Cloning and Characterization of a Novel Peptidase from Rat and Human Ileum*

(Received for publication, May 5, 1997, and in revised form, September 17, 1997)

Benjamin L. Schneider†‡, Sundararajah Thevananther‡†, M. Susan Moyer‡†, Holly C. Walters*, Piero Rinaldo**, Prasad Devarajan†, An Qi gang Sun‡‡, Paul A. Dawson† and Meenakshisundaram Ananthanarayanan‡‡‡

From the Department of Pediatrics, Divisions of ‡Pediatric Gastroenterology/Hepatology and **Pediatric Nephrology, and the ††Department of Genetics, Yale University School of Medicine, New Haven, Connecticut 06520, the Department of Internal Medicine Gastroenterology, Bowman Gray School of Medicine, Wake Forest University, Winston-Salem, North Carolina 27157, and the ‡‡‡Department of Pediatrics, Mount Sinai School of Medicine, New York, New York 10029-6574.

A novel 100-kDa ileal brush border membrane protein (I100) has been purified by anionic glycocholate affinity chromatography. Polyclonal antibodies raised against this protein were utilized to clone and characterize I100 in rats. A partial length human I100 cDNA was identified by hybridization screening. In the rat, the I100 protein is a 746-amino acid glycosylated (calculated core molecular mass of 80 kDa) type II integral membrane protein found on the apical surface of ileal villus enterocytes. Its 2.6-kilobase mRNA is expressed in distal small intestine in rats and in humans. The I100 cDNA is homologous to but distinct from human prostate-specific membrane antigen and rat brain N-acetylaspartylglutamate peptidase. It is expressed on both the basolateral and apical surfaces of stably transfected Madin Darby canine kidney cells. Analysis of these stably transfected Madin Darby canine kidney cells and I100 immunoprecipitates of rat ileal brush border membrane vesicles reveals that it has dipeptidyl peptidase IV activity. Future investigations will need to determine the exact substrate specificity of this novel peptidase.

Intestinal reclamation of conjugated bile salts occurs primarily on the apical surface of ileal enterocytes by sodium-dependent carrier-mediated uptake (1, 2). Photofinity labeling studies of ileal brush border membrane vesicles and of ileal enterocytes had implicated a 99-kDa protein in that process (3). Strategies to identify and clone the rat ileal sodium-dependent bile acid transporter included bile acid affinity chromatography (4) and ultimately expression cloning and hybridization (5, 6). The bile acid affinity chromatography employed an anionic glycocholate-Sepharose 4B affinity matrix (7) that incorporated two critical features for bile acid binding, namely a trihydroxysteroid nucleus and a monoanionic side chain (8, 9). In the course of studies to purify the ileal sodium-dependent bile acid transporter, a unique ileal 100-kDa protein was isolated and cloned. In this study, we identify this novel ileal protein as a peptidase that has dipeptidyl peptidase IV activity.

**EXPERIMENTAL PROCEDURES**

**Animal Care**—Sprague-Dawley rats (200–250-g males) were obtained from Charles River (Raleigh, NC) and were exposed to 12-h day/night cycles. Animals were housed, fed, and handled according to the protocol approved by the National Institutes of Health Guide for the Care and Use of Laboratory Animals and under a protocol approved by the Yale Animal Care and Use Committee. Human tissues were obtained under a protocol approved by the Clinical Practices Committee of the Bowman Gray School of Medicine.

**DNA/Amino Acid Sequence Analysis**—Standard DNA and amino acid sequence analyses were performed using either the Genetics Computer Group software package (Madison, WI) or software available through the National Center for Biotechnology and Information (Bethesda, MD).

**Affinity Column Purification of the Ileal 100-kDa Protein**—The I100 protein was initially isolated based upon its ability to bind to an anionic glycocholate-Sepharose 4B affinity column (7). Ileal brush border membrane vesicles (BBMV)1 were prepared from rat ileum by divalent cation precipitation as described previously (10). Approximately 2–4 mg of these ileal BBMV were solubilized in 50 mM n-octyl-β-D-glucopyranoside (OGS), 1 M NaCl, 10 mM Tris, pH 8.0, for 2.5 h at 4 °C and spun at 100,000g for 2.5 h. The supernatant was diluted to 15 mM OGS before applying to the bile acid affinity column. The bound proteins were eluted from the affinity column with 50 mM OGS, dialyzed against 25 mM NH4HCO3, 0.05% SDS, and then lyophilized (Fig. 1). The resulting affinity-purified protein was further purified by one-dimensional SDS-PAGE (11). The band of protein(s) of approximately 100 kDa molecular mass was excised from the gel, and the proteins were electroeluted using an Elutrap electroseparation system according to the manufacturer’s instructions (Schleicher & Schuell). The final protein preparation was analyzed by SDS-PAGE (Fig. 1).

**Polyclonal Antibody Preparation**—Polyclonal antibodies were raised in rabbits by subcutaneous and intramuscular injections of approximately 25 μg of the 100-kDa protein mixed in RIBI adjuvant (RIBI Immunochemicals, Hamilton, MT). Injections were performed on days 0, 19, and 35, and serum was obtained by ear vein bleeding on day 45.

**100-kDa Protein/Antibody Characterization**—The 100-kDa antibody was initially characterized by Western blotting and indirect immunofluorescence. Western blotting was performed by standard techniques using 100 μg of ileal BBMV and either a 1:500 or 1:2500 dilution of the antibody. The resulting antibody-purified protein was further purified by one-dimensional SDS-PAGE (11).

