Journal of Cell Science 106, 31-43 (1993)
Printed in Great Britain © The Company of Biologists Limited 1993

Molecular cloning of a 47 kDa tissue-specific and differentiation-dependent urothelial cell surface glycoprotein

Xue-Ru Wu and Tung-Tien Sun*

Epithelial Biology Unit, The Ronald O. Perelman Department of Dermatology and Department of Pharmacology, Kaplan Comprehensive Cancer Center, New York University School of Medicine, New York, NY 10016, USA

*Author for correspondence at the Department of Dermatology

SUMMARY

Despite the fact that bladder epithelium has many interesting biological features and is a frequent site of carcinoma formation, relatively little is known about its biochemical differentiation. We have shown recently that a 47 kDa glycoprotein, uroplakin III (UPIII), in conjunction with uroplakins I (27 kDa) and II (15 kDa), forms the asymmetric unit membrane (AUM) - a highly specialized biomembrane characteristic of the apical surface of bladder epithelium. Deglycosylation and cDNA sequencing revealed that UPIII contains up to 20 kDa of N-linked sugars attached to a core protein of 28.9 kDa. The presence of an N-terminal signal peptide sequence and a single transmembrane domain located near the C terminus, plus the N-terminal location of all the potential *N*-glycosylation sites, points to a type I (N-exo/C-cyto) configuration. Thus the mass of the extracellular domain (20 kDa plus up to 20 kDa of sugar) of UPIII greatly exceeds that of its intracellular domain (5 kDa). Such an asymmetrical mass distribution, a feature shared by the other two major uroplakins, provides a molecular explanation as to why the luminal leaflet of AUM is almost twice as thick as the cytoplasmic one. The fact that of the three major proteins of AUM only UPIII has a significant cytoplasmic domain suggests that this molecule may play an important role in AUM-cytoskeleton interaction in terminally differentiated urothelial cells.

Key words: bladder epithelium, differentiation, asymmetrical unit membrane, uroplakin III, cDNA cloning

INTRODUCTION

Bladder epithelium has several unique properties: (i) it elaborates, as a terminal differentiation product, a highly specialized biomembrane (the “asymmetric unit membrane” or AUM; see below) whose outer leaflet is almost twice as thick as the inner leaflet (Porter and Bonneville, 1963; Hicks, 1965; Koss, 1969). (ii) It is biologically flexible being able to “transdifferentiate” into a *bona fide* prostatic epithelium in response to certain mesenchymal influences (Cunha et al., 1983; Neubauer et al., 1983), and to give rise to an unusually wide range of neoplasms including transitional cell carcinomas, small cell carcinomas and adenocarcinomas (Mostofi, 1954; Melicow, 1974; Koss, 1975; Bryan, 1983). (iii) It can induce ectopic bone formation (Huggins, 1931; Roberts et al., 1974). Despite these interesting biological properties, and the fact that bladder carcinoma is one of the leading causes of cancer death in the western hemisphere, relatively little is known about the biochemistry of bladder epithelial differentiation. In this paper we describe for the first time the molecular cloning of the cDNA of a differentiation marker of bladder epithelium, a 47 kDa glycoprotein, which is a major constituent of the asymmetrical unit membrane.

As mentioned, the apical surface of terminally differentiated bladder epithelial cells is covered with numerous, morphologically unique, membrane plaques (0.1-0.5 μm in diameter) containing the asymmetric unit membrane. Ultrastructural studies of isolated AUM using negative staining revealed that the outer (luminal) leaflet of the AUM is highly organized, consisting of a quasi-crystalline, hexagonal array of 12 nm protein particles with P6 plane group symmetry and a center-to-center spacing of 16 nm (Hicks and Ketterer, 1969, 1970; Vergara et al., 1969; Staehelin et al., 1972; Knutton and Robertson, 1976; Severs and Hicks, 1979; Robertson and Vergara, 1980). In ultra-thin cross-sections, the AUM can be seen as having a unit membrane architecture but with an added layer of evenly spaced particles attached to the outer leaflet, which is therefore much thicker than the inner leaflet (8 nm vs 4 nm). These particles are about 4 nm high with a center-to-center distance of ~16 nm (Staehelin et al., 1972; Robertson and Vergara, 1980). The dimension and spacing of these particles indicate that they correspond to the 12 nm protein particles seen by negative staining (Hicks and Ketterer, 1969, 1970; Vergara et al., 1969; Staehelin et al., 1972; Knutton and Robertson, 1976). Interestingly, these particles can be stripped from AUM by trypsin, leaving behind a smooth and much

thinner (4 nm) outer leaflet (Caruthers and Bonneville, 1980). Taken together, these results indicate that the thickened appearance of the outer leaflet of the AUM is largely due to the presence of the 12 nm protein particles. Also in ultra-thin sections, the cytoplasmic leaflet of the AUM can be seen to interact closely with a dense meshwork of subcortical cytoskeletal elements (6-10 nm in diameter) via some short, thin, cross-bridging filaments (Chlapowski et al., 1972; Staehelin et al., 1972; Minsky and Chlapowski, 1978; Sarikas and Chlapowski, 1986, 1989). It has been proposed that such an interaction serves to anchor the AUM, and to distribute evenly the tension exerted on the urothelial surface, thus preventing it from rupturing during bladder distension.

Several features of the AUM make it an attractive system for studying the biochemical differentiation of bladder epithelium. First, we and others have developed a sucrose gradient/detergent-wash procedure allowing the preparation of milligram quantities of highly purified and morphologically intact AUM (Hicks and Ketterer, 1970; Chlapowski et al., 1972; Ketterer et al., 1973; Wu et al., 1990; also see Vergara et al., 1974). Second, we showed recently that highly purified (bovine) AUMs contain three integral membrane proteins, which we named uroplakins I (27 kDa), II (15 kDa) and III (47 kDa; Wu et al., 1990; Yu et al., 1990). Peptide mapping, partial amino acid sequencing and immunological data established that these three uroplakins represent distinct molecules (Wu et al., 1990). By raising monospecific antibodies to each of these proteins, we showed that all three uroplakins are urothelium-specific and AUM-associated in situ (Wu et al., 1990; Yu et al., 1990). Third, we have recently described a cell culture system for the serial and clonal cultivation of bovine and human bladder epithelial cells that form stratified colonies capable of synthesizing and processing all three uroplakins (Surya et al., 1990; also see Wu et al., 1990; Yu et al., 1990, 1992). Finally, we showed that all the uroplakins I and II epitopes that are recognized by our polyclonal and monoclonal antibodies are localized exclusively on the luminal leaflet of apical AUM. In contrast, the epitopes of uroplakin III were found to be associated with both luminal and cytoplasmic leaflets of AUM, suggesting a transmembrane configuration (Wu et al., 1990). This last finding raises the interesting possibility that, among the three uroplakins, the 47 kDa uroplakin III may be the only one that possesses a cytoplasmic tail capable of interacting with cytoskeletal filaments.

In this paper, we describe the molecular characterization of the 47 kDa bovine uroplakin III. Our data indicate that uroplakin III is a glycoprotein containing up to 20 kDa of hybrid- or complex-type, *N*-linked sugars attached to a core protein of 28.9 kDa, and that the N-terminus of the nascent protein chain contains a typical signal peptide sequence. These data suggest that UPIII is synthesized and processed via the classical RER/Golgi pathway instead of by free polysomes as suggested earlier (Pauli et al., 1983). The mature UPIII protein contains a single potential transmembrane domain, which divides the molecule into a long (luminal) N-terminal 189 amino acids containing all the potential *N*-glycosylation sites and cysteine residues, and a (cytoplasmic) C-terminal 52 amino acids enriched in charged residues with multiple potential phosphorylation sites. Thus the molecular mass of the luminal domain of UPIII (20 kDa plus glycosylation) greatly exceeds that of the cytoplasmic domain (5 kDa). Such an asymmetry of mass distribution, a feature shared by the other two major uroplakins (J. Yu, J.-H. Lin, X.-R. Wu, G. Kreibich, and T.-T. Sun, unpublished data), provides for the first time a molecular explanation for the morphological "asymmetry" of the two leaflets of AUM. Finally, our data indicate that, of the three major protein subunits of AUM, the 47 kDa UPIII is indeed the only one possessing a cytoplasmic tail, suggesting that it may play an important role in AUM-cytoskeleton interaction in terminally differentiated urothelial cells.

MATERIALS AND METHODS

Determination of amino acid sequence and composition

Asymmetric unit membranes (AUMs) were purified from bovine bladder epithelium (yield, 1 mg/bladder) using sucrose density gradient centrifugation followed by a detergent wash (Wu et al., 1990). These AUMs were dissolved in 1% sodium dodecyl sulfate (SDS), 25 mM Tris-HCl (pH 7.4), and the denatured uroplakins were resolved by SDS-polyacrylamide gel electrophoresis (PAGE) according to Laemmli (1970; 17% acrylamide with an acrylamide/bisacrylamide ratio of 120:1). The separated uroplakins were electrophoretically transferred to polyvinylidene difluoride (PVDF) membrane (Millipore, Bedford, MA) and stained with Coomassie Blue (Matsudaira, 1987). An area of the membrane containing approximately 3 μg (60 pmol) of UPIII (47 kDa) was excised and subjected to amino acid composition as well as N-terminal microsequencing analyses. Alternatively, pieces of polyacrylamide gel containing UPIII were cut out, washed with 70% formic acid, and treated with 5% cyanogen bromide (CNBr) in 70% formic acid for 2 hours at room temperature in the dark (Lonsdale-Eccles et al., 1981). The gel pieces were washed with 10% acetic acid and neutralized with 50 mM Tris-HCl (pH 7.4) containing 2% SDS for 2 hours. Peptides in the gel pieces were then resolved using a tricine SDS-PAGE system (16.5% total acrylamide/bisacrylamide and 6% bisacrylamide; Schagger and von Jagow, 1987). The peptides were electrophoretically transferred either to nitrocellulose membrane for immunoblotting, or to PVDF membrane for N-terminal sequencing (performed by Dr William Lane of the Harvard Microchemistry facility). For amino acid composition analysis, intact UPIII on a piece of PVDF membrane was hydrolyzed with 6N HCl at 11°C for 24 hours and then derivatized using ABI 420A derivatizer with phenylisothiocyanate (PTIC). The resulting PTC-amino acids were analyzed on-line via HPLC (ABI 130A).

