The Tight Junction Protein ZO-2 Contains Three PDZ (PSD-95/Disks-Large/ZO-1) Domains and an Alternatively Spliced Region*  

(Received for publication, July 11, 1996, and in revised form, August 9, 1996)  

Martin Beatch‡§, Lynne A. Jesaitis§§, Warren J. Gallin¶, Daniel A. Goodenough¶, and Bruce R. Stevenson**  

From the ‡Department of Anatomy and Cell Biology and §Department of Biological Sciences, University of Alberta, Edmonton, Alberta T6G 2H7, Canada and the ¶Department of Cell Biology, Harvard Medical School, Boston, Massachusetts 02115

The complete cDNA sequence for canine ZO-2, a tight junction-specific protein, is presented. A single open reading frame encodes a polypeptide of 1,174 amino acids with a predicted molecular mass of 132,085 daltons. As noted previously (1), ZO-2 is a member of the membrane-associated guanylate kinase-containing (MAGUK) protein family, a family which includes an additional tight junction-associated protein, ZO-1. These proteins contain a region homologous to guanylate kinase, an SH3 domain, and variable numbers of PSD-95/Disks-Large/ZO-1 (PDZ) domains, shown to be involved in protein-protein interactions. ZO-2 and ZO-1 contain three PDZ domains in the N-terminal half of the molecule. Between the first and second PDZ domains, ZO-2 displays a basic region (pI = 10.27) containing 22% arginine residues. Both ZO-1 and ZO-2 have proline-rich C-terminal regions that are not homologous to other MAGUK family members. Sequence analysis of multiple ZO-2 cDNAs reveals a 36-amino acid domain in this C-terminal region present in only some of the cDNAs. Overall, ZO-2 is highly homologous to ZO-1, showing 51% amino acid identity; however, the C-terminal ends of the molecules show only 25% amino acid identity. This suggests that the C-terminal ends of ZO-1 and ZO-2 have different functions.

Several molecular constituents of the tight junction have been identified. These include ZO-1 (2), cingulin (3), ZO-2 (1, 4), 7H6 (5), Rab3B (6), and occludin (7). The presence of other proteins at the tight junction, including spectrin (8, 9), another Rab protein (Rab13) (10), and p130 (11), has been implicated but not confirmed.

* This work was supported in part by operating grants from the Medical Research Council of Canada and the Kidney Foundation of Canada (to B. R. S.) and National Institutes of Health Grants GM28932 and GM18974 (to D. A. 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) L27152.

** Senior Scholar of the Alberta Heritage Foundation for Medical Research. To whom correspondence should be addressed. Tel.: 403-492-1841; Fax: 403-492-0450; E-mail: brus@anat.med.ualberta.ca.

The complete mouse (12) and human (13) protein and cDNA sequences of ZO-1 have been published. A partial sequence for ZO-2 has also been characterized (1). Analyses of these sequences indicate that ZO-1 and ZO-2 share significant homology with each other and with several other proteins, including the lethal(1)disks-large-1 (dlg) tumor suppressor gene product (dlg-A) of Drosophila (14), erythrocyte membrane-associated p55 (15), and PSD-95/SAP90, a protein found at brain presynaptic membranes (16, 17). These proteins share several conserved regions, including a region homologous to guanylate kinase (GUK),1 an enzyme which converts GMP to GDP, a single src homology (SH3) domain, hypothesized to be involved in protein-protein interactions necessary for signal transduction (18, 19), and a variable number of N-terminal repeats termed PDZ domains (from PSD-95, disks-large, ZO-1), shown to bind integral plasma membrane proteins (20–23). As all proteins in this group are associated with the plasma membrane, they have collectively been termed the MAGUK family (membrane-associated guanylate kinase-containing) (24).