1 The abbreviations used are: BBMV, brush border membrane vesicles; I100, ileal 100-kDa protein; MDCK, Madin Darby canine kidney; OGS, n-octyl-β-D-glucopyranoside; 5′-RACE, 5′-rapid amplification of cDNA ends; PCR, polymerase chain reaction; DPP IV, dipeptidyl peptidase IV; NAAALDase, N-acetylated α-linked acidic dipeptidase; NAAG, N-acetylglutamate synthetase; PSM, prostate-specific membrane antigen; PAGE, polyacrylamide gel electrophoresis; PBS, phosphate-buffered saline; bp, base pair(s).

† To whom correspondence should be addressed: Division of Pediatric Gastroenterology, Box 1198, Mt. Sinai Medical Center, One Gustave L. Levy Place, New York, NY 10029-6574. Tel.: 212-824-7755; Fax: 212-876-5631.

‡ This work was supported in part by National Institutes of Health Grants DK-02076, DK-34989, DK-43509, DK-47987, HD-20632, and HD-27757 and The March of Dimes Basil O'Connor Starter Scholar Grants DK-02076, DK-34989, DK-43509, DK-47987, HD-20632, and HD-27757 and the **Department of Genetics, Yale University School of Medicine, New Haven, Connecticut 06520, the ††Department of Internal Medicine Gastroenterology, Bowman Gray School of Medicine, Wake Forest University, Winston-Salem, North Carolina 27157, and the ‡‡‡Department of Pediatrics, Mount Sinai School of Medicine, New York, New York 10029-6574.

§ To whom correspondence should be addressed: Division of Pediatric Gastroenterology, Box 1198, Mt. Sinai Medical Center, One Gustave L. Levy Place, New York, NY 10029-6574. Tel.: 212-824-7755; Fax: 212-876-5631.

This paper is available on line at http://www.jbc.org
formed as described, except Triton X-100 was omitted as a permeabilizing agent (6). A 1:100 dilution of immune serum or 1:50 dilution of preimmune serum was utilized. Triton X-114 counterphase partitioning studies were utilized to assess whether the 100-kDa protein was an integral membrane protein (13, 14). 750 μg of ileal BBMV were partitioned. The final aqueous and detergent phases were mixed with a 4 × SDS-PAGE gel loading buffer to yield 12.5% (v/v) glycerol, 0.7 M β-mercaptoethanol, 2% (w/v) SDS, 60 mM Tris, pH 6.8, and were analyzed by Western blotting using the 100-kDa polyclonal antibody. Deglycosylation experiments were carried out as described previously (6). Briefly, 50 μg of denatured ileal BBMV were treated with 0.6 units of N-glycanase (Genzyme, Cambridge, MA) overnight at 37 °C and were then analyzed by Western blotting using the 1100 antibody.

The specificity of I100 binding to the anionic glycolate-Sepharose 4B affinity column was determined by elution of the bound proteins with increasing concentrations of taurocholate. Ileal BBMV were solubilized as described above and applied to the anionic glycolate-Sepharose 4B affinity column. After equilibration and binding to the column, proteins were eluted with increasing concentrations (0.10–1.0 M) of sodium taurocholate in 15 mM OGS, 1 mM NaCl, 10 mM Tris, pH 8.0. Finally, the column was stripped of residually bound proteins using 50 mM OGS. Pools of eluted proteins were dialyzed against 25 mM NH₄HCO₃, 0.05% SDS and then lyophilized. The lyophilized proteins were solubilized in 1.57 M glycerol, 0.7 M β-mercaptoethanol, 0.7 M SDS, 0.05 M Tris, pH 6.8, and were analyzed by Western blotting.

N-terminal Sequencing of I100— I100 antibodies were affinity purified using ileal BBMV proteins as antigen. 3.5 mg of ileal BBMV were separated by preparative one-dimensional SDS-PAGE and electrophotochemically transferred to nitrocellulose. The filter was blocked with 150 mM NaCl, 10 mM Tris (pH 7.5), 3% bovine serum albumin, 0.2% Triton X-100 (TBSA), and a region of the membrane was excised that encompassed proteins that migrated with an apparent mass of 100 kDa. This membrane strip was incubated overnight at 4 °C with 20 ml of a 1:20 dilution of immune serum in TBSA. After several washes with TBSA, the bound antibody was released with 0.2 M glycine, pH 2.8, neutralized with 1 M Tris, pH 8.0, and concentrated using a Centricron-100 device (Amicon, Beverly, MA). This purified antibody was then used to immunoprecipitate the 100-kDa protein from 0.75 mg of ileal BBMV protein that had been solubilized in 1% Nonidet P-40, 0.8% bovine serum albumin, 150 mM NaCl, 10 mM Tris (pH 7.4). Antigen-antibody complexes were captured with protein A-Sepharose and eluted from the Sepharose by boiling in 25% glycerol, 1.4 M β-mercaptoethanol, 4% (w/v) SDS, 120 mM Tris-HCl, pH 6.8. The immunopurified 100-kDa protein was resolved by one-dimensional SDS-PAGE, transferred to Immobilon-P transfer membrane (Millipore, Bedford, MA), and stained with Coomassie Blue. The protein was excised and subjected to automated Edman degradation to yield N-terminal sequence at the W. M. Keck Foundation Biotechnology Resource Laboratory at Yale University (15).

Cloning of the I100 cDNA— The I100 cDNA was cloned by conventional antibody screening of a rat ileal λ Zap II library kindly provided by Raymond DuBois (Vanderbilt University, Nashville, TN). Nonspecific IgG binding globulin binding was reduced by preabsorbing the filters with a 1:50 dilution of guinea pig serum in 150 mM NaCl, 10 mM Tris (pH 7.4), 1% bovine serum albumin at 4 °C overnight. A 1:1000 dilution of Escherichia coli/phase lysate precleared immune serum was used to screen 0.5 × 10⁶ colonies. Positive colonies were detected using the pico Blue immunoscreening kit according to the manufacturer’s instructions (Stratagene). Clones that were positive on tertiary screening were directly subcloned into pBluescript SK+ using the ExAssist/SOLR system according to the manufacturer’s instructions (Stratagene). Positive clones were sequenced on both strands at the William Keck Biotechnology Resource Laboratory at Yale University using a series of oligonucleotide primers, AmpliFinder DNA polymerase, and an ABI 373A automated sequencing system (ABI, Foster City, CA) (16). Sequence analysis was performed using software of the Genetics Computer Group (Madison, WI).

