in fact represent products of alternative splicing of ZO-1.

Phylogenetic analysis shows that MAGUK proteins divide into four main branches (Fig. 8A). Three of them contain proteins pertaining to the septate junction (Tamou, Dlg, and Lin2), suggesting that this type of cell–cell junction may be an ancestor of TJs and synapses. Lin2 and p55 appear in independent branches of the unrooted tree. It is noteworthy that certain key amino acids in the PDZ domains in these two proteins are different from those in the PDZ regions of the other MAGUK proteins in Fig. 8A. This enables PDZ domains to identify different motifs in the proteins that they recognize. Thus, while the PDZ groups in LIN2 and p55 proteins target to a carboxyl-terminal motif bearing a hydrophobic or aromatic amino acid (phenylalanine) in the -2 position, PDZ groups in the other

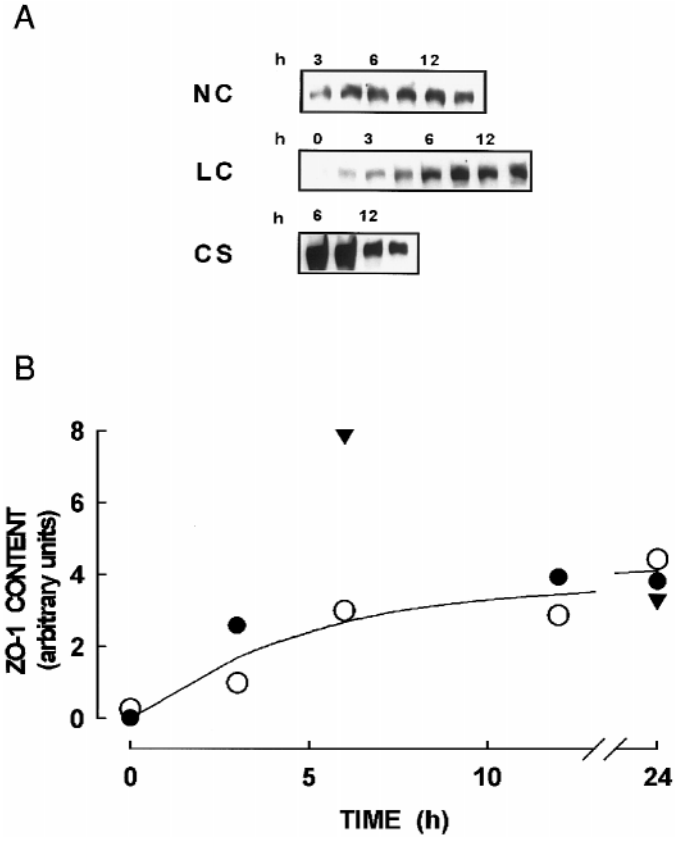


FIG. 10. Kinetics of ZO-1 protein expression in the presence or absence of Ca^{2+}. (A) Western blot of ZO-1 at several experimental time points in samples taken from monolayers cultured with calcium (NC) or without this ion (LC) or subjected to a calcium switch (CS). Numbers refer to hours. Each time point is represented by two lanes. (B) Ratio of densitometric values found for MDCK-ZO-1 are shown in arbitrary units for each time point. Filled circles (NC), empty circles (LC), triangles (CS) The curve was drawn by eye.

MAGUK proteins recognize peptides with an −OH (tyrosine, threonine, or serine) in such position [32].

Dlg appears more closely related to proteins of the synaptic branch than to MAGUK proteins associated to TJs. Furthermore, it is pertinent to notice that certain mutations in Dlg generate defects in synapses formation [21]. On top of these similarities, human Dlg presents three insertion sequences (I1–3) also present in SAP97 and Chapsyn-110.

Although proteins in the tree presented in Fig. 8A show no common ancestor, all the structures in which the listed MAGUK proteins participate follow a general pattern: (a) A transmembrane protein (occludin, a Shaker K-channel, neurexin, glycophorin C), (b) a MAGUK protein (ZO-1, Dlg, Lin2/CASK, or p55), and (c) a cytoskeleton-associated protein. In the second, third, and fourth branches of Fig. 8A protein 4.1 fulfills this role. The corresponding protein in the first branch has not yet been identified. Nevertheless, it is well known that the cytoskeleton participates in the assembly of TJs [13, 26]. In nonepithlelial cells, ZO-1 functions as a crosslinker between the cadherin–catenin complex and the actin-based cytoskeleton through direct interaction with α-catenin and actin filaments at is amino- and carboxyl-terminal halves, respectively [20].