Generation of cDNAs by polymerase chain reaction

Fully degenerate primers were designed according to the three partial amino acid sequences as indicated in Fig. 1, except S1 and AS1 in which the third position of leucine codon was substituted with deoxyinosine (Martin and Castro, 1985). Extra *Xba*I and *Sal*I sites were added to the 5′-ends of sense and antisense primers, respectively, to facilitate the cloning of the PCR products into plasmids. Three additional nucleotides were added to the 5′-ends of these cloning sites to improve the efficiency of restriction digestion (Kaufman and Evans, 1990). An internal probe was designed using the codon preference of human genome (Lathe, 1985). For making cDNA, total RNA was prepared from both bovine uro-

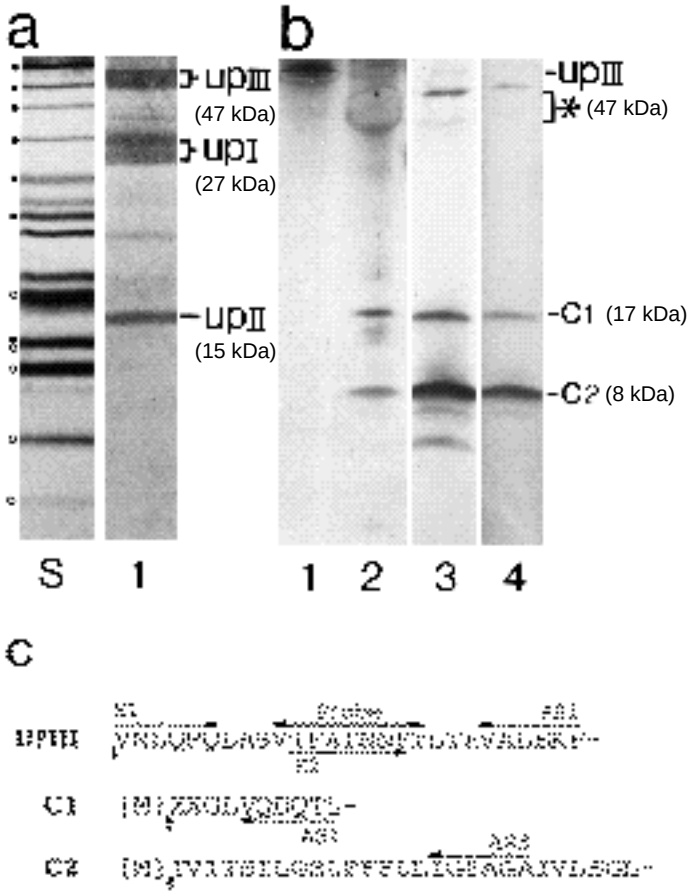

Fig. 1. Partial amino acid sequences of uroplakin III (UPIII). (a) A bovine AUM preparation. Gradient-purified asymmetric unit membrane (AUM) was dissolved in 1% SDS, and the three uroplakins were separated by SDS-PAGE (tricine-type) and visualized by silver nitrate staining (lane 1). Note that the AUM contains three major uroplakins, UP I (27 kDa), UP II (15 kDa) and UPIII (47 kDa). S denotes the molecular mass standards in which filled circles are, from top to bottom: bovine serum albumin, 66 kDa; egg albumin, 45 kDa; glyceraldehyde-3-phosphate dehydrogenase, 36 kDa; carbonic anhydrase, 29 kDa; trypsinogen, 24 kDa; trypsin inhibitor, 20.1 kDa; and α-lactalbumin, 14.2 kDa. The open circles are, from the top: myoglobin fragments, 17, 14.4, 8.2, 6.2 and 2.5 kDa. (b) Cyanogen bromide fragments of UPIII. Lane (1) shows electrophoretically purified UPIII (silver nitrate-stained). (2) CNBr-cleavage products of UPIII (silver nitrate-stained); note the generation of three major cleavage products: C1 (17 kDa), C2 (8 kDa), and a diffuse 35 kDa band. These peptides were electrophoretically transferred onto nitrocellulose sheets, and reacted with (lane 3) an affinity-purified rabbit anti-UPIII antibody, or (4) an independent rabbit antiserum raised against denatured UPIII (Wu et al., 1990). Note the strong immunoreactivities of C1 and C2 peptides. (c) The N-terminal sequences of intact UPIII and the C1 and C2 peptides. X denotes undetermined amino acids; vertical arrowhead denotes the beginning of mature UPIII sequence; and vertical arrows denote CNBr-digestion sites following presumptive methionine residues. Horizontal arrows (labeled S1 and S2) denote the amino acid sequences, on the basis of which the two PCR sense primers were designed; those labeled AS1, 2 and 3, the antisense primers. "Probe" denotes the sequence based on which the "internal probe" was designed.

thelium and esophageal epithelium (see later) and reverse-transcribed in a 20 μl reaction mixture containing 10 μg of total RNA, 50 mM Tris-HCl (pH 8.3), 75 mM KCl, 10 mM $MgCl_2$, 0.5 mM spermidine, 10 mM dithiothreitol, 1 mM each of dNTP, 1 μg/μl RNasin (Promega, Madison, WI), 0.24 μg of oligo(dT)15 primer (Boehringer Mannheim Biochemicals, Indianapolis, IN), and 10 units of AMV reverse transcriptase (Promega). PCR was carried out in a total volume of 50 μl containing 1 μl of cDNA synthesis mixture, 60 pmol of each primer, 10 mM Tris-HCl (pH 8.3), 50 mM KCl, 0.001% gelatin, 1.5 mM $MgCl_2$, 0.2 mM of each dNTP, 1.25 units of recombinant *Taq* polymerase (Perkin Elmer, Norwalk, CT). The mixture was subjected to 35 thermal cycles, each composed of 2 min of 94°C denaturation (for the first cycle, 5 min), 1 min of 50°C or 55°C annealing, and 2 min of 73°C extension (for the last cycle, 8.3 min). Portions of the PCR products were resolved on an agarose gel, blotted onto a Hybond nylon membrane (Amersham Life Science Products, Arlington Heights, IL) and hybridized with the above-mentioned, end-labeled, internal oligonucleotide probe. The 500 bp urothelium-specific product was gel-purified, digested with *Xba*I and *Sal*I, subcloned into a similarly treated pBluescript SK+ (Stratagene, La Jolla, CA), and sequenced. This insert was also labeled with a random-primed DNA labeling kit according to the manufacture's instructions (US Biochemical Corp., Cleveland, OH) and used as a probe to screen a λgt10 urothelial cDNA library.

Isolation and sequencing of UPIII cDNA clones

Total RNA was extracted from bovine bladder epithelium using guanidine thiocyanate, and purified by cesium chloride centrifugation. A urothelial cDNA library was constructed using poly(A)$^+$ RNA purified with an oligo(dT) column (Maniatis et al., 1982). Both oligo(dT)15 and random primers were used in synthesizing the first-strand cDNA in two separate reactions, and the resulting first-strand cDNAs were pooled for second-strand synthesis. The double-stranded cDNAs, with a predominant size range of 1 to 5 kb, were linked to *Eco*RI adaptor and inserted into *Eco*RI-digested λgt10 arms and packaged in vitro to produce an unamplified bacteriophage library (titer 1.6×10^6 pfu/μg cDNA).

For the isolation of UPIII cDNA clones, *Escherichia coli* (strain C600hfl; ClonTech, Palo Alto, CA) were infected with approximately 10^6 amplified bacteriophage and plated on LB agar plates. The resultant plaques were lifted onto nitrocellulose sheet (Schleicher & Schuell, Keine, NH) in duplicates. The sheet was vacuum-dried and hybridized with ^{32}P-random-primer-labeled UPIII probe at 64°C in 6× SSPE (1 M NaCl, 60 mM NaH_2PO_4, 5 mM EDTA), 5× Denhardt's solution (0.1% polyvinylpyrrolidone, 0.1% bovine serum albumin, 0.1% Ficoll), 0.1% SDS, and 100 μg/ml denatured salmon sperm DNA. Washing steps were carried out first at room temperature in 2× SSC, 0.1% SDS, and later at 68°C in 1× SSC and 0.1% SDS. One hundred clones were identified by autoradiography during an initial screening, and 20 of them were selected after a secondary screening. Inserts (>500 bp) from 14 clones were subcloned into the *Eco*RI-digested pBluescript SK+ and sequenced.

The sequences of both DNA strands were determined by the dideoxynucleotide chain termination method of Sanger et al. (1977), using deoxyadenosine 5′-[α-^{35}S]triphosphate as the label (Sequenase Kit; US Biochemicals Corp.). The T7 and T3 sequences were used as the initial sequencing primers to read the DNA sequences from both directions, and additional primers were synthesized based on the preceding sequences. For improved sequence reading in G/C compression regions, 7-deaza-dNTP or single-strand DNA binding proteins (US Biochemical Corp.) were

included in some of the sequencing reactions. The sequence was analyzed using a GCG software package, version 7.1 (Devereux et al., 1984).

Northern blot analysis

Bovine epithelial tissues were homogenized (with a Polytron) in RNA extraction buffer (4 M guanidine isothiocyanate, 0.05 M Tris-HCl (pH 7.6), 0.01 M EDTA, 1% β-mercaptoethanol, and 0.5% sodium lauryl sarcosine). Total RNA was obtained by centrifugation on a cesium chloride cushion (Chirgwin et al., 1979). For Northern blot analysis, 10 μg of total RNA was resolved on a 1.2% formaldehyde-agarose gel containing 200 mM 3-[*N*-morpholino]propanesulfonic acid (MOPS), 50 mM sodium acetate (pH 7.0) and 10 mM disodium ethylenediamine tetraacetate (EDTA). The RNA was then blotted onto Hybond nylon membrane (Amersham), which was later vacuum-dried at 80°C. The 1.7 kb cDNA insert was labeled using a random-primed kit (above), and used to probe the blot at 42°C for 16 h in 50% formamide, 5× SSPE, 2× Denhardt's solution (above), 0.1% SDS and 50 μg/ml salmon sperm DNA. Membranes were rapidly washed in 0.2× SSC and 0.1% SDS followed by a final wash at 60°C for 20 min. The size of RNA was determined by a comparison with RNA markers (Bethesda Research Laboratories) and with 28 S and 18 S ribosomal RNA.