ZO-1 and ZO-2 are different from other family members but similar to each other in that they display C-terminal acidic and proline-rich regions (1, 13). ZO-1 also contains an alternatively spliced region, termed the α motif, in the C-terminal region (25, 26). The exact function of this splice domain remains unknown. Here we present the complete sequence encoding ZO-2. We find that ZO-2, like ZO-1, contains three N-terminal PDZ domains. ZO-2 displays a basic (pI = 10.27), arginine-rich domain between PDZ1 and PDZ2, similar to a previously unnoted option in ZO-1. Finally, ZO-2 also contains an alternatively spliced region near the C terminus which we call the β motif. Amino acid (aa) sequence comparison between ZO-2 and ZO-1 indicates a high degree of homology, although divergence at the C termini of the molecules suggests these proteins will have functional differences.

EXPERIMENTAL PROCEDURES

Library Screening—To extend the cDNA sequence for ZO-2, a hybridization probe corresponding to the 5' most of the partial ZO-2 cDNA (1) was generated by polymerase chain reaction using oligonucleotides 5'-GCA CGG AGA CCG-3' and 5'-TTC ATC TTC AGG ACT-3'. The resulting 255-bp product was used to rescreen an oriented oligo(dt) primed Madin-Darby canine kidney (MDCK) cell cDNA library in Uni-ZAP XR (kindly provided by Mariano Zerial). Of ~500,000 plaques analyzed, 10 ZO-2 cDNAs were isolated, the longest of which (pM7.7) was 4.1 kb. Analysis of the 4.1-kb cDNA sequence and the size of mature ZO-2 transcript in Northern blots (about 5.2 kb) indicated that the 4.1-kb clone was not full-length. We then generated an oriented MDCK cDNA library specific for ZO-2 in Uni-Zap XR (Stratagenec). MDCK poly(A)+ RNA was reverse-transcribed with the ZO-2-specific antisense linker-primer 5'-GAG AGA GAG AGA GAG AGA GAA CTA GTC TCG AGG TAG TCA TCG TCA TGG TCC-3' containing an XhoI site. This primer was situated approximately 250 bp downstream of the 5' end of the known ZO-2 sequence contained in pM7.7. The cDNA was then ligated with EcoRI adapters, digested with XhoI, size-fractionated, and ligated into the Uni-ZAP XR vector. The new library was screened with an oligonucleotide containing the sequence 5'-TGG CTG CTG CCG CTC CTC TCG CTG TCG TAC CCG CTC GCG GAA CTT CTT CT-3' end-labeled with isotope (γ[32P]ATP (ICN Pharmaceuticals). This sequence is the complement of a known region of the human ZO-2 sequence.

1 The abbreviations used are: GUK, guanylate kinase; α, amino acid(α); β, base pair(s); k, kilobase(s); MDCK, Madin-Darby canine kidney; RT-PCR, reverse transcriptase-polymerase chain reaction; PAGE, polyacrylamide gel electrophoresis; nNOS, neuronal nitric-oxide synthase.
ZO-2 located 5’ of the primer used to create the library. Approximately 200–300 plaques were sequenced, and positive clones were plaque-purified.

Inserts were sized by restriction digestion, and the 3’ 200–300 bp were sequenced to positively identify five ZO-2 clones among 28 primary picks. The longest clone was chosen for double-stranded sequencing (Sequenase, United States Biochemical, or Amplicycle, Perkin Elmer).

**Sequence Analysis—**Nucleotide sequences, isoelectric points of various domains, and the overall molecular weight of the protein were analyzed using the University of Wisconsin Genetics Computer Group (GCG) software package. Amino acid sequence comparisons to determine the percent identity and the percent similarity (% identity + % conserved substitutions) were done using the default parameters in the GCG GAP alignment program based on the algorithm of Needleman and Wunsch (27). The amino acid numbers of the compared sequences are shown in Table I.