With the information obtained from the cDNA sequence, the genomic environments around α (1.0 kb), β (4.32 kb), and γ (0.42 kb) splicing sites, and the determination allowed by the Northern blot of the size of the 5′ untranslated region (0.6 kb), we estimate that the *MDCK-ZO-1* gene is at least 13.16 kb long, not considering promoter sequences.

While the size of *MDCK-ZO-1* mRNA illustrated in Fig. 9A is 7.4 kb, mRNAs from *h-* and *m-ZO-1* are 7.9 and 7.3 kb long, respectively [19, 37]. The messenger level in cells newly plated in LC is very low. This is strikingly different from the levels found in newly plated Caco-2 cells that had previously been maintained in suspension without Ca^{2+} for 48 h [2]. Upon plating with NC, Caco-2 cells show a threefold reduction of mRNA, which is in keeping with results presented in Fig. 9C. The observation reported in Fig. 9C indicates that, as for *h-ZO-1, MDCK-ZO-1* mRNA levels are profoundly influenced by calcium. This may be due to transcription turnoff or to an increase in hydrolysis. The fact that the significant difference between the levels of mRNA in NC- and LC-cultured cells is not reflected at the protein level suggests that, in the absence of Ca^{2+}, either this abundant *ZO-1* mRNA is not translated into ZO-1 protein or the protein produced would have a shorter half life.

As stated in the Introduction, the TJ, which was born as an obscure seal at the outermost end of the intercellular space, is now regarded as a highly complex structure that, in addition to regulating epithelial permeability, plays a role in other fundamental cell functions, such as proliferation, tumor suppression, and differentiation. In keeping with this variety of functions, the analysis of ZO-1 performed in this work shows that this molecule, in addition to having several specific domains characteristic of the MAGUK family, may be expressed in several splicing forms, whose biological functions remain unknown. This variation of ZO-1, together with changes in other molecules of the junctional complex, might provide the TJ with the versatility it needs to cope with the fundamental functions mentioned above, as well as with the ever-changing environment to which epithelial cells are constantly subjected.

We thank Dr. J. M. Anderson (Yale University) and Dr. Isaura Meza (CINVESTAV) for kindly providing clone B11 from *h-ZO-1* and a human actin probe, respectively. We also thank José Velazquez Cervantes and Porfirio Reyes Lopez from the Department of Physiology, Biophysics, and Neuroscience at CINVESTAV for their computer help, and the Department of Genetics and Molecular Biology of this Center for Generously allowing us the use of their computer facilities. This work was supported by research grants from the National Research Council of México (CONACYT).

REFERENCES

1. Anderson, J. M., Stevenson, B. R., Jesaitis, L. A., Goodenough, D. A., and Mooseker, M. (1998). Characterization of ZO-1, a protein component of the tight junction from mouse liver and Madin–Darby canine kidney cells. *J. Cell Biol.* **106,** 1141–1149.
2. Anderson, J. M., Van Itallie, C. M., Peterson, M. D., Stevenson, B. R., Carew, E. A., and Mooseker, M. S. (1989). ZO-1 mRNA and protein expression during tight junction assembly in Caco-2 cells. *J. Cell Biol.* **109,** 1047–1056.
3. Anderson, J. M., and Van Itallie, C. M. (1995). Tight junctions and the molecular basis for regulation of paracellular permeability. *Am. J. Physiol.* **269,** G467–G475.
4. Balda, M. S., Gonzalez-Mariscal, L., Contreras, R. G., and Cereijido, M. (1991). Assembly and sealing of tight junctions: Possible participation of G-proteins, phospolipase C, protein kinase C and calmodulin. *J. Membr. Biol.* **122,** 193–202.
5. Balda, M. S., and Anderson, J. M. (1993). Two classes of tight junctions are revealed by ZO-1 isoforms. *Am. J. Physiol.* **264,** C918–C924.
6. Balda, M. S., Anderson, J. M., and Matter, K. (1996). The SH3 domain of the tight junction protein ZO-1 binds to a serine protein kinase that phosphorylates a region C-terminal to this domain. *FEBS Lett.* **399,** 326–332.
7. Baumgartner, S., Littleton, J. T., Broadie, K., Bhat, M. A., Harbecke, R., Lengyel, J. A., Chiquet-Ehrismann, R., Prokop, A., and Bellen, H. J. (1996). A Drosophila neurexin is required for septate junction and blood–nerve barrier formation and function. *Cell* **87,** 1059–1068.
8. Bradford, M. (1976). A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein dye binding. *Anal. Biochem.* **72,** 248.
9. Cereijido, M., Robbins, E. S., Dolan, W. J., Rotunno, C. A., and Sabatini, D. D. (1978). Polarized monolayers formed by epithelial cells on a permeable and translucent support. *J. Cell Biol.* **77,** 853–880.
10. Cereijido, M. (1992). Evolution of the ideas on the tight junction. *In* "Tight Junctions" (M. Cereijido, Ed.), pp. 1–13, CRC Press, Boca Raton, FL.

