Occludin is an essential membrane protein component of cellular tight junctions, participating in both cell-cell adhesion in the paracellular space and anchoring of the junctional complex to the cytoskeleton. The latter function is accomplished through binding of the C-terminal cytoplasmic region to scaffolding proteins that mediate binding to cytoskeletal actin. We isolated a structural domain from both the bacterial-expressed C-terminal cytoplasmic region of human occludin and native cellular occludin, extracted from epithelial (Madin-Darby canine kidney) or endothelial (human brain) cells, by limited proteolysis with trypsin. This human occludin domain contains the last 119 amino acids as identified by N-terminal sequencing and peptide mass fingerprinting using matrix-assisted laser desorption ionization-time of flight mass spectrometry. Based on the sequence and secondary structure prediction, this domain contains 4 of 5 α-helices in the C-terminal region and is linked to the fourth membrane-spanning region by a loosely structured tethering polypeptide. Comparison of circular dichroism spectra of recombinant proteins corresponding to the entire C-terminal region versus only the binding domain region also supports the interpretation that the helical structural elements are concentrated in that domain. Co-immunoprecipitation of this domain with ZO-2 demonstrated preservation of the specificity of the scaffolding protein-binding function, and binding studies with immobilized ZO-2 suggest the presence of multiple ZO-2 binding sites in this domain. These results provide a basis for development of a structural model of the ZO-binding site that can be used to investigate regulation of tight junction anchoring by intracellular signaling events.

Tight junctions (TJ) play a key role in limiting the movement of water, ions, and macromolecules across both epithelial and endothelial surfaces (1–3). Occludin, the first membrane protein isolated from TJ complexes, is a ~60 kDa protein predicted to have four transmembrane regions, two extracellular loops, and N- and C-terminal cytoplasmic regions (4). Several members of the claudin family of membrane proteins were later isolated from TJ (5–10) and are thought to interdigitate with each other and occludin in linear strands to form the TJ complexes by sealing with the complementary strands on the apical surfaces of adjacent cells (7, 11). Both occludin and the TJ-associated claudins also bind to the cytoplasmic scaffolding proteins, ZO-1, -2, and -3, which serve to link the TJ complexes to the cell cytoskeleton (12–15). These scaffolding proteins are members of the MAGUK (membrane-associated guanylate kinase) family of proteins, which contain variable numbers of PSD-95/discs-large/ZO-1 (PDZ) domains, a Src homology (SH3) domain, and a domain homologous to guanylate kinase (GUK) in the same sequential order (16). Occludin binds to the GUK domain of each of the ZO proteins, while the claudins bind to the first PDZ (PDZ-1) domain (12, 14). The identification of the GUK domains of these three proteins is based on the amino acid comparison with GUK-containing proteins from the international protein data base (13, 17). The sequence of the GUK domain in each of these proteins is different, but there is high homology between them.

The binding between occludin and the ZO proteins is determined by the structural formation of the C-terminal cytoplasmic region. It is highly likely that this binding occurs in a compactly folded domain within this region; however, previous studies have not elucidated the precise identity of such a domain or the manner of its linkage to the rest of the occludin molecule. It is also not determined whether the binding affinity of ZO-1 with occludin differs from ZO-2 or ZO-3. Furuse et al. (18) have shown that truncation of C-terminal segments of occludin weakens or eliminates the binding to ZO-1, while deletion of short segments near the N terminus of the C-terminal cytoplasmic region does not alter this binding function. Nusrat et al. (19) studied a synthetic peptide, which they described as a “coiled-coil domain,” corresponding to amino acids 440–469 of the C-terminal region of human occludin. This peptide was shown to bind to native ZO-1 from T84 epithelial cells. While truncation studies can localize the functional elements with regard to the primary sequence, and binding studies using synthetic peptides can identify a specific region primarily responsible for recognition, neither provides complete information regarding the structural motif responsible for binding the GUK domain of the scaffolding proteins or the identity of potential regulatory elements within the binding domain of occludin. In order to investigate both binding and regulation, the full domain of occludin that interacts with the GUK domain needs to be identified and isolated.

Limited proteolysis, on the other hand, allows intact compact...
structural domains to be separated from more loosely coiled regions of the protein sequence, largely preserving the natural domain structure (20, 21). In the study described in this report, we used this method to isolate a structural domain from the region of occludin known to bind the ZO proteins, then tested this domain for retention of the binding function using co-immunoprecipitation and estimated the stoichiometry and dissociation constant using a binding technique. This method has the advantage of preserving the natural structure of occludin, either expressed in bacteria or extracted from cells, and isolating the functionally designated N-terminal domain of minimal size to facilitate further structural investigation. N-terminal sequencing and peptide mass fingerprinting (22) were employed to identify the precise sequence of the domain.