The 5’-end of the I100 cDNA was cloned by 5’-rapid amplification of cDNA ends (17) using the 5’ AmpliFinder RACE kit according to the manufacturer’s instructions (CLONTECH). I100 specific cDNA was synthesized from ileal poly(A)+ RNA using an antisense primer located at nucleotide 652 of clone BS15C1 (CACGGTACACCAACACCTCC). The initial cDNA synthesis experiment utilized an anchor primer and an oligonucleotide beginning at nucleotide 252 (CACGCCAGGCTGCCCAGGGCC), which was nested within the primer used for cDNA synthesis. Reamplification was performed with the same primers, and the resulting product was subcloned into a TA subcloning vector (Invitrogen).

A λ GT10 human ileal cDNA library (18) was screened by standard plaque hybridization techniques using a random hexamer-primed 32P-labeled I100 cDNA probe (KpnI/EcoRI digest of clone BS15C1, utilizing both fragments for labeling, i.e. nucleotides 1–1315 and 1316–2512). Hybridization was performed for 18 h at 42 °C in buffer containing 30% (v/v) formamide, 5 × SSPE, 5 × Denhardt’s solution, 0.1% (w/v) SDS, 100 μg/ml salmon sperm DNA. After hybridization, filters were washed to a final stringency of 0.2 × SSC, 0.1% SDS at 55 °C. Positive clones were plaque purified, and plate lysate DNA was isolated (19) and subcloned into pBluescript KS II (Stratagene).

Tissue Distribution of the I100 mRNA— The tissue distribution of the I100 mRNA was assessed by Northern blotting of total RNA from rat liver, kidney, ileum, jejunum, colon, placenta, brain, heart, and lung. The localization of the I100 transcript along the longitudinal axis of the intestine was then assessed by Northern blotting of total RNA from the esophagus, stomach, deciles of the small intestine, cecum, and proximal and distal halves of the colon. RNA loading of the blots was assessed by 28 S ribosomal oligonucleotide hybridization (20). Northern blot analysis of human I100 was performed using 5 μg of poly(A)+ RNA from human liver, cecum, and ileum and a human multiple tissue Northern blot (CLONTECH), which were hybridized with a random hexamer-primed 32P-labeled human I100 cDNA (Human 1, Fig. 6, corresponding

---

**Fig. 1. Purification of the I100 Protein.** Top, chromatogram of the elution profile of OGS-solubilized ileal brush border membrane proteins. Approximately 2 mg of protein was applied to an anionic glycolate-affinity column. Nonabsorbed proteins flow through the column in peak 1, while absorbed proteins are eluted from the column using 50 mM OGS (peak 2). A_{280nm} refers to the absorbance of 600 nm of light after completion of the Bio-Rad protein assay according to the manufacturer’s instructions (Bio-Rad). Bottom left, SDS-PAGE analysis of peak 2 proteins from ileum and jejunum. Aliquots of peak 2 proteins from the indicated tissues were resolved by SDS-PAGE analysis on a 10% polyacrylamide gel and stained with Coomassie Blue. An approximately 100-kDa (arrow) protein(s) is enriched from ileum but not jejunal membranes. Bottom right, the ileal 100-kDa protein(s) was excised and eluted from the gel on the left and reanalyzed by SDS-PAGE before use for antibody preparation.
Cloning and Characterization of a Novel Ileal Peptidase

Generation of an I100 Expression Construct—The I100 cDNA was subcloned into a TA-based eukaryotic expression vector (Invitrogen) for functional analysis. Nested primers were designed to the 5’ and 3’ untranslated portions (sense 1, GCCCGGGCCACAGGATGGCCGGCGAAAATAC; antisense 1, GAGAAACAGTTGCAAAGGTTGC; antisense 2, GAAAAGCCCAAGGACCGGGG) of I100, and a full-length I100 cDNA was prepared by nested reverse transcription-PCR using rat ileal RNA as a template. The reverse transcription-PCR product was subcloned into the bidirectional eukaryotic expression vector, and the properly oriented I100 cDNA was sequenced in its entirety. PCR-induced mutations were obtained in all full-length clones. To obtain a mutation-free I100 expression plasmid, a BagI fragment (nucleotides 301–1586) from BS15C1 was subcloned into one of the RT-PCR-based expression constructs. The full-length I100 expression construct was then used to transiently transfect COS-7 cells and to stably transfect MDCK cells.

Transfections of MDCK I Cells and Selection of Stably Transfected Colonies—MDCK-I cells were transfected overnight with the I100 eukaryotic expression constructs (1.0 µg/well (12-well plate)) by the lipofectin method (3.5 µg of lipofectin (Life Technologies, Inc.), 1.0 µg of plasmid). Untransfected cells were exposed to lipofectin and treated similarly to the transfected ones. Approximately 60–72 h after transfection, the COS-7 cells were lysed in 100 µl of 1 × SDS-PAGE loading buffer and analyzed by Western blotting.