11. Chomczynski, P., and Sacchi, N. (1987). Single step method of RNA isolation by acid guanidium thiocyanate–phenol–chlorophorm extraction. *Anal. Biochem.* **162,** 156–159.
12. Contreras, R. G., Miller, J. H., Zamora, M., González-Mariscal, L., and Cereijido, M. (1992). Interaction of calcium with the plasma membrane of epithelial (MDCK) cells during junction formation. *Am. J. Physiol.* **263,** C313–C318.
13. Gonzalez-Mariscal, L., Chavez de Ramirez, B., and Cereijido, M. (1985). Tight junction formation in cultured epithelial cells. *Am. J. Physiol.* **259,** C978–C986.
14. Gonzalez-Mariscal, L., Contreras, R. G., Bolivar, J. J., Ponce, A., Chavez de Ramirez, B., and Cereijido, M. (1990). Role of calcium in tight junction formation between epithelial cells. *Am. J. Physiol.* **259,** C978–C986.
15. Gottardi, C. J., Arpin, M., Fanning, A., and Louvard, D. (1996). The junction-associated protein, zonula occludens-1, localizes to the nucleus before the maturation and during remodeling of cell–cell contacts. *Proc. Natl. Acad. Sci. USA* **93,** 10779–10784.
16. Harlow, E., and Lane, D. (1988). Immunizations. *In* "Antibodies: A Laboratory Manual," pp. 55–137, Cold Spring Harbor Laboratory Press, New York.
17. Haskins, J., Gu, L., Wittchen, E. S., Hibbard, J., and Stevenson, B. R. (1998). ZO-3, a novel member of the MAGUK protein family found at the tight junction, interacts with ZO-1 and occludin. *J. Cell Biol.* **141,** 199–208.
18. Hata, Y., Butz, S., and Sudhof, T. C. (1996). Cask: A novel dlg/PSD95 homologue with an N-terminal calmodulin dependent protein kinase domain identified by interaction with neurexins. *J. Neurosci.* **16,** 2488–2494.
19. Itoh, M., Nagafuchi, A., Yonemura, S., Kitani-Yasuda, T., Tsukita, S., and Tsukita, S. (1993). The 220-kD protein colocalizing with cadherins in non-epithelial cells is identical to ZO-1, a tight junction-associated protein in epithelial cells: cDNA cloning and immunoelectron microscopy. *J. Cell Biol.* **121,** 491–502.
20. Itoh, M., Nagafuchi, A., Moroi, S., and Tsukita, S. (1997). Involvement of ZO-1 in cadherin-based cell adhesion through its direct binding to α catenin and actin filaments. *J. Cell Biol.* **138,** 181–192.
21. Lahey, T., Gorozyca, M., Jia, X., and Budnik, V. (1994). The *Drosophila* tumor supressor gene *dlg* is required for normal synaptic bouton structure. *Neuron* **13,** 823–835.
22. Lue, R. A., Marfatia, S. M., Branton, D., and Chishti, A. H. (1994). Cloning and characterization of hdlg: The human homologue of the *Drosophila* discs large tumor suppressor binds to protein 4.1. *Proc. Natl. Acad. Sci. USA* **91,** 9818–9822.
23. Marfatia, S. M., Lue, R. A., Branton, D., and Chishti, A. H. (1994). In vitro binding studies suggest a membrane associated complex between erythroid p55, protein 4.1 and glycophorin C. *J. Biol. Chem.* **269,** 8631–8634.
24. Marfatia, S. M., Morais Cabral, J. H., Lin, L., Hough, C., Bryant, P. J., Stolz, L., and Chishti, A. H. (1996). Modular organization of the PDZ domains in the human discs-large protein suggests a mechanism for coupling PDZ domain-binding proteins to ATP and the membrane cytoskeleton. *J. Cell Biol.* **135,** 753–766.