Enzymatic deglycosylation

Gradient-purified AUMs were dissolved in 10 mM HEPES-NaOH (pH 7.5) and 0.1% SDS. After 15 min at room temperature, the mixture was adjusted to contain 1% octyl glucoside, 0.05% NaN_3, 10 mM EDTA and certain glycosidases. The glycosidases used (all from Boehringer Mannheim) were *Streptomyces plicatus* endoglycosidase H (1 mU), *Vibrio cholerae* neuraminidase (10 mU), *Flavobacterium meningosepticum N*-glycosidase F (0.4 U), and *Diplococcus pneumoniae O*-glycosidase (1 munit). The buffering component was 50 mM sodium acetate (pH 5.5) for endoglycosidase H and neuraminidase, or 40 mM sodium phosphate (pH 7.4) for all other enzymes. After incubation at 37°C for 16 h, samples were resolved by SDS-PAGE and then either stained with Coomassie blue or transferred onto nitrocellulose membrane to react with anti-UPIII antibodies.

Antibody purification and immunoblotting

Antibodies to native AUM and denatured UPIII were generated as described previously (Wu et al., 1990). Six peptides were chemically synthesized, purified by HPLC and their purity checked by mass spectrophotometry (all by Dr William Lane of the Harvard Microchemistry Facility). The sequences of these peptides are: P1 (V19 to N33; see Figs 4 and 5); P2 (N34 to P45); P3 (S87 to K101); P4 (Q179 to L191); P5 (A243 to K257) and P6 (T261 to D276). Monospecific antibodies to these synthetic peptides were affinity-purified as follows. One milligram of peptide was suspended in 1 ml of phosphate-buffered saline (PBS), blotted onto a 3 cm×7 cm nitrocellulose membrane and air-dried. The membrane was incubated with 3% bovine serum albumin, followed by an antiserum made against native AUM (a-AUM). After 1 h at room temperature, the membrane was washed extensively with PBS, and the bound antibodies were then eluted with 50 mM diethylamine (pH 11.5), which was immediately neutralized with 1 M Tris-HCl (pH 7.4). The eluted antibody was concentrated using Centricon-30 (Amicon, Danvers, MA).

For immunoblotting, proteins or peptides electrophoretically transferred to nitrocellulose membrane (Towbin et al., 1979) were incubated with rabbit or chicken primary antibodies, and then with peroxidase-conjugated, goat anti-rabbit (BioRad, Rockville Centre, NY) or rabbit anti-chicken antibodies (ICN, Irvine, CA; Woodcock-Mitchell et al., 1982). Slot blotting was carried out using a mini-slot blotter (Bethesda Research Laboratories).

Immunofluorescent staining

Bovine bladder mucosa was fixed in Zamboni's solution (15% picric acid and 2% paraformaldehyde made in PBS) and cut into 5 μm frozen sections, which were then stained by the indirect immunofluorescence technique (Sun and Green, 1978).

RESULTS

Generation of a tissue- and sequence-specific partial cDNA of UPIII by polymerase chain reaction

To generate partial amino acid sequences of uroplakin III, we resolved the proteins of gradient-purified bovine asymmetric unit membrane (AUM) by SDS-PAGE (Fig. 1a) and microsequenced the 47 kDa UPIII band yielding an N-terminal sequence of 27 amino acid residues (Fig. 1c). We also digested the electrophoretically purified UPIII with cyanogen bromide (CNBr), generating two well-defined bands (Fig. 1b, lane 2; C1/17 kDa and C2/8 kDa) both of which retained their immunoreactivities with antisera against UPIII (Fig. 1b, lanes 3 and 4). Microsequencing of the 17 kDa and 8 kDa bands yielded 8 and 27 N-terminal amino acids, respectively (Fig. 1c).

On the basis of these three stretches of amino acid sequence (Fig. 1c), we designed a sense primer (S1) and three different antisense primers (AS1 to 3). Polymerase chain reactions (PCR) were performed using bovine urothelial cDNA as a template. Two PCR products, generated using S1 coupled with AS2 (500 bp) or AS3 (650 bp; Fig. 2a), have the following properties. They react strongly and specifically with an "internal" probe (Fig. 2a′). They are not formed using esophageal cDNA template indicating that they are urothelium-specific (Fig. 2a). When sense primer 2, which was designed according to a sequence 10 amino acids (or 30 bp equivalents) downstream from that of S1, was used in combination with AS2 and AS3, we obtained two PCR products that are, as expected, 30 bp shorter than the 500 bp and 650 bp products, respectively (Fig. 2b). When we cloned the 500 bp cDNA and used it to probe northern blots, it detected a 1.7 kb mRNA present in urothelial cells but not in esophageal epithelial cells (Fig. 2c). Finally, when we sequenced this cloned 500 bp PCR product, parts of its deduced amino acid sequence matched precisely the known N-terminal sequence of intact UPIII and that of the C1 CNBr-peptide (Fig. 1c). The tissue- and sequence-specificity of this 500 bp PCR product clearly establishes that it is a partial cDNA of UPIII.

Cloning and sequencing of full-length UPIII cDNAs

To isolate full-length UPIII cDNA, we used the cloned 500 bp PCR product to screen a λgt10 cDNA library of bovine urothelium and obtained four clones. The longest one contains a 1.7 kb insert; the other three shorter clones share the same restriction pattern and most likely represent overlapping partial clones of the same cDNA (Fig. 3).

The nucleotide sequence of the 1.7 kb cDNA contains 40 bp of 5′-untranslated sequence, followed by a single

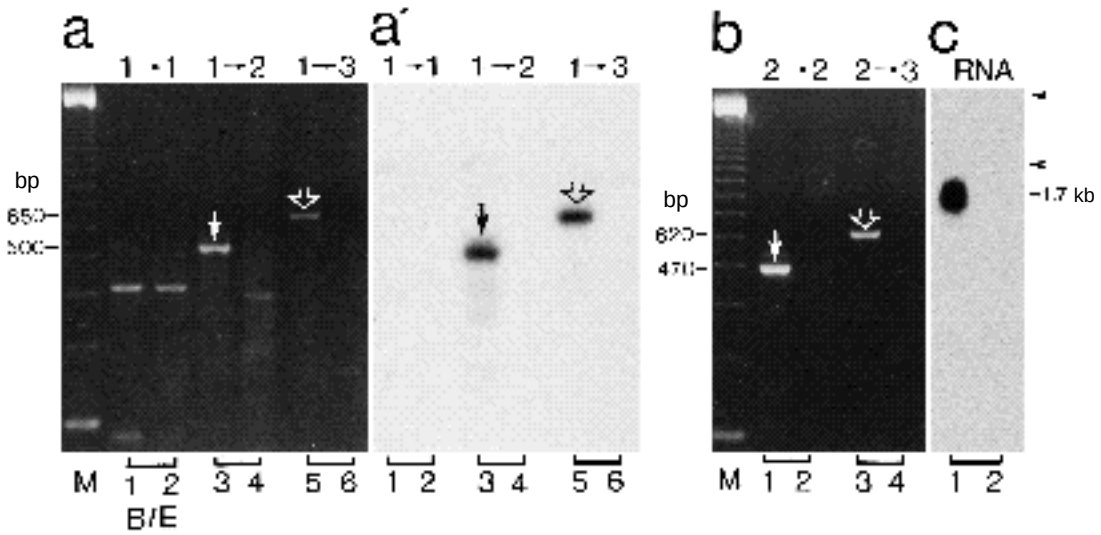

Fig. 2. Generation of a tissue- and sequence-specific partial UPIII cDNA by polymerase chain reaction (PCR). (a) PCR products generated using sense primer S1 coupled with antisense primers (lanes 1 and 2) AS1, (lanes 3 and 4) AS2, or (lanes 5 and 6) AS3. The PCR template was either urothelial cDNA (lanes 1, 3 and 5), or esophageal epithelial cDNA (lanes 2, 4 and 6). Note the generation of a 400 bp (S1→AS1) PCR product that is not urothelium-specific, and 500 bp (S1→AS2) and 650 bp (S1→AS3) PCR products that are urothelium-specific. (a′) The PCR products, as shown in (a), were probed with an "internal" probe (see Fig. 1c). Note the reactions with the 500 bp (filled arrow) and 650 bp (open arrow) PCR products, but not with the 400 bp product. (b) PCR products generated using sense primer S2 (30 bp downstream of sense primer S1) coupled with antisense primers AS2 and AS3. Note the generation of two PCR products of expected sizes of 470 bp (filled arrow) and 620 bp (open arrow). (c) Northern blot analysis. The 500 bp cDNA (Fig. 2a, lane 3, filled arrow) was cloned, and used to probe bovine (lane 1) urothelial or (lane 2) esophageal epithelial mRNA. Note the detection of a 1.7 kb urothelium-specific mRNA.

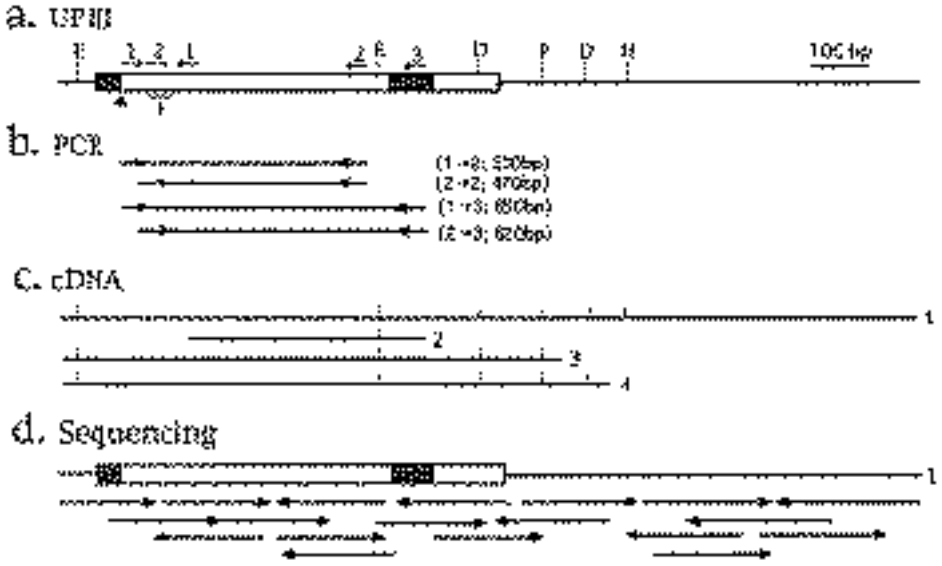

Fig. 3. Strategies for UPIII cDNA cloning and sequencing. (a) A schematic diagram of the UPIII cDNA sequence. The thick bar represents the coding region (the hatched areas denote the hydrophobic domains), while the lines denote the untranslated cDNA sequences. Arrowhead marks the N terminus of mature UPIII sequence, numbered arrows denote the position and orientation of PCR primers (Fig. 1c), and the small horizontal bar (labeled P) denotes the "internal probe". Restriction sites are: E (*Eag*I), D (*Dra*II), P (*Pst*I), and H (*Hin*dII). (b) The relationship among various PCR products. (c) cDNA clones. (d) The sequencing strategy.

open reading frame of 861 bp encoding 287 amino acids, plus 800 bp of 3′-untranslated region (Fig. 4). The presumed translation initiation methionine (encoded by the first ATG) precedes the N-terminal sequence of mature uroplakin III by 18 mostly hydrophobic amino acids, which may serve as a signal peptide (see below). Although this cDNA does not have a poly(A) tail, its 3′-untranslated region contains two polyadenylation signals, and the size of the cDNA (1.7 kb) is identical to that of UPIII mRNA (Fig. 2c), indicating that it is very close to full length.