**EXPERIMENTAL PROCEDURES**

**Cell Culture**—MDCK (ATCC, Manassas, VA) cells were cultured in medium containing MEM (minimum essential medium) supplemented with non-essential amino acids, 15 mM of HEPEs and 10% fetal bovine serum. Cultures were maintained in a humidified CO2 incubator (5% CO2/95% air) at 37°C.

**Expression Plasmid Construction**—Methods followed the procedures of Sambrook et al. (23) with minor modifications. Two expression plasmids (for C-Occ and DN-Occ, defined below) were constructed by cloning the PCR-amplified occludin cDNA fragment into the expression vector pRSETA (Invitrogen; Carlsbad, CA). The template plasmid pPSK-hoc, a generous gift from Drs. C. M. Van Itallie and James M. Anderson (24) contains the entire coding region for human occludin (GenBank™ accession number U53823). PCR was performed to generate cDNA for the entire C-terminal cytoplasmic region of occludin (C-Occ) using two primers, 5′-TGCTGGATCCACTGGAAGAAGATG-3′ and 5′-CACTTGGCATCAAGATTCTAG-3′, and to generate a 121 amino acid deletion in the C-Occ cDNA between the His tag and the last four predicted helices (resistant protein) designated (DN-Occ) using the upstream primer, 5′-CAAAGGGGATCCCGAAGG-3′ and the same downstream primer as for C-Occ. The amplification cycle was programmed to denature the template at 95°C for 1 min, anneal the primer to the template at 45°C for 1 min and synthesize the cDNA at 68°C for 1 min for a total of 30 cycles. The PCR fragment and the cloning vector pRSETA were digested with restriction enzymes, BamHI and EcoRI, and then the digested fragments were separated from undigested fragments by electrophoresis. The digested fragments were purified by electrophoresis using dialysis tubing with 12-kDa molecular mass cutoff (Spectrum Medical Industries, Houston, TX). After phenol/chloroform extraction and ethanol precipitation, these two purified fragments were ligated together at 16°C overnight. After cloning the PCR product into pRSETA, the insert was sequenced from both directions to confirm its identity using T7 forward primer and pRSET reverse primer.

**His-tagged Protein Expression and Purification**—Plasmid with cDNA insert was transformed into BL21 DE3 bacteria, and expression of the His-tagged proteins (C-Occ or DN-Occ) was stimulated by 0.1 mM of isopropyl-1-thio-β-D-galactopyranoside. Four hours after induction, bacteria were collected by centrifugation at 8630g for 5 min at 4°C, and then lysed in high salt sodium phosphate buffer (20 mM sodium phosphate and 500 mM sodium chloride, pH 7.4) containing 100 µg/ml of lysozyme and 1× protease inhibitor mixture (Roche Applied Science). Protease inhibitor mixture was added as 100 µl per ml of solution of 10× inhibitor mixture prepared according to manufacturer’s directions in high salt sodium phosphate buffer. Total soluble proteins were collected after centrifugation at 31,000 × g for 30 min at 4°C, and then incubated with nickel-charged ProBond resin (Invitrogen) for 30 min. The protein was then eluted with washing (150–400 mM) concentrations of imidazole from the ProBond column after washing with high salt sodium phosphate buffer containing 50–100 mM imidazole to remove other bacterial proteins nonspecifically bound to the nickel resin. The purified proteins were analyzed for purity using SDS-PAGE and Western blotting with occludin antibody (Zymed Laboratories Inc., South San Francisco, CA). Concentration of the resultant protein solution was ~100 µg/ml. Full-length canine ZO-2 was expressed in SF9 insect cells (a gift from Dr. Bruce R. Stevenson) using the Baculovirus expression system (15). SF9 cells (1 × 106 cells/ml) were infected with ZO-2 cDNA containing baculovirus for 48 h, and then cells were harvested by centrifugation at 437 × g for 5 min. Procedures used to purify the His-tagged ZO-2 protein are the same as those for His-tagged C-Occ.

**SDS-PAGE and Western Blotting**—For Western blotting analysis, proteins were separated by SDS-PAGE (10% for the in vitro binding assay, and 12.5% or 15% for limited proteolysis and co-immunoprecipitation), and transferred onto a polyvinylidene difluoride membrane (Millipore, Billerica, MA). The membrane was incubated with blocking solution containing 0.1% Tween 20, 5% nonfat milk, and 2.5% goat serum in phosphate-buffered saline for 1 h. The membrane was then transferred into fresh blocking solution containing primary antibody (1 µg/ml) to occludin or ZO-2, and incubated for 1 h. Unbound and non-specifically bound antibody was removed by three changes of washing solution (0.1% Tween 20 in phosphate-buffered saline) over a period of 15 min. Biotinylated secondary antibody (1:1000 dilution) was then added and incubated with the membrane for 30 min in washing solution with 5% nonfat milk added. After washing, the membrane was incubated with either alkaline phosphatase- or horseradish peroxidase-conjugated streptavidin (1:1000 dilution) for 15 min. All incubations were performed at room temperature with gentle rocking. The immunoreactive protein bands were visualized by incubating the membrane with alkaline phosphatase substrate, BCIP/NBT (Zymed Laboratories Inc.), or horseradish peroxidase substrate, stable DAB (Research Genetics Inc.; Huntsville, AL), respectively.