Sodium-dependent Taurocholate Uptake of Transiently Transfected COS-7 Cells—Sodium-dependent taurocholate uptake was determined 72 h post-transfection by incubating transfected COS-7 cells with 1.0 µM [3H]taurocholate in 0.5 ml of 116 mM NaCl (or 116 mM choline chloride), 2.0 mM KCl, 1.1 mM KH2PO4, 0.8 mM MgSO4, 1.8 mM CaCl2, 10 mM glucose, 10 mM HEPES (pH 7.4, 37 °C). After incubating for 15 min at 37 °C, the cells were washed three times with 1.0 ml of ice-cold choline containing incubation buffer and lysed with 0.5 ml of 0.5% Triton X-100 in water. Aliquots were taken to determine cell-associated protein and radioactivity. The same assay was also used to assess sodium-dependent taurocholate uptake in untransfected MDCK and MDCK cells stably transfected with the I100 expression construct (below).

Transfection of MDCK-I Cells—MDCK-I cells were transfected overnight with the I100 eukaryotic expression construct (5.0 µg of plasmid and 18 µl of lipofectin in 5.0 ml of Opti-MEM reduced serum media). On day 6, the cells were split 1:5 in media containing 0.9 mg/ml Geneticin (G418) and transferred to 100-mm dishes. Stably transfected neomycin-resistant colonies were picked with the aid of cloning cylinders and transferred to 22-mm wells containing 1.0 ml of media containing G418.

Analysis of the Stably Transfected MDCK Clones—The neomycin-resistant colonies were lysed in 1.0% SDS and analyzed by SDS-PAGE and Western blotting to determine the presence and confirm the size of the expressed protein. Preliminary analysis of membrane expression was performed by cell surface labeling. Pilot studies were performed by labeling rat ileal BBMV with sulfo-NHS biotin (Pierce) and and the 100-kDa protein was immunoprecipitated with the polyclonal antibody. Briefly, 3 mg of rat ileal BBMV were suspended in PBS, 1 mM MgCl2, 0.1 mM CaCl2 (PBSCM) and allowed to react with 1 mg of sulfo-NHS biotin in a final volume of 3 ml at 4 °C for 30 min. Unreacted sulfo-NHS biotin was removed by washing the vesicles repeatedly with PBSCM. The vesicles were then solubilized by incubation at 4–5 °C for 30 min in immunoprecipitation buffer (IP) containing 150 mM NaCl, 1% (v/v) Nonidet P-40, 10 mM Tris-HCl, pH 7.4, and insoluble proteins were removed by centrifugation at 150,000 × g. I100 protein immunoprecipitation was performed in IP buffer containing 0.4% bovine serum albumin and a 1:100 dilution of either immune or preimmune serum. Immunocomplexes were precipitated with protein A-Sepharose and washed repeatedly with IP buffer containing 0.05% SDS. The immunoprecipitate was released with 4 × SDS-PAGE loading buffer, analyzed by SDS-PAGE, transferred to nitrocellulose, and detected using horse-
radish peroxidase-conjugated avidin and chloronapthol as a visualizing agent. Cell surface labeling of the MDCK cells was further assessed by indirect immunofluorescence using confocal microscopy. Untransfected and stably transfected MDCK cells were grown on Transwell filters (Costar, Cambridge, MA) in MEM-E medium supplemented with 10% fetal bovine serum, 2 mM glutamine, 100 μg/ml streptomycin, and 100 units/ml penicillin until they just reached confluence (3–4 days). Cells were washed with PBS and fixed in situ with 100% acetone for 20 min. After three washes with PBS, cells were incubated with normal goat serum for 60 min to block nonspecific staining. Labeling was achieved by exposing the cells to a 1:1000 dilution of the polyclonal antibody in PBS with 2% bovine serum albumin and 10% goat serum for 60 min at room temperature. The cells were then extensively washed with PBS and incubated in the dark for 30 min with Cy3-labeled goat anti-rabbit IgG (Amersham Life Sciences), diluted 1:500 in PBS with 0.5% bovine serum albumin. After three final washes with PBS, the filters were mounted on glass slides with the cells facing upward, covered with Crystal/mount (Biomeda Corp., Foster City, CA) and coverslipped. Confocal microscopy was performed with a Bio-Rad MRC-600 scanning laser microscope attached to a standard microscope (Carl Zeiss, Inc., Thornwood, NY) with a Leitz 50 × water immersion lens (Leica, Inc., Deerfield, IL). For Z section microscopy, the confocal aperture was set at its minimum section thickness (1 μm), and images were collected in 1.2-μm steps.

Sodium-dependent uptake of taurocholate was assayed as described above using either untransfected MDCK cells or one stably transfected with the I100 expression construct.

Peptidase Activity of the I100 Protein—Dipeptidyl peptidase activity of the I100 protein was determined by fluorescent analysis (excitation at 380 nm and emission at 450 nm) of the hydrolysis of Gly-Pro 7-amido-4-methylcoumarin (Sigma) (21). Assays were performed using I100 that was immunoprecipitated from rat ileal BBMV or by analysis of stably transfected MDCK cells. For the immunoprecipitation assays, 1 mg of BBMV was solubilized in 1 ml of IP buffer by stirring for 30 min. at 4 °C. The detergent extract was pipetted up and down with a 23-gauge needle 10 times and subsequently centrifuged at 150,000 × g for 20 min. 100 μl of supernatant was subjected to immunoprecipitation for 2.5 h at room temperature with a 1:166 dilution of either preimmune or immune rabbit serum (372 μl of IP buffer, 100 μl of BBMV supernatant, 25 μl of 10% BSA, and 3 μl of serum). Immune complexes were captured at 4 °C for 50 min with 50 μl of a 50% solution of protein A-Sepharose. The Sepharose beads were washed four times with 1.0 ml of IP buffer and then incubated with 2.0 ml of buffered substrate solution at room temperature for 60 min.