The authenticity of the deduced amino acid sequence, as shown in Fig. 4, is suggested by several pieces of data. First, this sequence contains all three known partial amino acid sequences (underlined in Fig. 4), i.e. the N-terminal sequences of the intact UPIII and the C1 and C2 CNBr-peptides (Fig. 1c). Second, the size difference (~150 bp) between the S1→AS2 (~500 bp) and S1→AS3 (~650 bp) PCR products (Fig. 2a) is in excellent agreement with the cDNA-derived value (144 bp; Fig. 4). Third, the cDNA-predicted amino acid composition matches that of the electrophoretically purified UPIII (Table 1). Fourth, several polyclonal rabbit and chicken antisera that we previously

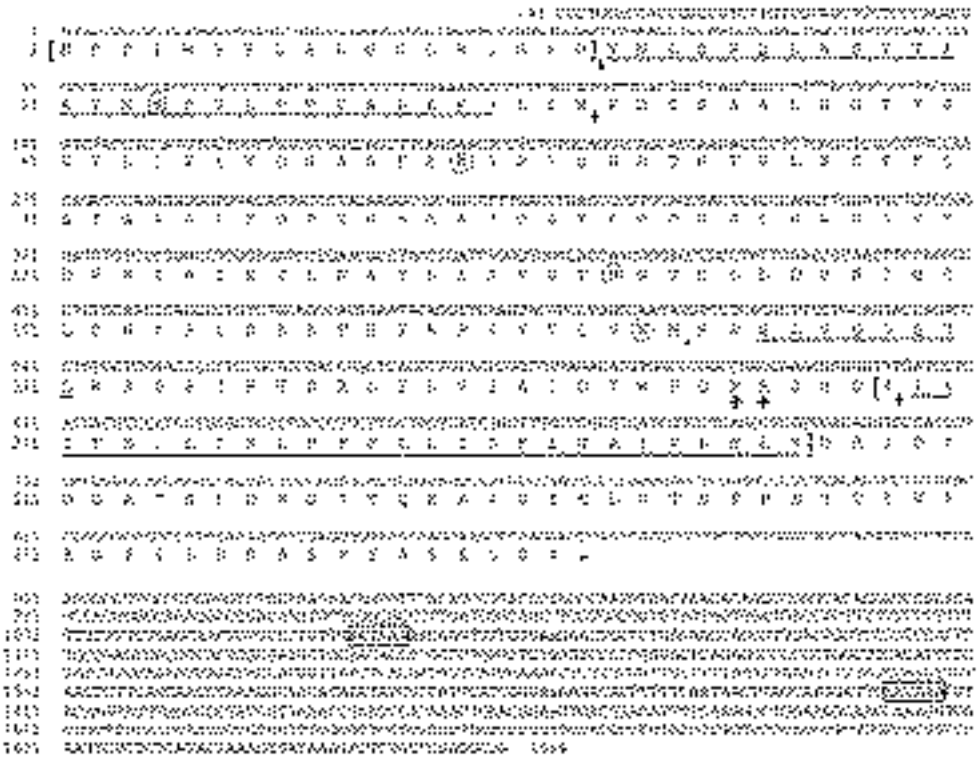

Fig. 4. The nucleotide and deduced amino acid sequence of a UPIII cDNA. The nucleotide sequence of the 1.7 kb cDNA clone is shown, along with its single open reading frame of 287 amino acids (single letter code). The initiation codon of methionine is designated as position 1; the translation termination codon is marked by an asterisk; and the two polyadenylation signals are boxed. The amino acid sequences matching those of the known partial sequences (Fig. 1c) are underlined. Arrows denote the three potential CNBr-cleavage sites. The arrowhead marks the N terminus of mature UPIII. The first and second sets of brackets mark a putative signal peptide and an internal putative transmembrane domain, respectively. The circled asparagine residues denote the four potential *N*-glycosylation sites. A dibasic sequence, marked ++, resides on the luminal side of the transmembrane domain (see text).

raised against intact UPIII molecule were found to contain antibodies recognizing six chemically synthesized UPIII oligopeptides (Figs 5 and 6). Since three of these peptides (peptides 3, 5 and 6) were actually synthesized according to purely hypothetical (cDNA-deduced) sequences (Fig. 5), this result confirms their authenticity. This experiment also enabled us to affinity-purify site-specific antibodies (Fig. 6f; see below). Finally, all these affinity-purified antibodies including those recognizing the (hypothetical) peptide 5 react with the 47 kDa UPIII (Fig. 7). Moreover, such antibodies stain immunohistochemically, as expected, the AUM-enriched urothelial umbrella cells (Fig. 8). Taken together, these data strongly support the authenticity of our cDNA clone and the correctness of its deduced amino acid sequence.

The above immunological studies also showed that different antisera react differently towards various synthetic peptides (Fig. 6). In general, antibodies raised against native AUM tend to react strongly with peptides 3 and 5 (Fig. 6a and c), and only antibodies raised against SDS-denatured UPIII react strongly with peptide 6 (Fig. 6d and e). This raises the possibility that peptide 6, located near the C terminus, may be normally buried and is "presented" antigenically only when the UPIII molecule is SDS-denatured.

UPIII is *N*-glycosylated

One puzzling aspect of the deduced amino acid sequence of UPIII, as shown in Fig. 4, is that it corresponds to a molecular mass of only 28.9 kDa, which is almost 20 kDa smaller than that of the intact UPIII (47 kDa). Since the hypothetical UPIII sequence contains four potential *N*-glycosylation sites (circled in Fig. 4), glycosylation could account for this discrepancy. To test this possibility, we dissolved purified AUM in SDS, treated the uroplakin mixture with glycosidases, and followed the SDS-PAGE mobility of UPIII using specific antibodies. The results indicate that Endo-H, an enzyme specific for high mannose-type sugars, did not change the electrophoretic mobility of UPIII (Fig.

Table 1. The amino acid composition of the cDNA-deduced UPIII sequence (labeled cDNA) agrees with that of the electrophoretically purified UPIII (labeled protein)

	Amino acid composition of uroplakin III (mole %)	
Amino acid	cDNA	Protein
Ala	8.6	9.1±0.5
Pro	5.9	6.0±0.2
Val	7.1	6.4±0.6
Ile	3.7	4.1±0.2
Leu	11.5	11.7±0.3
Phe	3.7	4.2±0.7
Met	1.1	1.8±1.0
Asx*	10.4	9.5±0.1
Glx†	7.1	8.2±0.8
Arg	4.8	5.3±0.9
Lys	2.2	2.5±0.1
His	0.7	1.0±0.2
Cys	1.5	0.4±0.0
Ser	11.5	11.4±0.8
Gly	6.7	9.0±0.4
Thr	8.9	6.9±1.2
Tyr	3.7	3.0±0.4
Trp	0.7	ND‡

The first column lists amino acids in the three-letter code, in which Asx represents a mixture of Asn and Asp, and Glx a mixture of Gln and Glu. The protein data represent the average of the values from two independent batches of UPIII. All numbers are in mole percentage (±s.d.).

*Asp + Asn.

†Glu+Gln.

‡Not determined.

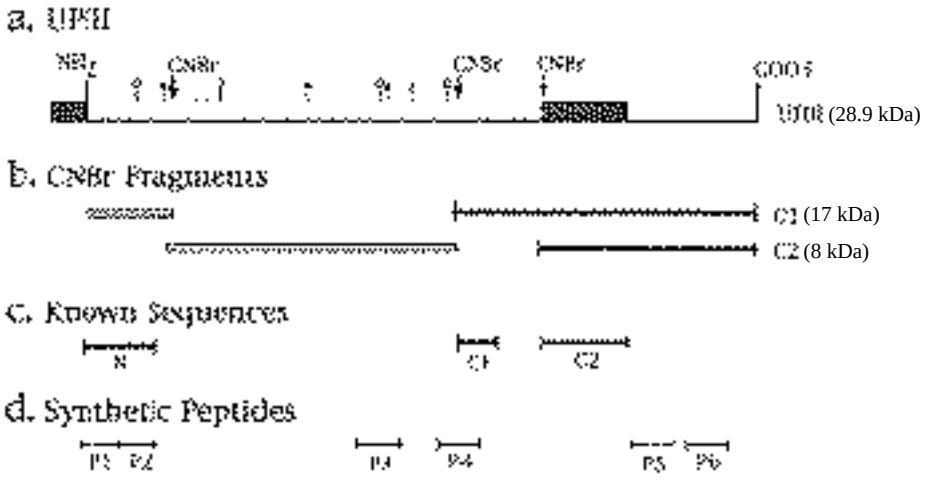

Fig. 5. A schematic representation of the predicted UPIII protein sequence showing the location of CNBr-fragments, known sequences and synthetic peptides. (a) Full-length amino acid sequence with a predicted molecular mass of 28.9 kDa. The first and second hatched bars represent the signal peptide sequence and a putative transmembrane domain, respectively. The N and C termini of mature UPIII are marked by NH_2 and COOH, respectively. The three potential cyanogen bromide cleavage sites (methionine) are as indicated. The N-terminal domain contains four potential *N*-glycosylation sites (open circles) and four cysteines (filled circles). (b) CNBr fragments. The two filled bars denote the C1 and C2 fragments from which some partial amino acid sequences were generated. The two open bars are fragments predicted from the cDNA sequence; the longer one, containing all three favorable *N*-glycosylation sites (see text), most likely corresponds to the diffuse 35 kDa band shown in Fig. 1b (asterisk). (c) Known amino acid sequences obtained from intact UPIII (N), and C1 and C2 CNBr-peptides. (d) Chemically synthesized oligopeptides. Note that while peptides 1, 2 and part of 4 were synthesized on the basis of known protein sequences (compare with c), peptides 3, 5 and 6 were based solely on cDNA-predicted sequences.