**Limited Proteolysis**—Limited proteolysis was performed at room temperature using C-Ooc prior to elution from the nickel resin. An aliquot of 50 µl of C-Occ (~5 µg) bound to the resin was used for each reaction. Trypsin (Sigma, cat. T-8128, EC 3.4.24.11) in a low range protease derived from Streptomyces griseus (Sigma, cat. P-8811, EC 3.4.24.31) was added to the solution to a final enzyme concentration of 0.5, 5, 50, or 500 µg/ml, and incubated for 5 min. The reactions were stopped by adding 50 µl of solution containing 2× SDS-PAGE buffer and 1× of diethiothreitol prior to electrophoresis. After SDS-PAGE and Western blotting with occludin antibody, the protein masses were calculated by comparison to known protein markers (Sigma) using the computer program SigmaGel (Jandel Scientific Software, San Rafael, CA) (25).

To demonstrate that the digestion pattern of purified C-Ooc protein is identical to that of naturally occurring cellular occludin, MDCK cells (in the presence of 100 U/ml of human basic fibroblast growth factor) were lysed in high salt sodium phosphate buffer containing 1% Nonidet P-40 and 0.5% sodium deoxycholate without protease inhibitors. After sonication for 1 min on ice, the solution was diluted 1:4 with 20 mM sodium phosphate solution prior to trypsin digestion (20 µg/ml for 30 min at room temperature). Protease inhibitor mixture in low salt sodium phosphate buffer (20 mM sodium phosphate, 125 mM sodium chloride, pH 7.4) was added as described above to stop the reaction after the incubation period, and cell debris was removed by centrifugation at 31,000 × g for 15 min at 4°C. The digestion products were isolated by immunoprecipitation with a polyclonal occludin antibody (Zymed Laboratories Inc.; raised using a polysaccharide consisting of the last 50 amino acids of human brain microvascular endothelial cells) by SDS-PAGE and identified by Western blotting with the same occludin antibody.

**In Vitro Binding Assay**—The binding assay was performed before eluting the C-Occ protein from the ProBond nickel resin. MDCK cells were used as a source of native TJ scaffolding proteins because of their high level of expression. After binding of total solubil C-Occ protein from bacterial lysate to nickel resin and washing with high salt sodium phosphate buffer containing 100 mM imidazole, an aliquot of resin (100 µl) was transferred to a microcentrifuge tube and equilibrated with low salt sodium phosphate buffer. Tight junction proteins were extracted from confluent MDCK monolayers using high salt sodium phosphate buffer. Total soluble proteins were collected after centrifugation to separate the supernatant for electrophoresis. The His-tag region alone, expressed in bacteria containing the original pRSETA plasmid without the occludin insert, was used as a negative control for the binding assay.

**Immunoprecipitation (IP)**—MDCK cell lysate (as described above)
was diluted with 20 mM sodium phosphate to a final chloride concentration of 125 mM, and then incubated with 50 μl of protein G-conjugated agarose (Roche Applied Science) in the presence of 5 μg of antibody to either occludin, ZO-2 or GFAP (Zymed Laboratories Inc.) for 4 h at 4 °C. The antibody-bound molecules were then separated from unbound ones by centrifugation at 2940 × g for 10 s. The supernatant was used as a negative control in the binding assay, and the GFAP antibody (directed against an antigen not present in MDCK cells) was used as a non-relevant antibody control. The protein G-agarose was finally washed five times with 1 ml of buffer each time. It was followed by the addition of 100 μl of 1× SDS-PAGE loading buffer and boiling for 10 min to solubilize the attached proteins and centrifugation to remove the agarose. The supernatant was then analyzed by SDS-PAGE and Western blotting.

In the co-IP experiments, the purified C-Occ was treated with trypsin (5 μg/ml) for 1 h at room temperature, and then protease inhibitor mixture was added to inactivate trypsin. The resultant solution was incubated with diluted MDCK cell lysate for 2 h at 4 °C prior to immunoprecipitation with ZO-2 antibody according to the above protocol.

Amino Acid Sequencing—The purified C-Occ was digested with trypsin (50 μg/ml) for 5 min at room temperature, separated by 12.5% SDS-PAGE, transferred onto a polyvinylidene difluoride membrane, stained with Coomassie Brilliant Blue R-250 to identify the protein bands and then destained with a solution containing 30% methanol and 10% acetic acid in water. The outlined protein bands were cut out from the membrane and subjected to N-terminal amino acid sequencing by Edman degradation (26) using the 494/HT PROCISe Sequencing System (Applied Biosystems).