The stably transfected MDCK cells were studied either as homogenates or as intact confluent cells. For homogenate analysis cultured...
cells were washed three times with ice-cold 0.9% NaCl. The cells were scraped with rubber policeman and lysed in 5 mM HEPES, pH 7.4, by repeated aspiration through a 23-gauge needle and stored frozen at −70 °C. The enzyme activity was initiated by the addition of 13 µg of cell lysate from control or stably transfected MDCK-I cells to 2.0 ml of buffered substrate solution (50 mM Gly-Pro 7-amido-4-methylcoumarin in 150 mM glycine, pH 8.6). For analysis of intact cells, untransfected and I100 stably transfected MDCK cells were grown to confluence in 35-mm dishes. The confluent monolayer of cells was washed twice with 1× PBS and overlaid with 1.0 ml of buffered substrate solution for 15 min at 37 °C.

RESULTS

Affinity Column Purification of the I100 Protein—Approximately 50% of the OGS solubilized ileal brush border membrane proteins were bound to the anionic glycocholate-Sepharose 4B affinity column and were eluted by a high concentration of OGS (Fig. 1). One-dimensional SDS-PAGE analysis of the high detergent eluted fraction (peak 2) revealed protein with a molecular mass of approximately 100 kDa, which was enriched in ileal but not jejunal BBMV (Fig. 1). The 100-kDa protein(s) was further purified by one-dimensional SDS-PAGE and electroelution from excised gel slices (Fig. 1).

Polyclonal Antibody Preparation—Immune serum recognized a diffuse band of protein(s) at approximately 100 kDa when ileal BBMV were analyzed by Western blotting (Fig. 2A). This band was not recognized by a 5-fold greater concentration of preimmune serum. The I100 protein was specifically immunoprecipitated from biotinylated rat ileal BBMVs (Fig. 2B).

100-kDa Protein/Antibody Characterization—By indirect immunofluorescent analysis of rat ileal sections, the I100 protein was localized to the brush border membrane of ileal villus enterocytes (Fig. 3). No brush border staining was observed using immune serum in proximal jejunum or using preimmune serum in ileum. Triton X-114 countercurrent phase partitioning studies revealed that the I100 protein was exclusively found in the detergent phase, which is consistent with it being an integral membrane protein (Fig. 4A). After treatment with N-glycanase, the apparent molecular mass of I100 decreased by approximately 20 kDa, and I100 migrated as a single species of approximately 90 kDa (Fig. 4B). The specificity of the binding
to the anionic glycocholate-Sepharose 4B affinity column was
determined by elution of the bound proteins with increasing
concentrations of taurocholate. Western blot analysis of column
fractions revealed that I100 protein could be eluted from the
anionic glycocholate-Sepharose 4B affinity column with as lit-
tle as 0.5 mM taurocholate (not shown) and was completely
eluted from the column with 1.0 mM taurocholate (Fig. 5).

N-terminal Sequencing of I100 and Cloning of the I100
cDNA—Approximately 20 mg of I100 protein was purified by
immunoprecipitation. N-terminal Edman degradation yielded
the following unique primary sequence:

XXX(S/A)KIL-
GVGIXXX(S/A)XLL.

To obtain an I100 cDNA clone, a rat
Zap II library was screened using immune serum. Nine distinct
clones, ranging in size from 1.1 to 2.5 kilobases, were isolated.
One of the clones, BS15C1, was sequenced completely on both
strands, while two other clones were sequenced and found to be
identical shorter fragments of BS15C1 (see Fig. 6). PCR anal-
ysis of the remaining six clones revealed that they encoded
homologous sequences. Clone BS15C1 contained a 2527-bp in-
sert with a 2217-bp open reading frame. The clone encom-
passed a poly(A) tract that was 16 bp downstream of a consen-
sus polyadenylation signal (AATAAA) (22). Translation of the
open reading frame revealed sequence identity with the N-
terminal peptide sequence; however, the BS15C1 cDNA clone
was lacking the first six amino acids of the N-terminal peptide
and the 5'-untranslated region. Rescreening the library with
an I100 5'-end specific probe did not yield a more 5'-clone.
Therefore, 5'-RACE was employed to obtain the missing 5'
sequence. A total of 54 separate 5'-RACE subclones were anal-
alyzed by restriction enzyme digestions and by analytical gel
electrophoresis. Six clones were sequenced completely on both
strands and contained open reading frames that were identical
to the primary sequence of the I100 protein. Four clones ex-
tended to a presumed initiator methionine and one extended to
an in-frame 5'-stop codon, which was 27 bp proximal to the
initiator methionine triplet (Fig. 6). The predicted initiator
methionine lies within an appropriate consensus for initiation
of translation (23).

The I100 cDNA was therefore presumed to be 2590 bp in
length with an open reading frame encoding a 746-amino acid
protein with a predicted molecular mass of 80,640 Da (Fig. 7).
Kyte Doolittle hydropathy analysis predicted a single potential
membrane-spanning domain at the N terminus of the protein
(24). There were 10 potential N-glycosylation sites, and motif
analysis revealed an ATP/GTP binding site (P-loop) at amino
acid 225 (25). TBLASTN analysis of the I100 cDNA using the
data base at the National Center for Biotechnology Information

### Table I

| Accession no. | Description         | log probability | Reference |
|--------------|---------------------|-----------------|-----------|
| U75973       | Rat NAAG-peptidase  | -151            | Unpublished |
| M909457      | Human PSM           | -149            | 30        |
| M78417       | C. elegans ORF      | -46             | 52        |
| M58040       | Rat transferrin receptor | -34         | 53        |
| Z49626       | S. cerevisiae ORF   | -21             | Unpublished |