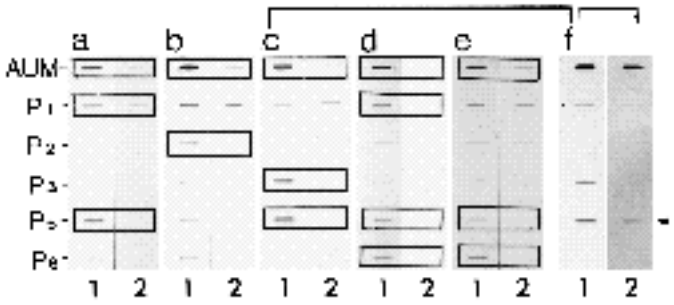

Fig. 6. Peptides chemically synthesized according to predicted UPIII sequence are actually recognized by some antibodies raised against intact UPIII. Fifty micrograms of AUM and each of the synthetic peptides (P 1, 2, 3, 5, and 6) were suspended in phosphate-buffered saline, blotted onto nitrocellulose sheet, and reacted with various antisera (lanes 1) or corresponding preimmune sera (lanes 2). Antisera used are: (a) a rabbit antiserum made against native AUM; (b) a chicken anti-AUM; (c) another rabbit anti-AUM; (d) a chicken antiserum raised against SDS-denatured UPIII; and (e) a rabbit antiserum raised against SDS-denatured UPIII. Immunoreactions that are significantly higher than preimmune sera are boxed. Note that although all the antisera react with AUM, they show different reactivities against various synthetic peptides. (f) Monospecific antibodies against peptide 5 were affinity-purified from a rabbit anti-AUM antiserum (lanes c1 and f1). Note that the purified antibodies (lane f2) now recognize peptide 5 specifically (arrowhead).

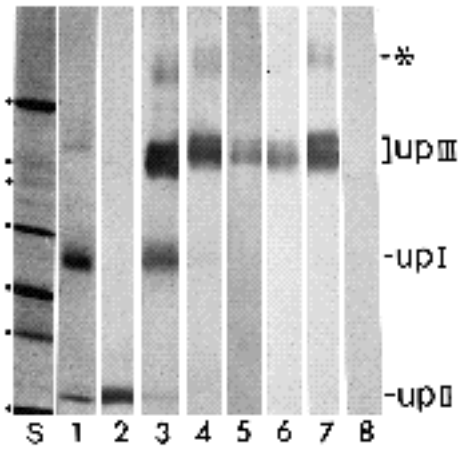

Fig. 7. Affinity-purified antibodies to synthetic peptide 5 react specifically with intact uroplakin III without crossreacting with UPI or II. Lane S shows the molecular mass standards including, from top to bottom, bovine serum albumin (66 kDa), egg albumin (45 kDa), glyceraldehyde-3-phosphate dehydrogenase (36 kDa), carbonic anhydrase (29 kDa), trypsinogen (24 kDa), trypsin inhibitor (20.1 kDa), and α-lactalbumin (14.2 kDa). Lane 1 shows the three uroplakins of gradient-purified AUM, which were electrophoretically transferred onto nitrocellulose sheet, and visualized by Fast Green staining. Similar sheets were stained with the following rabbit antibodies: lane 2, an antiserum raised against a synthetic peptide corresponding to an N-terminal sequence of the 15 kDa UP II (Yu et al., 1990). Lane 3 a rabbit anti-AUM antiserum that reacts with all three uroplakins; asterisk marks a minor 75 kDa band that contains UPIII and possibly UPI (see Fig. 9). Lane 4, a UPIII-specific antibody originally raised against native AUM; lane 5, an antiserum raised against electrophoretically purified UPIII; lanes 6 and 7, two independent batches of antibodies affinity-purified from a rabbit anti-AUM using synthetic peptide 5 of UPIII; and lane 8, a preimmune serum control. Note that the affinity-purified antibodies to peptide 5 (synthesized according to a cDNA-predicted UPIII sequence) react strongly and specifically with intact UPIII protein without crossreacting with the other two uroplakins.

9a and b). However, treatment with *N*-glycosidase F, an enzyme that cleaves all the major *N*-linked sugars including the complex and hybrid types, resulted in the complete disappearance of the 47 kDa species, which now ran at ~30 kDa (Fig. 9). *O*-glycosidase (with or without neuraminidase pre-treatment) had no effect on the 47 kDa UPIII (Fig. 9), even though in control experiments the enzyme cleaved efficiently *O*-linked sugars of fetuin (data not shown). These results clearly indicate that UPIII is heavily glycoslyated with *N*-linked sugars of the complex or hybrid

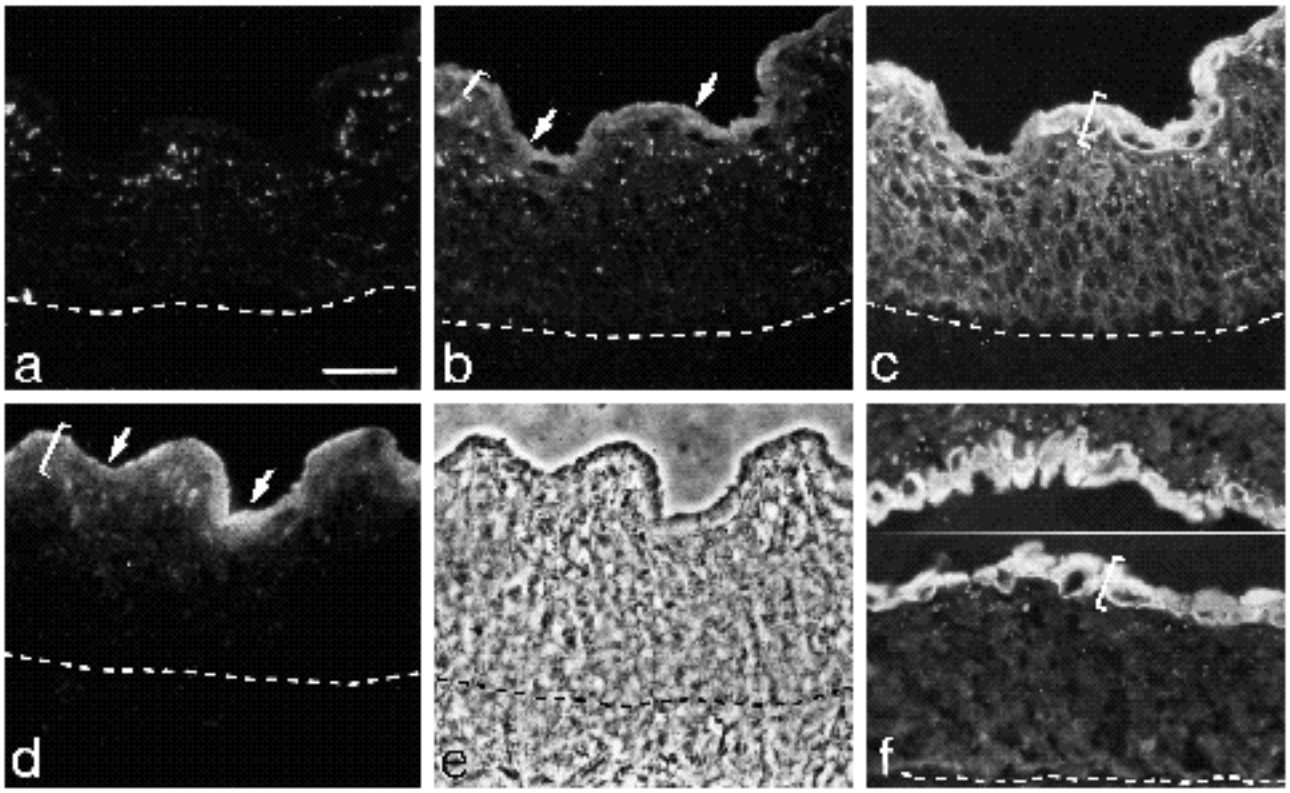

Fig. 8. Affinity-purified antibodies to synthetic peptide 5 stain specifically the apical surface of terminally differentiated bovine urothelial cells. Frozen sections (5 μm) of bovine bladder mucosa were stained with various antibodies by indirect immunofluorescence. (a) As a control, the specimen was stained with a preimmune serum showing the nonspecific staining of some cytoplasmic granules. (b) Affinity-purified rabbit antibodies monospecific for peptide 5 of UPIII (see Fig. 6f, lane 2; Fig. 7, lane 6). Note the specific staining of the apical surface (arrows) of the superficially located and terminally differentiated umbrella cells. (c) Double-staining of the same section in (b) for keratin using a mixture of mouse AE1 and AE3 antibodies (Woodcock-Mitchell et al., 1982). Note the staining of all urothelial cell layers, and the particularly intense cytoplasmic staining of the umbrella cells (bracketed). (d) An antiserum raised against intact UPIII (Fig. 7; lane 5); note the specific staining of apical surface similar to that produced by anti-peptide 5. (e) Phase-contrast picture of the same field as in (d). (f) A mouse monoclonal antibody to K8 keratin (courtesy of Dr. Roland Moll of the University of Mainz, Germany). Note, in two different fields, the selective staining of superficial umbrella cells (bracketed). The broken lines mark the basement membrane. Bar (a), 25 μm; all panels are of the same magnification.

type, and further establish the authenticity of our UPIII clone.