Peptide Mass Fingerprinting—After enzymatic digestion of C-Occ (~5 μg/50 μl) with trypsin (50 μg/ml for 5 min at room temperature), the peptide fragments were separated by SDS-PAGE, stained with Coomassie Blue, and each individual protein band was eluted and digested with chymotrypsin (20 μg/ml) at 37 °C for 8 h. The enzyme was then inactivated using trifuoroacetic acid to adjust the pH between 2 and 3. The peptide samples were cleaned using a reversed-phase ZipTip (Millipore) and eluted with 50% acetonitrile in 0.1% trifluoroacetic acid, mixed with the ionization matrix and spotted on specimen grids for mass spectrometry. Mass spectrometric analysis was performed using a Voyager-DE Elite MALDI-TOF mass spectrometer (Applied Biosystems, Foster City, CA), operated in the positive ion mode using angiotensin as the calibrant. Ionization of the peptide species under study was produced by pulses from a 337-nm nitrogen laser, and an accelerating voltage of 20 kV was used. The ionization matrix for this analysis was o-cyano-4-hydroxycinnamic acid dissolved in a 1:1 water/acetonitrile mixture. A minimum of 100 laser pulses was obtained, and the results were averaged to produce the resulting output spectrum. The spectra were identified by comparing the masses of the peaks with the fingerprint database generated from the mass calculation of each predicted chymotrypsin-digested fragment using the Peptide Mass program from ProSight (27).

Circular Dichroism (CD)—CD analysis of the purified His-tagged protein (20 mM sodium phosphate, 500 mM sodium chloride, pH 7.4) was performed using an Aviv 60DS spectropolarimeter (Aviv Associates, Lakewood, New Jersey). Measurements were made using a quartz microcell with a path length of 0.1 cm. Three scans from 185 to 260 nm were recorded at a speed of 0.5 nm/s, with a 0.5-nm interval. The scans were averaged and corrected by subtracting a buffer baseline (28). The relative percentages of secondary structures were determined using the CDFIT program (29).

Secondary Structural Prediction—Secondary structure prediction of the C-terminal cytoplasmic region of human occludin was performed using internet server-based programs, PHD (30), and Jpred (31).

Interaction between ZO-2 and dN-Occ—In each binding reaction, purified ZO-2 was immobilized on protein A/G-agarose using 2 μg of ZO-2 antibody, and then incubated with various concentrations of dN-Occ ranging from 1 to 100 μM. This range of dN-Occ concentration was used based on preliminary studies using a broader concentration range from 0.1 to 100 μM. Irrelevant antibody control to GFAP was used as a negative control to estimate nonspecific binding. The binding reactions were performed with gentle rocking at 4 °C for 3 h. Protein A/G-agarose with bound (ZO-2-dN-Occ) complexes was separated from free dN-Occ by centrifugation at 2940 × g for 2 min and washing the agarose pellet twice with low salt sodium phosphate binding buffer. The agrose-attached protein complexes was boiled for 10 min in 75 μl of 1× SDS loading buffer before separating the proteins by electrophoresis (7.5% gel for ZO-2 and 12.5% for dN-Occ). To visualize the specific protein bands, separated proteins were transferred onto a polyvinylidene difluoride membrane for Western blotting as described above. The intensity of each protein band was analyzed using SigmaGel software (Jandel Scientific Software, San Rafael, CA), and fitting of the binding data was performed using SigmaPlot (SPSS Inc., Chicago, IL) to estimate the association constants and stoichiometry, using the model shown by Equation 1,

\[ Z + O_n \rightleftharpoons Z(O_n) \]  
(Eq. 1)

and the association constant is defined by Equation 2.

\[ K_a = \frac{[Y][Z]}{[X]^n} \]  
(Eq. 2)

In Equation 2, \( K_a \) is the association constant, \([Y]\) is the concentration of bound dN-Occ, \([X]\) is the concentration of free dN-Occ, \([Z]\) is the immobilized ZO-2 concentration, which is constant in each reaction as measured by Western blotting, and \(n \) is the number of dN-Occ molecules binding to each molecule of ZO-2. Concentrations of bound dN-Occ were determined by interpolation of the band intensities in a standard curve constructed from the Western blot of a series of dN-Occ dilutions of known concentration.