FIG. 7. Nucleotide and amino acid sequences of I100. The nucleotide and amino acid sequences of the I100 protein are shown. An in-frame stop codon in the 5'-untranslated region and the termination of the nested primers for generation of the expression construct are underlined with a dotted line. Lysine residues surround the potential membrane-spanning domain, which is double underlined. Potential N-glycosylation sites are circled. A predicted ATP/GTP binding site is indicated by the underlined italicized amino acids.
(Bethesda, MD) demonstrated that the I100 cDNA was unique (26). Significant homologies ($<1 \times 10^{-10}$) were found with genes encoding a presumed N-acetylaspartylglutamate (NAAG) peptidase from rat brain (gbU755793), prostate-specific membrane antigen (gbM99487), the transferrin receptor, and uncharacterized genes from Caenorhabditis elegans (embZ78417) and Saccharomyces cerevisiae (Table I). Because of the marked homology with the rat brain NAAG peptidase and the human prostate-specific membrane antigen, a human ileal cDNA library was screened to permit direct comparison of the human ileal I100 and the human prostate-specific membrane antigen sequences. Two cDNA clones have been isolated from the human ileal library that demonstrate marked sequence homology with the rat brain NAAG peptidase (Table II).

**Analysis of the Tissue Distribution of the I100 mRNA**

The tissue distribution of the I100 transcript was assessed by Northern blotting, which showed expression of a 2.6-kilobase transcript only in the small intestine (Fig. 8A). The localization of the I100 transcript along the longitudinal axis of the intestine was then assessed by Northern blotting of total RNA from the stomach, deciles of the small intestine, cecum, and proximal and distal halves of the colon. The I100 transcript was primarily expressed in the distal small intestine (Fig. 8B). A similar sized transcript was also detected in human ileum but not in spleen, thymus, prostate, testis, ovary, colon, peripheral leukocytes, cecum, or liver (Fig. 9, A and B).

**Functional Analysis of I100**

To assess the function of the I100 protein, an expression vector was prepared. The 100-kDa protein could be clearly detected by Western blotting of tissue lysates by immunoprecipitation using the I100 antibody (Fig. 2). Cell surface labeling of the stably transfected MDCK cells indicated that the I100 protein could be labeled in rat ileal BBMV with NHS biotin and immunoprecipitated using the I100 antibody (Fig. 2). Cell surface labeling of the stably transfected MDCK cells indicated that the I100 protein was expressed on the surface of the MDCK cells (data not shown). Confocal analysis confirmed this with labeling of both apical and lateral membranes of the MDCK cells (Fig. 11). Sodium-dependent bile acid transport activity was not seen in either the transiently transfected COS-7 cells.

**TABLE II**

| Rat     | Human |
|---------|-------|
| Ileum   | Brain | Prostate | Ileum |
|---------|-------|----------|-------|
| Rat ileum | 100  | 60       | 59    | 88    |
| Rat brain | 60   | 100      | 90    | 60    |
| Human prostate | 59 | 90       | 100   | 59    |
| Human ileum | 88 | 60       | 59    | 100   |

**FIG. 9. Northern blot analysis of the localization of the I100 mRNA in humans.** A, human multiple tissue Northern blot (CLONTECH). B, Northern blot of poly(A)$^+$ RNA from the indicated human tissues. Both blots were hybridized with a human I100 cDNA probe and exposed to film with an intensifying screen for 96 h. A transcript is detected in small intestines, specifically ileum. The bottom panels indicate equivalent loading of the lanes as shown by hybridization with a glyceraldehyde 3-phosphate dehydrogenase probe.
or the stably transfected MDCK cells, despite the membrane localization of the I100 protein (data not shown). The high degree of homology of the I100 protein and prostate-specific membrane antigen, an NAALADase, suggested a peptidase activity for the I100 protein. Significant homology with dipeptidyl peptidase IV (DPP IV) led to analysis of this activity in these initial studies. Enhanced dipeptidyl peptidase IV activity could be observed in both cellular homogenates (untransfected, 259 pmol/min; transfected, 375 pmol/min) and intact stably transfected MDCK cells (untransfected, 860 ± 43 pmol/min; transfected, 2527 ± 38 pmol/min, mean ± standard deviation, n = 3, p < 0.0001). Immune and preimmune sera immunoprecipitates of rat ileal BBMV were analyzed and revealed 10-fold increased DPP IV activity in the 100K immunoprecipitate (preimmune, 4.5 ± 2.0 pmol/min; immune, 46.8 ± 25.6 pmol/min; n = 3, p < 0.05).

**DISCUSSION**

Affinity chromatography was utilized to isolate a novel ileal protein (I100). Various classical biochemical investigations demonstrated that the I100 protein is an integral membrane glycoprotein with an apparent molecular mass of approximately 90 kDa following deglycosylation, which localizes to the apical membrane of villus enterocytes. The lack of a requirement for permeabilization of the ileal sections in the staining process suggests that the antigenic epitopes of the I100 protein most likely reside on the luminal surface of the enterocyte. Cloning of the I100 gene was ultimately performed by antibody screening of a rat ileal cDNA expression library. A nearly full-length cDNA clone, BS15C1, was isolated with a long open reading frame that matched N-terminal peptide sequence obtained from the purified protein. 5′-RACE was utilized to determine the 5′ sequence of the I100 cDNA, translation of which matched unresolved N-terminal sequence data (Fig. 6).

Analysis of the open reading frame of the I100 cDNA sequence suggests that it is a type II integral membrane protein (27). Hydropathy analysis indicates that it contains a single potential membrane-spanning domain extending from amino acid 7 to approximately 28 with lysine residues found on either side of the potential membrane-spanning domain (Ref. 24; see also Fig. 7). The Triton X-114 phase distribution studies are evidence that this potential membrane-spanning domain actually traverses the plasma membrane and anchors the I100 protein into the brush border membrane. The deglycosylation experiments coupled with the observation that immunostaining did not require cell permeabilization strongly argue that the majority of the I100 protein is extracellular. Thus, I100 is a type II integral membrane protein with a short intra-cellular N terminus (amino acids 1–6) and a long extracellular C terminus (amino acids 29–746) that is glycosylated.