**Splice Domain Analysis—**Total RNA from MDCK cells was isolated using Trizol (Life Technologies, Inc.) and subjected to reverse transcription by avian myeloblastosis virus reverse transcriptase (Promega). The

| Table I | Amino acid numbers of compared sequences |
|---------|----------------------------------------|
| Canine ZO-2 | Human ZO-2 | Human ZO-1 |
| PDZ1 10–96 | 33–119 | 11–97 |
| PDZ2 291–368 | 307–384 | 174–251 |
| PDZ3 496–574 | 512–590 | 412–490 |
| SH3 592–650 | 608–666 | 508–569 |
| GUK 714–863 | 730–879 | 632–782 |
| Acid 883–947 | 899–963 | 806–876 |
| Post-acid 948–end | 964–end | 877–end |
| Splice 1008–1043 | 1024–1059 | 910–989 |

**RESULTS**

The full-length nucleotide and aa sequence of canine ZO-2 is shown in Fig. 1. This sequence contains a single open reading frame coding for a 1,174-aa polypeptide with a predicted molecular mass of 132,085 daltons, considerably less than the mass of 160 kDa determined by SDS-PAGE (1). The disparity between the predicted molecular mass and that determined by SDS-PAGE could be due to anomalous electrophoretic migration or to post-translational modifications. A full-length sequence for human gene X104 related to the Friedreich ataxia locus was previously published, but no characterization of the gene product was presented (28). A comparison of the X104 aa sequence to canine ZO-2 indicates a high degree of similarity: the canine and human sequences are 87% identical at the aa level (Table II). The similarity is even greater within the characteristic MAGUK and acidic domains, with 92–100% aa identity. This indicates that the X104 gene product is the human homolog of canine ZO-2.

As noted previously (1), ZO-2 is highly homologous to ZO-1, showing an overall aa sequence identity of 51% (Table III). However, as in the comparison of canine and human ZO-2, the similarities between ZO-2 and ZO-1 are greater in the PDZ, SH3, and GUK domains. In addition, both ZO-2 and ZO-1
display acidic regions at the C-terminal end of the GUK domain followed by proline-rich regions. Interestingly, while the acidic domains of ZO-2 and ZO-1 show significant similarity (60% aa identity), the proline-rich, C-terminal regions following the acidic domains are significantly less similar (25% aa identity).

Fig. 1 indicates there are three potential in-frame start codons for the canine ZO-2 sequence. Several criteria indicate that the third (nucleotide 398) site is used as the start site. First, the nucleotide sequence surrounding the third site most closely matches the eukaryotic initiation consensus sequence (Fig. 2) (29). In addition, the 100 aa following this start site closely match both the initial aa sequence of human ZO-1 (70% identity, 80% similarity) and the 100 aa following a methionine at position 24 of the human ZO-2 sequence (99% identity, 100% similarity) (28). Finally, the aa sequence generated from upstream of this third canine ZO-2 initiation codon shows no significant homology with corresponding sequences upstream of the human ZO-1 start site or methionine 24 of human ZO-2. The canine ZO-2 sequence shows no initiation codon corresponding to that published for human X104.

Fig. 1 indicates the relative positions of three N-terminal PDZ repeats, the SH3 and GUK domains, and an acidic region. This arrangement of domains is similar to that found in ZO-1 (13). We further note that the region between PDZ1 and PDZ2 of canine ZO-2 is rich in arginine residues (22%) and is highly basic, with a pi of 10.27. Although not previously documented, a similar basic region exists in the corresponding region of ZO-1 (12, 13). Fig. 1 also illustrates a 36-aa domain encoded by a 108-bp sequence present in 6 of 10 partial cDNAs from the C-terminal region of canine ZO-2. To confirm that the presence or absence of this sequence resulted from alternatively spliced mRNAs, we designed primers which straddled the 108-bp region and performed RT-PCR using total RNA isolated from MDCK cells. This analysis resulted in two distinct PCR products which ran at the expected sizes (413 and 305 bp; Fig. 3), confirming the presence of at least two unique species of RNA for canine ZO-2. We propose that this alternatively spliced region be called the β motif. It was previously demonstrated that ZO-1 contains an alternatively spliced 80-aa region termed the α motif (25, 26). The ZO-1 α motif is larger than that of ZO-2, and an aa comparison of the ZO-1 α and ZO-2 β motifs shows little similarity (17% aa identity; Table III). The similarity between the canine and human ZO-2 β domains (72% aa identity) is lower than that found over the entire length of these molecules (87% aa identity), but similar to that found over the postacidic, C-terminal domains (68% aa identity; Table II).