RESULTS

Limited Proteolysis of the C-Occ Protein—A structural domain is, by definition, obtained as a stable proteolytic core, while interdomain sequences are generally susceptible to digestion by different proteases. Limited proteolysis can therefore be used to obtain the shortest segments of occludin that contain compact structural domains. Two different enzyme preparations, trypsin (T), and protease, were used in these experiments. Trypsin cleaves peptides at specific amino acids (lysine and arginine). Protease used here is a crude preparation from S. griseus that has a broader range of peptide cleavage. The use of these two enzyme preparations provides two independent methods of limited proteolysis and several different types of protease activity to test for shorter functional fragments than those generated by trypsin proteolysis alone. Results of limited proteolysis of the purified C-Occ protein are shown in Fig. 1. When a low concentration of trypsin (0.5 μg/ml) was used to cleave C-Occ attached to the nickel resin, the product consisted of essentially the entire C-Occ peptide without smaller fragments. With a higher concentration (5 μg/ml), however, C-Occ was cleaved into many fragments, shown in Fig. 1 as a ladder pattern. A single 19.2-kDa band resulted at the highest concentration of trypsin used (500 μg/ml). The pattern of broad range protease digestion was similar.
Trypsin Digestion of Native Occludin—To test the hypothesis that the structural organization of the C-terminal cytoplasmic region of native occludin is similar to that of the bacterial expressed C-Occ, total cell lysates from cultured cells and human brain tissue were treated with trypsin (T), and the tryptic fragments were concentrated by immunoprecipitation. Because even severe lysis conditions are not sufficient to completely extract occludin from the membranes, total lysates including the membrane fractions were used in the trypsin digestion. The Western blot (Fig. 2.) shows that the two protein bands resulting from the digestion of lysates of both cultured MDCK cells (lane 3) and human brain tissue (lane 2) correspond to the 21.5- and 19.2-kDa tryptic fragments of C-Occ. This result demonstrates that native occludin extracted from both epithelial (MDCK) cells and brain endothelial cells shares a similar susceptibility to trypsin digestion with bacterial-expressed C-Occ protein, suggesting a similar structural organization.

Identification of the N- and C-Terminal Sequences of the Tryptic Fragments—To identify the N-terminal ends of the tryptic fragments, Edman degradation was used to sequence the SDS-PAGE separated protein bands. Using this method, the first five N-terminal amino acids of the 21.5- and 19.2-kDa bands were determined to be serine, lysine, arginine, threonine, and glutamic acid for the 21.5-kDa band, and threonine, glycine, glutamic acid, and serine for the 19.2-kDa band. They are at amino acid positions 390–394 (21.5 kDa) and 404–408 (19.2 kDa) of human occludin. The cleavage site of the 21.5-kDa band was specific for trypsin at the C-terminal end of arginine. However, the cleavage site of the 19.2-kDa band was at the C-terminal end of threonine, which is not a typical trypsin cleavage site (arginine or lysine), likely due to impurities in the trypsin enzyme preparation used, as discussed previously. This scenario was supported by the experiment in which sequencing grade trypsin (Ts) was used for limited proteolysis, resulting in only the 21.5-kDa band.

Peptide mass fingerprinting was used to identify the C-terminal fragments generated from chymotrypsin digestion of the purified 21.5 kDa and 19.2 kDa bands. At least twelve peaks from the MALDI-TOF mass spectra have masses compatible with fragments from chymotrypsin digestion of the 21.5 kDa band (Fig. 3a), and those include the fragment (VGDY-DRKQT) containing the C-terminal amino acid of C-Occ (Fig. 3b). A peak with virtually identical mass (Fig. 3c) as that in the 21.5-kDa band is also present in the chymotrypsin digestion spectrum of the 19.2-kDa band but is absent in the chymotrypsin control. These results demonstrate that the C-terminal amino acid of both tryptic fragments is the C-terminal threonine of occludin. Therefore, the shortest fragment resulting from limited proteolysis is composed of the last 19 amino acids of C-Occ.

Secondary Structure Prediction and CD Analysis—I pred analysis of C-Occ predicts five α-helices and one β-strand, located in the human occludin sequence at Arg268–Arg733 (helix 1), Ser425–Arg734 (helix 2), Leu440–Asp465 (helix 3), Glu473–Gln486 (helix 4), Lys496–Val514 (helix 5), and Glu529–Lys599 (β-strand). The PHD method resulted in a prediction of 34.9% α-helix, 0.7% β-strand, and 64.4% random coil.

Two plasmins were constructed to express His-tagged proteins representing (1) the entire C-terminal cytoplasmic region of occludin (C-Occ), as used previously, and (2) a shorter peptide (dN-Occ) obtained from C-Occ by deleting 121 amino acids between the His tag and the last four predicted helical regions (Fig. 4). These His-tagged proteins, expressed in BL21 DE3 bacteria and purified as described above, migrate on a 12.5% SDS-PAGE to ~45 kDa and 27 kDa according to protein markers, and have predicted masses of 34 and 20 kDa, respectively.

A summary of the secondary structure analysis using CD is shown in Fig. 5. Analysis of CD spectra using the CDFIT program yields 43% α-helix, 12% β-strand, and 45% random coil structure for the C-Occ protein, and 57% α-helix, 13% β-strand, and 30% random coil structure for the dN-Occ protein. The result for C-Occ from the CD analysis is consistent with the PHD prediction that this protein has a significantly higher percentage of α-helix than β-strand, although the actual numbers vary considerably between the two methods (PHD 35% versus CD 43% for α-helix and PHD 1% versus CD 12% for β-strand). Both predictive and experimental methods clearly demonstrate that when the relatively less structural region is removed, the percentage of α-helix increases and that for random coil decreases considerably. These results add further support to the hypothesis that a structural domain is localized in the C-terminal end of C-Occ, and this domain is mainly composed of helices.