Initial functional studies of this protein were directed at determining its role in sodium-dependent bile acid transport. The canine renal cell line, MDCK, was chosen for these studies because we have not been able to detect endogenous sodium-dependent bile acid transport activity in this cell line. In addition, it has been used in the development of stably transfected lines of both the ileal and hepatic sodium-dependent bile acid transporters (28). Confocal immunofluorescent analysis of immunostained stably transfected I100 MDCK cells showed both apical and basolateral localization of the I100 protein. The imperfect sorting of the I100 protein may be the result of overexpression of this protein and/or an absence of appropriate sorting machinery, as the I100 gene is not normally expressed in the kidney. Alternatively, the basolateral staining may represent a pool of proteins that is sorted to the basolateral membrane and then endocytosed/transcytosed to the apical membrane. This phenomenon has been observed when the hepatic canalicular dipeptidyl peptidase IV gene is expressed in MDCK cells (29). The plasma membrane localization of the I100 protein thus permits accurate assessment of its bile acid translocation properties. Analysis of three independent uptake experiments revealed no sodium-dependent bile acid transport activity. Therefore, it is clear that the I100 protein is not capable of supporting sodium-dependent bile acid transport activity on its own or in the presence of the repertoire of membrane proteins found in MDCK cells. This is not surprising given the topology of the I100 protein and the recent cloning of a distinct gene, ASBT, which is directly involved in sodium-de-
Cloning and Characterization of a Novel Ileal Peptidase

Homology to PSM led to analysis of the peptidase activities of the I100 protein. PSM was initially characterized as a membrane protein found in the prostate carcinoma cell line LNCaP (30). It seems to be expressed in both normal and malignant prostate cells. One of its apparent functions was discovered by Carter et al. (31) who screened a rat brain cDNA library using antibodies that recognized a rat brain N-acetylated alpha-linked acidic dipeptidase (32). Six immunopositive clones were identified, which all seemed to be the same gene, and were 86% identical to human PSM. The PSM cDNA was PCR amplified from the LNCaP cell line and shown to confer N-acetylaspartylglutamate (NAAG) peptidase activity after transfection into cells. PSM has also been shown to have pteroyl polyglutamyl carboxypeptidase activity (33). At the time the I100 cDNA was cloned, it was not possible to determine if this cDNA was the rat homologue of human PSM, and no rat PSM sequences were available for comparison. Northern blotting of rat prostate RNA with the I100 cDNA probe revealed no cross-reactive species (data not shown). The human I100 gene was then partially cloned to definitively resolve the issue. There is a high degree of homology between the rat and human I100 genes, which are distinct from the human PSM gene. In the interim, an unpublished rat brain clone was deposited in the GenBank, which is the apparent rat homologue of the brain/prostate NAAG peptidase (accession number U75973). Direct comparison of the rat brain NAAG peptidase and the I100 cDNA reveals a degree of homology similar to that seen with human PSM (Table II). This indicates that the brain and prostate NAAG peptidase genes are closely related, if not identical, and are distinct from I100 (Fig. 12). Preliminary data indicate that I100 is localized to chromosome 11 and may potentially explain the second weaker fluorescent in situ hybridization band that has been previously reported for the PSM gene (34, 35). A peptidase activity for the I100 gene is much more consistent with its predicted topology and localization. In this study, we have examined its dipeptidyl peptidase IV activity. MDCK cells that were stably transfected with an I100 eukaryotic expression vector displayed enhanced ability to hydrolyze Gly-Pro 7-amido-4-methylcoumarin, a substrate of dipeptidyl peptidase IV. The enhanced activity of DPP IV in the MDCK cells is potentially the result of activation of endogenously produced DPP IV. Immunoprecipitation studies were performed to exclude this possibility. In preliminary studies, the I100 antibodies did not inhibit the DPP IV activity found in rat ileal BBMV (data not shown). The immunoprecipitates of solubilized BBMV proteins using immune serum contained markedly enhanced DPP IV activity, demonstrating that the I100 protein itself had DPP IV activity.

DPP IV is a widely expressed enzyme activity characterized as a serine peptidase that cleaves N-terminal aminooxyacid-proline or acyl-alanine (36). cDNAs for DPP IV have been cloned from the rat liver, human intestine, mouse fetal thymus, and yeast (37–41). Direct comparison of the primary amino acid sequences of rat liver DPP IV with the rat I100 protein reveals 20.5% identity and 44.9% similarity over the entire length of the peptide. Like I100, DPP IV is an apical type II integral membrane of approximately 110 kDa, with a short cytoplasmic N terminus and a glycosylated extracellular domain. DPP IV in 2 B. L. Shneider, H. C. Walters, and P. A. Dawson, unpublished data.
T-lymphocytes is also referred to as CD26. CD26-negative cell lines have been shown to have residual DPP IV activity, indicating the existence of an alternative peptidase with DPP IV activity (42). The physiologic role of DPP IV is not completely clear. In the intestine, it may be involved in the nutritional assimilation of proline containing protein sources like gliadin or collagen (43, 44). In addition, it may play an important role in the regulation/degradation of biologically active peptides like growth hormone releasing factor, neuropeptide Y, peptide YY, morphcin, and glucose-dependent insulinotropic polypeptide (43, 45–47). Intestinal DPP IV activity in humans has been shown to be greatest in the ileum, consistent with the expression of the I100 gene along the longitudinal axis of the human and rat intestine (48).