Uroplakin III has a signal peptide sequence and a potential transmembrane domain

As mentioned, the UPIII cDNA contains a single open reading frame of 287 amino acids. It bears two stretches of hydrophobic amino acids that scored >1.6 in the Kyte-Doolittle hydropathy algorithm (bracketed in Fig. 4; Fig. 10). The one located at the N terminus consists of 18 amino acids, and is clearly a pre-sequence. It begins with a segment of 11 hydrophobic amino acids, which may serve as a transmembrane signal sequence; and it ends with several polar amino acids with glycines occupying the −3 and −1 positions, thus fulfilling the "−3, −1 rule" of signal peptidase cleavage site (von Heijne, 1986, 1990). Although no positively charged region is present at the N terminus in this pre-sequence, the N-terminal methionine, which is unlikely to be formylated, may suffice (von Heijne, 1990). These features suggest strongly that this pre-sequence represents a signal peptide, which is removed during UPIII maturation.

Another stretch of 28 predominantly hydrophobic amino acids, long enough to span the membrane lipid bilayer, is located towards the C terminus. It divides the molecule into an N-terminal domain of 189 amino acids containing *all* the potential glycosylation sites and cysteine residues, and a C-terminal domain of 52 amino acids of predominantly hydrophilic residues (Figs 4, 5 and 10; and see below). Such a transmembrane configuration is consistent with our previous immunolocalization data showing that epitopes of UPIII could be found on both sides of the AUM (Wu et al., 1990). Although the precise boundaries of this transmembrane domain have not been defined, it is flanked by two (predicted) β-turns whose amino acid residues probably cannot form backbone hydrogen bondings and thus are unlikely to be embedded in the membrane (Fig. 10; Gennis, 1989).

The N-terminal domain of 189 amino acids is predicted to be enriched in β-sheet (Fig. 10). Of the four Asn-linked glycosylation consensus sequences located here (Figs 4 and 10), the efficiency of the first one (Asn34) is questionable, given a proline interrupting the Asn and Thr (Struck and Lennarz, 1980). Asn74 and Asn170 coincide with predicted

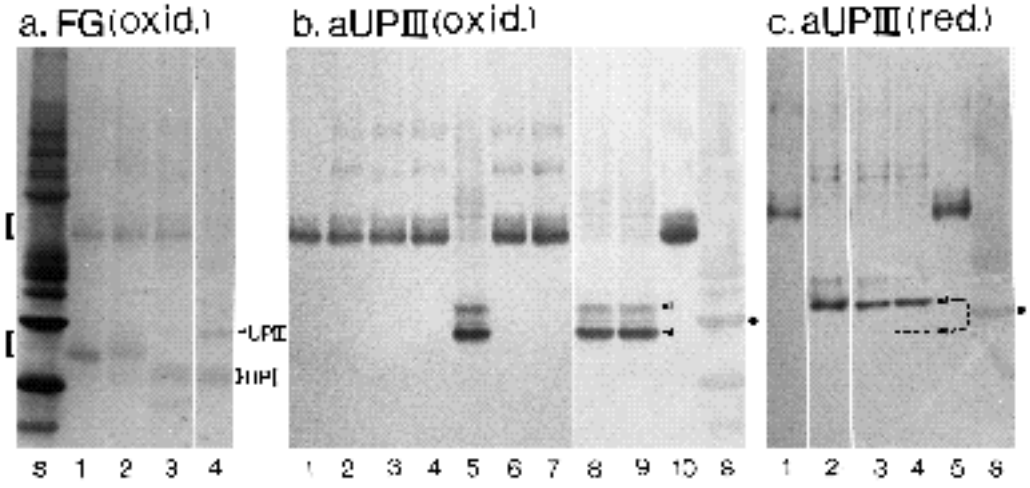

Fig. 9. Deglycosylation of UPIII reveals a core protein of ~30 kDa. Uroplakins of gradient-purified AUM were dissolved in 0.1% SDS, treated with glycosidases, and resolved by SDS-PAGE. The proteins were then electrophoretically transferred onto nitrocellulose paper, and stained with (a) Fast Green. Lane S shows the molecular mass standards as shown in Fig. 7; lane 1, control uroplakins; 2, another uroplakin control that has been incubated at 37°C for 16 h; 3, Endo-H treatment; 4, *N*-glycosidase F treatment. Note the decreased apparent size of UPI after both enzyme treatments, and that of UPIII only after *N*-glycosidase F treatment. (b) UPIII (without reduction) was visualized by anti-UPIII immunoblotting. Lanes 1, 2 and 10, control uroplakins. Lanes 3 through 9, uroplakins treated with the following glycosidases: 3, Endo-H; 4, neuraminidase; 5, *N*-glycosidase F; 6, *O*-glycosidase; 7, neuraminidase plus *O*-glycosidase; 8 *N*-glycosidase F plus *O*-glycosidase; and 9, neuraminidase, *N*-glycosidase F plus *O*-glycosidase. Note the decreased apparent size of UPIII (from 47 kDa to about 28 kDa) after *N*-glycosidase F treatment. Asterisk and arrow mark the 75 kDa and 95 kDa UPIII-containing crosslinked species (see text and Fig. 7). (c) UPIII was reduced with 10% β-mercaptoethanol (BME) before SDS-PAGE and immunoblotting. Lanes 1 and 5 are control uroplakins. Lanes 2, 3 and 4 correspond to lanes 5, 8 and 9 of (b), respectively. S denotes molecular mass standards partially reduced by BME diffused from the BME-containing sample of an adjacent (left) lane (5); protein bands that are skewed (due to decreased SDS-PAGE mobility) are clearly BME-sensitive. The filled circles in (b) and (c) mark the position of carbonic anhydrase (29 kDa), whose electrophoretic mobility is not altered by reduction and thus provides a reliable molecular mass reference. Filled and open arrowheads indicate the positions of oxidized and reduced forms of deglycosylated UPIII, respectively. Note the complete conversion of the lower molecular mass 28 kDa UPIII band by reduction to the 30 kDa position.

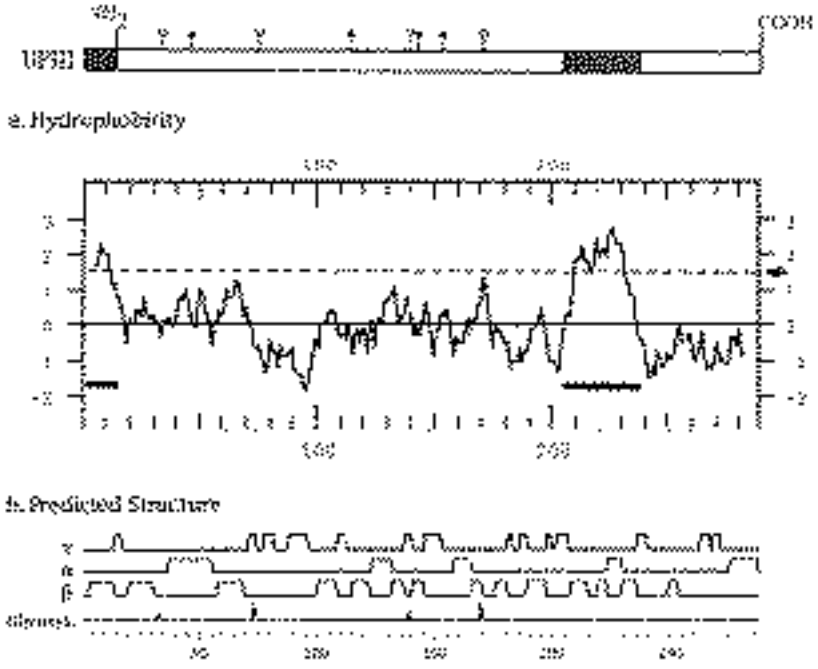

Fig. 10. The hydropathy plot and predicted secondary structure of UPIII. The top bar (as shown in Fig. 6) is included to facilitate the alignment of various structural domains. (a) A hydropathy plot using the Kyte-Doolittle algorithm with a window width of 13 amino acids; a stretch of >20 amino acids with an average hydrophobicity scale of >1.6 suggests of a transmembrane domain (marked by thick horizontal bars; Garnier and Robson, 1989). (b) Prediction of secondary structures using the Garnier-Osguthorpe-Robson (GOR) method. Abbreviations are: T (β-turns); α (α-helices); β (β-sheets); Glycosyl. (potential *N*-glycosylation sites).

β-turns, and thus might be preferentially glycosylated (Fasman, 1989). Although our data left no doubt that some of these sites must be glycosylated (Fig. 9), the precise extent of glycosylation of individual sites remains to be determined.

Another feature of the N-terminal domain is that it contains all four cysteines of the mature UPIII (Figs 4, 5 and 10). In this regard, it should be noted that UPIII, without being reduced, can be resolved by SDS-PAGE into a doublet (Figs 7 and 9). The resolution of these two bands was

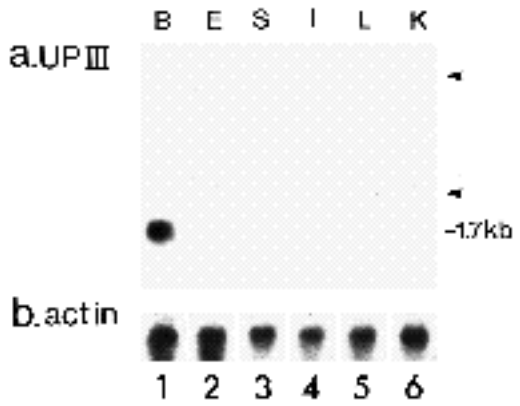

Fig. 11. Northern blot analysis establishes the urothelium-specificity of UPIII messenger RNA. Ten micrograms of total RNAs from various bovine tissues were resolved on a denaturing agarose gel, transferred to nitrocellulose paper, and probed with the 1.7 kb UPIII cDNA. The RNAs were from bovine (1) urothelium, (2) esophageal epithelium, (3) snout epithelium, (4) intestine mucosa, (5) liver and (6) kidney. (a) Shows the hybridization with UPIII cDNA. (b) As a control, the same blot as in (a) was washed under stringent conditions to remove the UPIII probe, and then re-probed with human β-actin cDNA. Note the detection of a 1.7 kb urothelium-specific UPIII mRNA. The arrowheads mark the positions of 28 S and 18 S ribosomal RNA.

greatly improved following deglycosylation (marked with open and closed arrowheads in Fig. 9b). These two bands share the same N-terminal sequence (Wu et al., unpublished), and both react with antibodies against two separate C-terminal synthetic peptides (P5 as shown in Fig. 7; P6, data not shown) suggesting that they are both intact UPIII molecules. Interestingly, the lower molecular mass band can be quantitatively converted to the higher molecular mass band by reduction (Fig. 9c), suggesting that the former contains intramolecular disulfide bonds whose reduction results in the unfolding and thus a slower electrophoretic mobility. The same immunoblotting experiment revealed a minor 75 kDa UPIII-related band (Fig. 9b), which disappeared upon reduction, suggesting that it represents a UPIII disulfide crosslinked with either another UPIII or some other proteins. Finally, another minor 95 kDa UPIII-related band was identified (Fig. 9b), which was unaffected by reduction, indicating that it represents a (non-cystine) crosslinked UPIII.