Functional Assay: Binding of C-Occ and the Tryptic Fragments to ZO-2 from MDCK Cells—The known function of the entire C-terminal cytoplasmic region of occludin is to form complexes with TJ scaffolding proteins to anchor the TJ to the cytoskeleton. Fig. 6 shows the results of the binding assay for bacterial expressed C-Occ, demonstrating that it binds to ZO-2 (first panel). The absence of the ZO-2 band when ZO-2 is first removed from the MDCK cell lysate by immunoprecipitation, and the presence of this band when a non-relevant antibody (to
GFAP) is used instead, demonstrate the specificity of the binding. To test the hypothesis that the domain region of C-Occ isolated by limited proteolysis retains this function, antibody to ZO-2 was used to co-immunoprecipitate the tryptic fragments with ZO-2 protein from MDCK cell lysate. Both 21.5-kDa and 19.2-kDa fragments were precipitated by this method, demonstrating their binding to ZO-2 (Fig. 7). The specificity of the co-immunoprecipitation was tested using irrelevant antibody against the astrocyte intermediate filament protein, glial fibrillary acidic protein (GFAP), which is not present in MDCK cells, to eliminate the possibility that the binding was through immunoglobulin instead of ZO-2. This result is shown in Fig. 7, lane 3, where neither 21.5- nor 19.2-kDa fragments are present in the precipitate. Results from the co-immunoprecipitation experiments demonstrate that the tryptic fragments alone can bind to ZO-2, and also suggest that similar binding occurs to the other scaffolding proteins (ZO-1 and ZO-3) by virtue of their similar structure. Under the lysis conditions used in this study,
there was less ZO-1 than ZO-2 present in the soluble fraction, suggesting that the binding affinities of occludin with the three different TJ scaffolding proteins might not be the same. Further experiments are needed to investigate the regulation and binding affinity between occludin and the different TJ-scaffolding proteins.

Molecular Interaction of Bacterial-expressed dN-Occ and Insect Cell Expressed ZO-2—The above-mentioned study does not provide information on whether ZO-2 interacts directly with occludin. It is conceivable that the ZO-2 binds to occludin through another protein in the cellular milieu. To further investigate the molecular interaction between the binding domain of occludin (dN-Occ) and a homogeneous ZO-2 sample, we used an in vitro binding assay. To obtain ZO-2 for this assay, we overexpressed the full-length ZO-2 with a His tag at the N-terminal end in insect cells using the Baculovirus system. The formation of dN-Occ-ZO-2 complex was monitored by immobilizing His-tagged ZO-2 with antibody on protein A/G-agarose, and then titrating His-tagged dN-Occ from 1 to 100 μM into the binding reactions. Western blot results using occludin antibody (Fig. 8a) show increasing amounts of dN-Occ bound to ZO-2 as higher concentrations of dN-Occ were added in the binding reaction. After fitting the data to Equation 1, the predicted binding curve is shown as a solid line in Fig. 8b.

When simulations were performed using a model in which one ZO-2 molecule binds to two dN-Occ molecules or, alternatively, 2 B. H. Peng and G. Campbell, unpublished data.

FIG. 4. Distribution of predicted secondary structures related to His-tagged proteins. A. Jpred method predicts five α-helices and one β-strand in the depicted locations. B. His-tagged protein (C-Occ) encodes the entire C-terminal cytoplasmic region of human occludin (292 amino acids, including the His tag). C, the shorter His-tagged protein (dN-Occ) is constructed by deleting 121 amino acids between the His tag and the last four predicted helices. It has a high percentage of helical structure and corresponds in sequence to the limited proteolysis product.

FIG. 5. Circular dichroism analysis of purified His-tagged occludin proteins. Top panel, CD spectra of C-Occ and dN-Occ. Bottom panel, calculated secondary structure composition of these proteins using CDFIT program. As expected from the secondary structure prediction, dN-Occ has a much higher percentage of helical structure than C-Occ.

FIG. 6. Specific binding of C-Occ to ZO-2. The ZO-2 band (arrow) is evident on the Western blot after binding to C-Occ immobilized on ProBond nickel resin. No binding is seen when the His tag peptide is substituted for C-Occ (panel 1). Prior immunoprecipitation of ZO-2 from the MDCK cell lysate removes this band from the binding assay product (panel 2). The ZO-2 band remains after immunoprecipitation of the lysate with antibody to the irrelevant antigen, GFAP (panel 3). Panel 4 shows the results for the total cell lysate prior to binding (MDCK) and for the precipitate remaining (PIP-Z2) after the cell lysate was treated with ZO-2 antibody prior to the binding experiments shown in panel 2. The lower band in all lanes is an unidentified cellular protein with which the ZO-2 antibody cross reacts, but which does not specifically bind to C-Occ. (Western blot stained using ZO-2 antibody and HRP-DAB system.)