**DISCUSSION**

The full-length sequence for canine ZO-2 further confirms that this protein is a member of the MAGUK family. Moreover, ZO-2 shows the same distribution of MAGUK domains as ZO-1, and both ZO-1 and ZO-2 contain unique domains not found in other MAGUK family members, including acidic and proline-rich aa sequences (Fig. 4). This indicates that the tight junction contains at least two proteins with a high degree of similarity.

A comparison of the canine ZO-2 aa sequence presented here with that from a previously published human gene (X104) related to Friedreich ataxia (28) indicates that these proteins are homologous. However, several differences in the sequences of the proteins from these two species are apparent. The canine cDNA sequence contains three potential, in-frame start codons. We believe the third ATG is the most likely initiation site based on 1) the highest degree of similarity to the eukaryotic initiation consensus sequence (Fig. 2) (29) and 2) a comparison of the
ZO-2 Complete Sequence: PDZ Domains and Alternative Splice

A comparison of the amino acid sequences of canine ZO-2, human ZO-2, and human ZO-1 shows a high degree of similarity between these proteins (Tables II and III). The similarities are particularly high in the characteristic MAGUK domains, suggesting that these domains are functionally important. Of all the MAGUK protein domains, data demonstrating functional relevance have been provided for only the PDZ region. PDZ domains in PSD-95/SAP90 have been shown to bind to the N-methyl-D-aspartate receptor (20), Shaker-type potassium channels (21) and neuronal type nitric-oxide synthase (nNOS) (22). The PSD-95/SAP90-nNOS interaction appears to be mediated by PDZ domains in each molecule (22). Likewise, PDZ-DVDZ domain interactions are implicated in the binding of nNOS to syntrophin, part of the dystrophin protein complex in skeletal muscle (22). The aa sequences surrounding the start sites of canine and human ZO-2 and human ZO-1. Surprisingly, we found no match for the published start site of human ZO-2 in our canine sequence. In addition, the canine ZO-2 is 69 aa longer than human ZO-2 based on a best-fit amino acid sequence comparison (GCG GAP), and the 35 aa immediately preceding the termination of human ZO-2 show a lower aa similarity (13%) to the corresponding region of canine ZO-2 than is present over the rest of the molecules. Combined, these results suggest species-dependent divergence of the canine and human ZO-2 sequences.

Acknowledgments—We thank Julie Haskins, David Paul, and Dwayne Weber for their friendship and technical support. We are also grateful to Marino Zerial for the MDCK library.