The C-terminal domain contains 52 amino acids with no cysteine residues or potential glycosylation sites. It is highly hydrophilic with nearly 30% charged amino acids (compared with 15% of the N-terminal domain). It bears clusters of serine, threonine and tyrosine residues forming multiple repeats of T/S-X-(+/−) or T/S-X-X-T/S, which are the consensus sequences of some protein kinases (Roach, 1991). This domain also contains an RGD sequence, a cell adhesion motif present in many extracellular matrix proteins (Ruoslahti and Pierschbacher, 1987; Yamada, 1991). Its functional significance in UPIII remains unclear, however, given its cytoplasmic location (see below).

Tissue-specific and differentiation-related expression of uroplakin III

Using antibodies to UPIII, we have previously shown that this protein is urothelium-specific (Wu et al., 1990). To see whether this tissue specificity is regulated on mRNA level, we performed Northern blot analysis. The UPIII cDNA probe detects a 1.7 kb mRNA in bovine bladder epithelium; however, this message is undetectable in esophageal epithelium, snout epithelium, intestinal epithelium, liver and renal cortex, indicating that it is urothelium-specific (Figs 2c and 11). Moreover, as mentioned earlier, antibodies to intact UPIII as well as several synthetic peptides of UPIII decorate the superficially located umbrella cells (Fig. 8). Taken together, these results strongly suggest that UPIII is an excellent marker for an advanced stage of urothelial differentiation.

DISCUSSION

We describe in this paper the molecular characterization of uroplakin III, a novel integral membrane protein that is expressed in a bladder-specific and differentiation-dependent manner and is a major component of the asymmetric unit membrane. The 1.7 kb cDNA encodes 287 amino acids with a calculated molecular mass of 28.9 kDa. The authenticity of this clone and its deduced protein sequence is supported by several lines of evidence including partial amino acid sequences (Figs 1, 4 and 5), amino acid composition (Table 1), antibody reactivity (Figs 6 and 7), tissue-specificity (Figs 2 and 11), and differentiation-dependent expression (Fig. 8b). In addition, deglycosylation of UPIII yielded a core protein of ~30 kDa, which matches perfectly the cDNA-predicted molecular mass (Fig. 9). A search of the sequences available in Genbank using the TFASTA program (Pearson and Lipman, 1988) revealed no significant structural homology between UPIII and any known proteins.

Uroplakin III is a type I transmembrane glycoprotein

Our data indicate that UPIII in its nascent form contains two hydrophobic sequences (Figs 4 and 10). The one located at the N terminus is clearly a pre-sequence, which may serve as a signal peptide involved in initiating the translocation of the N-terminal domain into the lumen of the ER (Rapoport and Wiedmann, 1985; von Heijne, 1986, 1990; Gierasch, 1989). The other stretch of hydrophobic sequence most likely serves as a transmembrane domain (Figs 4, 5 and 10), which presumably can also function as a stop-transfer sequence, thereby preventing the C-terminal domain from further translocation (von Heijne and Gavel, 1988; Jennings, 1989). The presence of these two hydrophobic sequences predicts that UPIII adopts an N-exo/C-cyto topology (the “Type I topology”; von Heijne and Gavel, 1988). This N-exo/C-cyto topology is strongly supported by the fact that all the potential N-linked glycosylation sites, some of which we know must be glycosylated (Fig. 9), reside in the N-terminal domain (Figs 4 and 5). Since glycosyltransferases are known to be localized in the lumen of rough endoplasmic reticulum and Golgi com-

plex, the glycosylated N-terminal domain of UPIII is almost certainly located in the lumen (Hirschberg and Snider, 1987). Even though such a N-exo/C-cyto topology runs against the "positive-inside" rule, since the N-terminal end of UPIII carries at least three net positive charges (von Heijne and Gavel, 1988; Hartmann et al., 1989; Paterson and Lamb, 1990), the combined effects of signal peptide and stop transfer sequence might be sufficient to over-ride such a factor.

Although the amino acid sequence of UPIII does not seem to be homologous to those of any known integral membrane proteins, its overall structure bears a striking resemblance to the core proteins of some cell surface proteoglycans including syndecan (Saunders et al., 1989), *N*-syndecan (Carey et al., 1992), epican (Kugelman et al., 1992), and a TGF-β receptor subunit (Lopez-Casillas et al., 1991). The core proteins of most of these proteoglycans are relatively small (around 30 kDa), and possess a single transmembrane domain with a longer N-terminal sequence exposed on the luminal surface and a much shorter (3 to 5 kDa), hydrophilic C-terminal sequence located on the cytoplasmic side. The extracellular domains of these proteins are usually heavily glycosylated. Moreover, as in UPIII (Fig. 4), most of these molecules contain an extracellular dibasic sequence near the transmembrane domain; enzymatic cleavage of this site can thus release the N-terminal domain (Saunders et al., 1989). Such a dibasic cleavage site has also been found in some other surface membrane proteins such as protein tyrosine phosphatase LAR (Streuli et al., 1992). Since we have not yet seen a cleaved, soluble form of UPIII, we do not know whether such a cleavage of UPIII actually takes place.

The role of uroplakin topology in establishing the sidedness of "asymmetric" unit membrane

A unique feature of urothelial apical plaques is that their luminal leaflets (8 nm) are almost twice as thick as their cytoplasmic leaflets (4 nm; Hicks, 1965; Koss, 1969; Staehelin et al., 1972; Robertson and Vergara, 1980). Although it is generally felt that the staining of the leaflets of plasma membranes is mainly due to lipids, the thickened appearance of the outer leaflet of the AUM is known to be due to the 12 nm protein particles (see Introduction). The topology of UPIII as described here, and that of the other two uroplakins (J.-H. Lin, J. Yu, X.-R. Wu, G. Kreibich, and T.-T. Sun, unpublished data), supports and extends this idea. Our present data on UPIII revealed that the molecular mass of its luminal domain (189 residues) greatly exceeds that of the cytoplasmic domain (52 residues; Figs 4 and 10). Moreover, the luminal domain contains additional sugar residues of up to 20 kDa (Fig. 10). Thus, the total molecular mass of the luminal domain is much greater than that of the cytoplasmic domain. Such an asymmetry of mass distribution is even more striking for the other two uroplakins. Specifically, the 27 kDa uroplakin I has four transmembrane domains with two long hydrophilic loops both extending into the luminal space, thereby having extremely small cytoplasmic domains (J. Yu, X.-R. Wu, T.-T. Sun, unpublished data). The 15 kDa uroplakin II is, like UPIII described here, a type I transmembrane protein; however, its hydrophobic transmembrane domain is located at its very C terminus and again there is practically no cytoplasmic domain (J.-H. Lin, X.-R. Wu, G. Kreibich, and T.-T. Sun, unpublished data). Thus, the bulk of all three uroplakins projects into the luminal space, most likely interacting tightly with one another to form the 12 nm protein particles. Although it was thought previously that the protein particles were "floating" on the outer leaflet without any of their parts penetrating into the inner leaflet (Staehelin et al., 1972; Robertson and Vergara, 1980; also see Vergara and Chesnut, 1983), our data strongly suggest that all three uroplakins are integral membrane proteins with their particle-forming (luminal) domains anchored to AUM via their transmembrane domains that penetrate through the lipid bilayer. This model suggests that the asymmetrical topology of uroplakin molecules may account in part for why the outer leaflet of AUM is much thicker than the inner leaflet. Our model also suggests that, in the earlier studies of Caruthers and Bonneville (1980), trypsinization probably cleaves off the extracellular particle-forming domains of uroplakins, leaving behind their transmembrane domains still embedded in the lipid bilayer. Experiments are in progress to test these hypotheses.

The structural role of the 47 kDa UPIII in the formation of 12 nm AUM particles is not yet known. However, with the availability of milligram quantities of highly purified AUM containing quasi-crystalline protein structure, and antibodies against multiple epitopes of UPIII (Figs 5 and 6, and data not shown), we are in the process of mapping in detail the subdomain organization of UPIII and to define how UPIII interacts with the other two uroplakins in forming AUM.

UPIII contributes to the formation of urothelial glycocalyx

Ultrastructural studies showed that the luminal surface of urinary bladder epithelium is covered with a layer of glycocalyx, which is believed to play an important role in preventing bacterial adherence (Parsons et al., 1980; Pauli et al., 1983; Ruggieri et al., 1992). However, the biochemical nature and cellular origin of this glycocalyx are unknown. Our data indicate that UPIII, which is one of the three major proteins of the AUM, contains up to 20 kDa *N*-linked, complex-type sugar residues (Fig. 9). These carbohydrates of UPIII may constitute a significant portion of the surface glycocalyx. The other two uroplakins are much less important contributors to the glycocalyx, since UPI contains only a small amount (~2 kDa) of high-mannose-type sugars, while mature UPII does not seem to be glycosylated at all (Fig. 9 and data not shown).

UPIII may serve as an anchorage for an underlying cytoskeleton

Although earlier ultrastructural data have firmly established that the cytoplasmic side of AUM interacts with the cytoskeleton (see Introduction), little was known about the membrane protein that is involved in this interaction. From our studies, we now know that morphologically intact AUM contains only three integral membrane proteins (Wu et al., 1990; Yu et al., 1990), and that the 47 kDa uroplakin III is the only uroplakin that possesses a significant cytoplasmic tail (52 amino acid long) that can potentially interact

with the cytoskeleton. The sequence of this cytoplasmic domain does not resemble any of the membrane proteins known to be involved in membrane cytoskeleton interaction, and is therefore novel. This sequence contains many potential phosphorylation sites, raising interesting possibilities as to how this interaction may be regulated.