Examination of the amino acid sequences of I100 and PSM (Fig. 12) suggests that these two proteins may be members of the α/β hydrolase fold family of proteins and potentially explains the range of enzyme activities that have been observed for I100 and PSM/brain NAALADase (49). The α/β hydrolase fold family of proteins includes prolyl oligopeptidase, dipetidyl peptidase IV, and acylaminoacyl peptidase. I100 is clearly distinct from prolyl oligopeptidase, dipetidyl peptidase IV, and acylaminoacyl peptidase (Fig. 12). This report demonstrates DPP IV activity in I100. The PSM/brain NAALADase activity is reminiscent of that described for acylaminoacyl peptidase DPP IV activity in I100. The PSM/brain NAALADase activity is also different from prolyl oligopeptidase, dipeptidyl peptidase IV, and its presumed catalytic site. Interestingly, analysis of sequence alignment of I100, PSM, and brain NAALADase reveals conserved Ser-Asp-His residues at amino acids 623, 663, and 686, respectively, of rat I100 (Fig. 12). Therefore, it is possible that its catalytic site may be critical for the function of the enzyme.

REFERENCES

1. Lack, L., and Weiner, I. M. (1961) Am. J. Physiol. 200, 313–317
2. Schiff, E. R., Small, N. C., and dietschi, J. M. (1972) J. Clin. Invest. 51, 1351–1362
3. Kramer, W., Burnhardt, G., Wilson, F. A., and Kurz, G. (1983) J. Biol. Chem. 258, 3623–3627
4. Shneider, B., Ananthanarayan, M., Moyer, M., Insler, N., and Sufry, F. (1991) Gastroenterology 100, 796
5. Wong, M. H., Oelkers, P., and Dawson, P. A. (1995) J. Biol. Chem. 270, 27228–27234
6. Xu, S. Y. (1986) Gene Anal. 2, 90–91
7. Card, R., and Dautry, F. (1989) Nucleic Acids Res. 17, 7115
8. Blackman, D. L., Watson, A. J., and Montrose, M. H. (1992) Anal. Biochem. 200, 352–358
9. Proudfost, N. (1991) Cell 64, 671–674
10. Kozak, M. (1989) J. Biol. Chem. 264, 229–241
11. Kyte, J., and Doolittle, R. F. (1982) J. Mol. Biol. 157, 105–132
12. Walker, J. E., Saraste, M., Runswick, M. J., and Gay, N. J. (1982) EMBO J. 1, 945–951
13. Altschul, S. F., Gish, W., Miller, W., Myers, E. W., and Lipman, D. J. (1990) J. Mol. Biol. 215, 403–410
14. Parks, G. D., and Lamb, R. B. (1991) Cell 64, 777–787
15. Sun, A. Q., Ananthanarayan, M., Soroka, C., Shneider, B. L., and Sufry, F. (1995) J. Pathol. 182, 315–317
16. Mostly, R. B., and Minor, J. T. (1996) Proc. Natl. Acad. Sci. U.S.A. 93, 749–753
17. Slusher, B. S., Robinson, M. B., Tsai, G., Simmons, M. L., Richards, S. S., and Coyte, J. T. (1994) J. Biol. Chem. 269, 2343–2348
18. Pinto, J. T., Suffoleto, B. P., Zaxin, G. M., Lin, S., Tong, W. P., May, F., Mukherjee, B., and Heston, W. D. W. (1995) Clin. Cancer Res. 43, 235–239
19. Carter, R. E., Fieldman, A. R., and Coyte, J. T. (1996) Proc. Natl. Acad. Sci. U.S.A. 93, 4535–4539
20. Ophardt, E., and Sufry, F. (1996) J. Biol. Chem. 271, 21625–21631
21. Darmoul, D., Lacasa, M., Barcillault, L., Mayogt, D., Sapin, C., Trestot, P., Barbat, A., and Trugnan, G. (1993) J. Biol. Chem. 268, 4842–4843
22. Marguet, D., Bernard, A. M., Vivier, I., Darmoul, D., Naquet, P., and Pierres, M. (1992) J. Biol. Chem. 267, 2200–2208
23. Roberts, C. J., Nothwehr, S. F., and Stevens, T. H. (1992) J. Cell Biol. 119, 89–93
24. Jacotot, E., Callebaut, C., Blanco, J., Krust, B., Neubert, K., Barth, A., and Hovanesian, A. G. (1996) Eur. J. Biochem. 239, 248–258
25. Tiruppathi, C., Miyamoto, Y., Gananpally, V., and Lebisch, F. H. (1993) Am. J. Physiol. 265, G81–G89
26. Suzuki, Y., Erickson, R. H., Sedlmayer, A., Chang, S.-K., Ikemori, Y., and Kim, Y. S. (1993) Am. J. Physiol. 265, G1153–G1159
27. Bai, J. F. P., and Chang, L.-L. (1995) Proc. Natl. Acad. Sci. U.S.A. 92, 8998–9002
28. Cartier, H., and Dusoir, D. (1995) Eur. J. Biochem. 230, 352–360
29. von Dippe, P., Ananthanarayan, M., Drain, P., and Levy, D. (1986) Biochim. Biophys. Acta 862, 352–360
30. Marcus, S. N., Schteingart, C. L., Marquez, M. L., Hofmann, A. F., Xia, Y., Leek, J., Lench, N., Maraj, B., Alley, C. M., Anderssen, S., Cross, J., Whelan, P., Mensch, M. A., and Markham, A. F. (1995) Eur. J. Biochem. 229, 242–249
31. Marguet, D., Bernard, A. M., Vivier, I., Darmoul, D., Naquet, P., and Pierres, M. (1992) J. Biol. Chem. 267, 2200–2208
32. Roberts, C. J., Nothwehr, S. F., and Stevens, T. H. (1992) J. Cell Biol. 119, 89–93
33. Jacotot, E., Callebaut, C., Blanco, J., Krust, B., Neubert, K., Barth, A., and Hovanesian, A. G. (1996) Eur. J. Biochem. 239, 248–258