25. Merrifield, R. B. (1963). Solid phase peptide synthesis: The synthesis of a tetrapeptide. *J. Am. Chem. Soc.* **85,** 2149–2154.
26. Meza, I., Ibarra, G., Sabanero, M., Martinez-Palomo, A., and Cereijido, M. (1980). Occluding junctions and cytoskeletal components in a cultured transporting epithelium. *J. Cell Biol.* **87,** 746–754.
27. Nakai, K., and Kanehisa, M. (1992). A knowledge base for predicting protein localization sites in eucaryotic cells. *Genomics* **14,** 897–911.
28. Sakakibara, A., Furuse, M., Saitou, M., Ando-Akatsuka, Y., and Tsukita, S. (1997). Possible involvement of phosphorylation of occludin in tight junction formation. *J. Cell Biol.* **137,** 1393–1401.
29. Sambrook, J., Fritsch, E. F., and Maniatis, T. (1989). "Molecular Cloning: A Laboratory Manual," 2nd ed., Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.
30. Schneeberger, E. E., and Lynch, R. D. (1994). Tight junctions: Their modulation under physiological and pathological states. *In* "Molecular Mechanisms of Epithelial Cell Junctions: From Development to Disease" (S. Citi, Ed.), pp. 123–140, R. G. Landes, Austin, TX.
31. Singer, K., Stevenson, B. R., Woo, P. L., and Firestone, G. L. (1994). Relationship of serine/threonine phosphorylation/dephosphorylation signaling to glucocorticoid regulation of tight junction permeability and ZO1 distribution in nontransformed mammary epithelial cells. *J. Biol. Chem.* **269,** 16108–16115.
32. Songyang, Z., Fanning, A. S., Fu, C., Xu, J., Marfatia, S. M., Chishti, A. H., Crompton, A., Chan, A. C., Anderson, J. M., and Cantley, L. C. (1997). Recognition of unique carboxyl-terminal motifs by distinct PDZ domains. *Science* **275,** 73–77.
33. Swofford, D. (1992). PAUP: Phylogenetic Analysis Using Parsimony, version 3.1. Computer program and manual distributed by the Center of Biodiversity, Illinois Natural History Survey, Champaign, IL 61820.
34. Tam, J. P. (1988). Synthetic peptides vaccine design: Synthesis and properties of a high density multiple antigenic peptide system. *Proc. Natl. Acad. Sci. USA* **85,** 5409–5413.
35. Van Itallie, C. M., Balda, M. S., and Anderson, J. (1995). Epidermal growth factor induces tyrosine phosphorylation and reorganization of the tight junction protein ZO1 in A431 cells. *J. Cell Sci.* **108,** 1735–1742.
36. Wilde, C. D., Crowther, C. E., Cripe, T. P., Lee, M. G., and Cowan, N. J. (1982). Evidence that a human β tubulin pseudogene is derived from its corresponding mRNAS. *Nature* **297,** 83–84.
37. Willott, E., Balda, M. S., Fanning, A. S., Jameson, B., Van Itallie, C., and Anderson, J. M. (1993). The tight junction protein ZO-1 is homologous to the *Drosophila* disc-large tumor suppressor protein of septate junctions. *Proc. Natl. Acad. Sci. USA* **90,** 7834–7838.
38. Zomerdijk, J. C. B. M., Ouellette, M., ten Asbroek, A. L. M. A., Kieft, R., Bommer, A., Clayton, C. A., and Borst, P. (1990). The promoter for a variant surface glycoprotein gene expression site in *Trypanosoma brucei*. *EMBO J.* **9,** 2791–2801.

Received December 14, 1998