REFERENCES

1. Jesaitis, L. A., and Goodenough, D. A. (1994) J. Cell Biol. 123, 949–961
2. Stevenson, B. R., Siliciano, J. D., Mooseker, M. S., and Goodenough, D. A. (1986) J. Cell Biol. 103, 755–766
3. Citi, S., Sahay, H., Jakes, R., Geiger, B., and Kendirc-Jones, J. (1988) J. Cell Biol. 103, 272–276
4. Gumbiner, B. M., Lowenkopf, T., and Aparicio, D. A. (1991) Proc. Natl. Acad. Sci. U. S. A. 88, 3460–3464
5. Zhang, Y., Saito, T., Misawa, N., Kurnosov, Y., and Mori, M. (1993) J. Cell Biol. 120, 477–483
6. Weber, E., Berta, G., Tousson, A., St. John, P., Green, M. W., Gopalakrishnan, U., Jilling, T., Sturcher, E. J., Elton, T. S., Abrahamson, D. R., and Kirk, K. L. (1994) J. Cell Biol. 125, 583–594
7. Furuse, M., Hirase, T., Itoh, M., Nagafuchi, A., Yonemura, S., Tsukita, S., and Tsukita, S. (1993) J. Cell Biol. 123, 1777–1788
8. Itoh, M., Yonemura, S., Nagafuchi, A., Tsukita, S., and Tsukita, S. (1991) J. Cell Biol. 115, 1449–1462
9. Furuse, M., Itoh, M., Hirase, T., Nagafuchi, A., Yonemura, S., Tsukita, S., and Tsukita, S. (1993) J. Cell Biol. 123, 1777–1788
10. Balda, M. S., and Anderson, J. M. (1993) J. Cell Biol. 123, 583–594
11. Balda, M. S., Gonzalez-Marsical, L., Matter, K., Cereijido, M., and Anderson, J. M. (1993) J. Cell Biol. 123, 283–302
12. Itoh, M., Nagafuchi, A., Yonemura, S., Kitani-Yasuda, T., Tsukita, S., and Tsukita, S. (1993) J. Cell Biol. 121, 491–502
13. Willett, E., Balda, M. S., Fanning, A. S., Jameson, B., Van Itallie, C., and Anderson, M. J. (1993) Proc. Natl. Acad. Sci. U. S. A. 90, 7833–7838
14. Woods, D. F., and Bryant, P. J. (1991) Cell 66, 451–464
15. Ruff, P., Speicher, D. W., and Hua-nis-Chishiti, A. (1993) Proc. Natl. Acad. Sci. U. S. A. 90, 6595–6599
16. Cho, K. O. Hunt, C. A., and Kennedy, M. B. (1992) Neuron 9, 929–942
17. Kistner, U., Wenzel, B. M., Veh, R. W., Cases-Langhoff, C., Garner, A. M., Apfalterer, U., Gruen, B., Gundelfinger, E. C., and Garner, C. C. (1993) J. Biol. Chem. 268, 4580–4583
18. Booker, G. W., Gout, I., Downing, A. K., Driscoll, P. C., Boyd, J., Waterfield, M. D., and Campbell, A. D. (1993) Cell 73, 813–822
19. Mayer, B. J., and Baltimore, D. (1993) Trends Cell Biol. 3, 8–13
20. Kornau, H. C., Schenker, L. T., Kennedy, M. B., and Seeburg, P. H. (1995) Science 269, 1737–1740
21. Kim, E., Nielsenhammer, M., Rothschuld, A., Jan, Y. N., and Sheng, M. (1995) Nature 378, 88–88
22. Brennan, J. E., Chao, D. S., Gee, S. H., McGee, A. W., Craven, S. E., Santillano, D. R., Wu, Z., Huang, F., Xia, H., Peters, M. F., Freuchter, S. C., and Breit, D. W. (1996) Cell 84, 757–767
23. Gompertz, S. N. (1996) Cell 84, 659–662
24. Kim, S. K. (1995) Curr. Opin. Cell Biol. 7, 641–649
25. Kuribara, H., Anderson, J. M., and Farquhar, M. G. (1992) Proc. Natl. Acad. Sci. U. S. A. 89, 7075–7079
26. Balda, M. S., and Anderson, J. M. (1993) Am. J. Physiol. 264, C198–C203
27. Needleman, S. B., and Wunsch, C. D. (1970) J. Mol. Biol. 48, 443–453
28. Dulcic, F., Rodius, F., Mandel, J.-L., and Koenig, M. (1994) Hum. Mol. Genet. 3, 909–914
29. Kosak, M. (1987) Nucleic Acids Res. 15, 8125–8148
The Tight Junction Protein ZO-2 Contains Three PDZ (PSD-95/Discs-Large/ZO-1) Domains and an Alternatively Spliced Region
Martin Beatch, Lynne A. Jesaitis, Warren J. Gallin, Daniel A. Goodenough and Bruce R. Stevenson

J. Biol. Chem. 1996, 271:25723-25726.
doi: 10.1074/jbc.271.42.25723

Access the most updated version of this article at http://www.jbc.org/content/271/42/25723

Alerts:
• When this article is cited
• When a correction for this article is posted

Click here to choose from all of JBC's e-mail alerts

This article cites 0 references, 0 of which can be accessed free at http://www.jbc.org/content/271/42/25723.full.html#ref-list-1