FIG. 7. Co-immunoprecipitation of ZO-2 and C-Occ. ZO-2 was immunoprecipitated from the MDCK lysate and the precipitate was incubated in a binding reaction with C-Occ immobilized on Protein G-agarose. As expected from the secondary structure prediction, dN-Occ has a much higher percentage of helical structure than C-Occ.

Non-specific binding of dN-Occ to protein A/G-IgG (GFAP antibody) did occur at a level of about 50% of the maximal specific binding, but was not concentration-dependent. Non-specific binding at low dN-Occ concentration was greater than specific binding, probably due to the absence of the bulky ZO-2 molecule, which may block some of the non-specific binding sites. Thus, it was not appropriate to simply subtract non-specific from specific binding. This idea is also supported by the results of co-immunoprecipitation experiments (Fig. 7), which showed that non-specific binding of the C-Occ domain to IgG is not appreciable in the presence of high concentration of cellular proteins, which may also occupy non-specific binding sites.

References:

2 B. H. Peng and G. Campbell, unpublished data.
two ZO-2 molecules bind to one dN-Occ molecule, the results are more consistent with multiple ZO-2 molecules binding to one dN-Occ molecule with an apparent dissociation constant, $K_d$, of $-2 \mu M$.

**DISCUSSION**

We have isolated and identified a structural domain within the C-terminal cytoplasmic region of human occludin that retains the major function of this part of the protein, i.e. binding to the scaffolding proteins, demonstrated here specifically for ZO-2. This domain is protected from limited proteolysis of both recombinant occludin expressed in bacteria and native occludin extracted from both epithelial and endothelial cells. The existence of a stable domain that contains 22 potential digestion sites for trypsin alone, as a product of limited proteolysis, is strong evidence that this region is tightly folded. The domain sequence was analyzed by N-terminal sequencing and peptide mass fingerprinting, and found to consist of the last 119 amino acids (404–522) of the C-terminal end of the protein. Based on secondary structure predictions, and corroborated by circular dichroism data, this domain contains four of the five predicted $\alpha$-helices in this region, constituting the great majority of the structural elements. The structural model that emerges from this analysis is that of a compact helical domain tethered to the cytoskeleton.

Based on sequence analysis and secondary structural predictions, this domain appears to be highly conserved (>90% identity) between human, canine and mouse. The transition between the N terminus of the domain and the loosely coiled region occurs within a 7 amino acid sequence that is conserved across all species reported (32). Secondary structural predictions (Jpred, data not shown) show similar arrangements of $\alpha$-helices in this C-terminal region for each species.

Other approaches that have been used to investigate the binding region of occludin with the scaffolding proteins include the use of recombinant peptides corresponding to the C-terminal region with various truncations and deletions (18, 33), and a synthetic peptide with sequence identical to the presumptive binding site on occludin (19). Significantly, these previous studies have helped localize the binding domain to the C-terminal half of this region. The work by Wittchen et al. (33) also suggests that the binding, at least between recombinant proteins of the occludin binding region and ZO-2, occurs directly between the two molecules. Binding studies using cell lysates, as in our study, cannot completely rule out the possibility of binding through intermediate molecules.

In order to circumvent some of the limitations of previous studies, this in vitro study was initiated. The use of limited proteolysis to separate the domain from the protein has the advantage of allowing the natural folding of the entire region of the protein (either after bacterial expression or under natural cellular conditions) prior to isolation, rather than depending on folding of truncated or isolated sequences to emulate the nat-
Occludin Cytoskeleton-anchoring Domain

REFERENCES

1. Powell, D. W. (1981) Am. J. Physiol. 241, G275–G288
2. Gumbiner, B. (1997) Am. J. Physiol. 283, C749–C758
3. Rubin, L. L. (1999) Curr. Opin. Cell Biol. 11, 830–833
4. Furuse, M., Hirase, T., Itoh, M., Nagafuchi, A., Yonemura, S., Tsukita, S., and Tsukita, S. (1996) J. Cell Biol. 135, 1771–1782
5. Furuse, M., Fujita, K., Hiragi, T., Fujimoto, K., and Tsukita, S. (1998) J. Cell Biol. 141, 1539–1550
6. Mitic, L. L., Van Itallie, C. M., and Anderson, J. M. (2000) Am. J. Physiol. Gastrointest. Liver Physiol. 279, G250–G254
7. Morita, K., Furuse, M., Fujimoto, K., and Tsukita, S. (1999) Proc. Natl. Acad. Sci. U. S. A. 96, 511–516
8. Morita, K., Sasaki, H., Fujimoto, K., Furuse, M., and Tsukita, S. (1999) J. Cell Biol. 145, 579–588
9. Morita, K., Sasaki, H., Furuse, M., and Tsukita, S. (1999) J. Cell Biol. 147, 185–194
10. Nimi, T., Nagashima, K., Ward, J. M., Minos, P., Zimonjic, D. B., Popescu, N. C., and Kimura, S. (2001). Mol. Cell Biol. 21, 7380–7390
11. Furuse, M., Furuse, K., Sasaki, H., and Tsukita, S. (1999) J. Cell Biol. 143, 263–272
12. Fanning, A. S., Jameson, B. J., Jesaitis, L. A., and Anderson, J. M. (1998) J. Biol. Chem. 273, 29745–29753
13. Haskins, J. S., Wu, S. H., Bibb, J. M., and Stevenson, B. R. (1998) J. Cell Biol. 141, 199–208
14. Itoh, M., Furuse, M., Morita, K., Kubota, K., Saitou, M., and Tsukita, S. (1999) J. Cell Biol. 147, 1351–1360
15. Itoh, M., Morita, K., and Tsukita, S. (1999) J. Biol. Chem. 274, 5981–5986
16. Beatch, M., Jesaitis, L. A., Gallia, W. J., Goodenough, D. A., and Stevenson, B. R. (1996) J. Biol. Chem. 271, 25723–25726
17. Jesaitis, L. A., and Goodenough, D. A. (1994) J. Cell Biol. 124, 949–961
18. Furuse, M., Itoh, M., Hirase, T., Nagafuchi, A., Yonemura, S., Tsukita, S., and Tsukita, S. (1996) J. Cell Biol. 127, 1617–1626
19. Nusrat, A., Chen, J. A., Foeley, C. S., Liang, T. W., Tom, J., Cromwell, M., Quan, C., and Mrasty, R. J. (2000) J. Biol. Chem. 275, 28161–28162
20. Ramzan, M. A., Cookson, E. J., and Beynon, R. J. (1991) Biochem. Soc. Trans. 19, 2968–2968
21. Belova, G. I., Prasad, R., Nazimov, I. V., Wilson, S. H., and Slesarev, A. I. (1996) J. Biol. Chem. 271, 30661–30667
22. Scheler, C., Lamer, S., Pan, Z., Li, X. P., Salnikow, J., and Jungblut, P. (1998) Electrophoresis 19, 1014–1027
23. Sambrook, J., Fritsch, E. F., and Maniatis, T. (1989) in Molecular Cloning: A Laboratory Manual. Vol. 1, 2nd Ed, pp. 21–84. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY
24. Van Itallie, C. M., and Anderson, J. M. (1997) J. Cell Sci. 110, 1113–1121
25. SigmaGel for Windows, Rel. 1.05, 1995. Jandel Scientific Software: San Rafael, CA
26. Edman, P., and Begg, G. (1978) Eur. J. Biochem. 1, 80–91
27. Wilkins, M. R., Lindskog, I., Gasteiger, E., Bairioch, A., Sanchez, J. C., Hochstrasser, D. F., and Appel, R. D. (1997) Electrophoresis 18, 483–488
28. Cheng, X., and Luan, J. C. (1998) Biochemistry 37, 51–60
29. Johnson, W. C. (1999) Proteins 35, 307–312
30. Rost, B. (1996) Methods Enzymol. 266, 525–539
31. Cuff, J. A., Clamp, M. E., Siddiqui, A. S., Finlay, M., and Barton, G. J. (1998) Bioinformatics. 14, 882–893
32. Aono, Akatsuka, Y., Saitou, M., Hirase, T., Kishi, M., Sakakibara, I., Itoh, M., Yonemura, S., Furuse, M., and Tsukita, S. (1996) J. Cell Biol. 133, 45–47
33. Witschnig, P., Haskins, J. S., and Stevenson, B. R. (1999) J. Biol. Chem. 274, 35179–35185
34. Tsukamoto, T., and Nigam, S. K. (1997) J. Biol. Chem. 272, 16133–16139
35. Tsukamoto, T., and Nigam, S. K. (1999) Am. J. Physiol. 276, F737–F750
36. Kevi, C. G., Oshima, T., Alexander, B., Cse, L. L., and Alexander, J. S. (2000) Am. J. Physiol. Cell Physiol 279, C21–C30

Acknowledgments—We thank Lucy L. Lee for assistance with circu-
dichroism, J. Stephen Smith for N-terminal sequencing and Dr. Anthony M. Haag for assistance with mass spectrometry. The N-
terminal sequencing and mass spectrometry were carried out in the
Protein Chemistry Laboratory, University of Texas Medical Branch,
Galveston, TX.
In Vitro Protein Complex Formation with Cytoskeleton-anchoring Domain of Occludin Identified by Limited Proteolysis
Bi-Hung Peng, J. Ching Lee and Gerald A. Campbell

J. Biol. Chem. 2003, 278:49644-49651.
doi: 10.1074/jbc.M302782200 originally published online September 25, 2003

Access the most updated version of this article at doi: 10.1074/jbc.M302782200

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 31 references, 21 of which can be accessed free at
http://www.jbc.org/content/278/49/49644.full.html#ref-list-1