While there is no doubt that a cytoskeletal element interacts with AUM, relatively little is known about this filamentous component. Ultrastructural studies indicate that these cytoskeletal structures are 6-10 nm in diameter, and Bonneville and Chlapowski (1981) have shown, by isolating the luminal cells and extracting the cytoskeleton, that these are most likely keratin filaments. Consistent with this idea, we have found recently that highly purified AUMs still contain a small amount of a 50 kDa, AE1-positive acidic keratin (Wu et al., unpublished). In addition, a new keratin polypeptide (keratin 20) has recently been shown to be preferentially expressed in the cortical region of urothelial umbrella cells (Moll et al., 1990, 1992). Finally, a simple epithelial keratin, K18, was found to be stained selectively in the superficial umbrella cells (also see the similar staining pattern of K8 in Fig. 8f), even though some other monoclonal antibodies to this same keratin were found to stain the entire urothelium (Debus et al., 1982; Schaafsma et al., 1989). This result strongly implies that in umbrella cells certain epitopes of K18 (and possibly K8; see Fig. 8f) undergo conformational changes and thus become "unmasked". The fact that these keratin changes in umbrella cells coincide with the expression of uroplakins raises the possibility that such changes may be coordinated and involved in AUM-cytoskeleton interaction. Another possibility is that some membrane-linking proteins may be involved in mediating the uroplakin-cytoskeleton interaction. Experiments are underway to test these possibilities.

Our sincere thanks go to Gert Kreibich of the NYU Cell Biology Department for his advice and suggestions, and to William Lane of the Harvard Microchemistry Facility for his expert assistance in generating the partial amino acid sequences; our interactions with them have been most helpful, educational and enjoyable. We also thank Jun Yu and Jun-Hsiang Lin for stimulating discussions, Pamela Cowin and Michael O'Guin for critical reading of the manuscript, and Irwin M. Freedberg and Pablo Morales for their continued interest and support. This work was supported in part by grants from the National Institutes of Health (DK39753 and AM39749). Computing was supported by NSF (DIR-8908095). Genebank sequence accession number: L19542.

REFERENCES

Bonneville, M. A., and Chlapowski, F. J. (1981). Characterization of intermediate-type filaments in urothelium. *J. Cell Biol.* **91**, 234a.

Bryan, G. T. (1983). Etiology and pathogenesis of bladder cancer. In *The Pathology of Bladder Cancer*, vol. I. (ed. G. T. Bryan and S. M. Cohen), pp. 1–9. CRC Press, Boca Raton, Florida.

Carey, D. J., Evans, D. M., Stahl, R. C., Asundi, V. K., Conner, K. J., Garbes, P. and Cizmeci-Smith, G. (1992). Molecular cloning and characterization of N-syndecan, a novel transmembrane heparan sulfate proteoglycan. *J. Cell Biol.* **117**, 191-201.

Caruthers, J. S. and Bonneville, M. A. (1980). The asymmetric unit membrane (AUM) structure in the luminal plasma membrane of urothelium: the effect of trypsin on the integrity of the plaques. *J. Ultrastruct. Res.* **71**, 288-302.

Chirgwin, J. M., Przybyla, A. E., MacDonald, R. J. and Rutter, W. J. (1979). Isolation of biologically active ribonucleic acid from sources enriched in ribonuclease. *Biochemistry* **18**, 5294-5299.

Chlapowski, F. J., Bonneville, M. A. and Staehelin, L. A. (1972). Luminal plasma membrane of the urinary bladder. II. Isolation and structure of membrane components. *J. Cell Biol.* **53**, 92-104.

Cunha, G. R., Fujll, H., Neubauer, B. L., Shannon, J. M., Sawyer, L. and Reese, B. A. (1983). Epithelial-mesenchymal interactions in prostatic development. I. Morphological observations of prostatic induction by urogenital sinus mesenchyme in epithelium of the adult rodent urinary bladder. *J. Cell Biol.* **96**, 1662-1670.

Debus, E., Weber, K. and Osborn, M. (1982). Monoclonal cytokeratin antibodies that distinguish simple from stratified squamous epithelia: characterization on human tissues. *EMBO J.* **1**, 1641-1647.

Devereux, J., Haeberli, P. and Smithies, O. (1984). A comprehensive set of sequence analysis programs for the vax. *Nucl. Acids Res.* **12**, 387-395.

Fasman, G. D. (1989). Protein conformational prediction. *Trends Biochem. Sci.* **14**, 295-299.

Garnier, J. and Robson, B. (1989). The GOR method for predicting secondary structures in proteins. In *Prediction of Protein Structure and the Principles of Protein Conformation* (ed. G. D. Fasman), pp. 417-465. Plenum Press, New York.

Gennis, R. B. (1989). *Biomembranes: Molecular Structure and Function.* pp. 118-119.Springer-Verlag, New York.

Gierasch, L. M. (1989). Signal sequences. *Biochemistry* **28**, 923-930.

Hartmann, E., Rapoport, T. A. and Lodish H. F. (1989). Predicting the orientation of eukaryotic membrane-spanning proteins. *Proc. Nat. Acad. Sci. USA* **86**, 5786-5790.

Hicks, R. M. (1965). The fine structure of the transitional epithelium of rat ureter. *J. Cell Biol.* **26**, 25-48.

Hicks, R. M. and Ketterer, B. (1969). Hexagonal lattice of subunits in the thick luminal membrane of the rat urinary bladder. *Nature* **224**, 1304-1305.

Hicks, R. M. and Ketterer, B. (1970). Isolation of the plasma membrane of the luminal surface of rat bladder epithelium, and the occurrence of a hexagonal lattice of subunits both in negatively stained whole mounts and in sectioned membranes. *J. Cell Biol.* **45**, 542-553.

Hirschberg, C. B. and Snider, M. D. (1987). Topography of glycosylation in the rough endoplasmic reticulum and Golgi apparatus. *Annu. Rev. Biochem.* **56**, 63-87.

Huggins, C. B. (1931). The formation of bone under the influence of epithelium of the urinary tract. *Arch. Surg.* **22**, 377-408.

Jennings, M. L. (1989). Topography of membrane proteins. *Annu. Rev. Biochem.* **58**, 999-1027.

Kaufman, D. L. and Evans, G. A. (1990). Restriction endonuclease cleavage at the termini of PCR products. *BioTechniques* **9**, 304-306.

Ketterer, B., Hicks, R. M., Christodoulides, L. and Beale D. (1973). Studies of the chemistry of the luminal plasma membrane of rat bladder epithelial cells. *Biochim. Biophys. Acta* **311**, 180-190.

Knutton, S. and Robertson, J. D. (1976). Regular structures in membranes: the luminal plasma membrane of the cow urinary bladder. *J. Cell Sci.* **22**, 355-370.

Koss, L. G. (1969). The asymmetric unit membranes of the epithelium of the urinary bladder of the rat: an electron microscopic study of a mechanism of epithelial maturation and function. *Lab. Invest.* **21**, 154-168.

Koss, L. G. (1975). Tumors of the urinary bladder. In *Atlas of Tumor Pathology*, 2nd series, Fasc. II. Washington, DC: Armed Forces Institute of Pathology.

Kugelman, L. C., Ganguly, S., Haggerty, J. G., Weissman, S. M. and Milstone, L. M. (1992). The core protein of epican, a heparan sulfate proteoglycan on keratinocytes, is an alternative form of CD44. *J. Invest. Dermatol.* **99**, 381-385.

Laemmli, U. K. (1970). Cleavage of structure proteins during the assembly of the head of bacteriophage T4. *Nature* **227**, 680-685.

Lathe, R. (1985). Synthetic oligonucleotide probes deduced from amino acid sequence data. Theoretical and practical considerations. *J. Mol. Biol.* **183**, 1-12.

Lonsdale-Eccles, J. D., Lynley, A. M. and Dale, B. A. (1981). Cyanogen bromide cleavage of proteins in sodium dodecyl sulfate/polyacrylamide gels. Diagonal mapping of proteins from epidermis. *Biochem. J.* **197**, 591-597.

Lopez-Casillas, F., Cheifetz, S., Doody, J., Andres, J. L., Lane, W. S. and Massague, J. (1991). Structure and expression of the membrane

proteoglycan betaglycan, a component of the TGF-β receptor system. *Cell* **67**, 785-795.

Maniatis, T., Fritsch, E. F. and Sambrook, J. (1982). *Molecular Cloning: A Laboratory Manual.* Cold Spring Harbor, New York: Cold Spring Harbor Laboratory.

Martin, F. H. and Castro, M. M. (1985). Base pairing involving deoxyinosine: implications for probe design. *Nucl. Acids Res.* **13**, 8927-8938.

Matsudaira, P. (1987). Sequence from picomole quantities of proteins elctroblotted onto polyvinylidene difluoride membranes. *J. Biol. Chem.* **262**, 10035-10038.

Melicow, M. M. (1974). Tumors of the bladder. A multifaceted problem. *J. Urol.* **112**, 467-478.

Minsky, B. D. and Chlapowski, F. J. (1978). Morphometric analysis of the translocation of transitional epithelial cells during the expansion-contraction cycles of mammalian urinary bladder. *J. Cell Biol.* **77**, 685-697.

Moll, R., Lowe, A., Laufer, J. and Franke, W. W. (1992). Cytokeratin 20 in human carcinomas. A new histodiagnostic marker detected by monoclonal antibodies. *Amer. J. Pathol.* **140,** 427-447.

Moll, R., Schiller, D. L. and Franke, W. W. (1990). Identification of protein IT of the intestinal cytoskeleton as a novel type I cytokeratin with unusual properties and expression patterns. *J. Cell Biol.* **111**, 567-580.

Mostofi, F. K. (1954). Potentialities of bladder epithelium. *J. Urol.* **71**, 705-714.

Neubauer, B. L., Chung, L. W. K., McCormick, K. A., Taguchi, O., Thompson, T. C. and Sunha, G. R. (1983). Epithelial-mesenchymal interactions in prostatic development. II. Biochemical observations of prostatic induction by urogenital sinus mesenchyme in epithelium of the adult rodent urinary bladder. *J. Cell Biol.* **96**, 1671-1680.

Parsons, C. F., Stauffer, C. and Schmidt, J. D. (1980). Bladder surface glycosaminoglycans: an efficient mechanism of environmental adaptation. *Science* **208**, 605-607.

Paterson, R. G. and Lamb, R. A. (1990). Conversion of a class II integral membrane protein into a soluble and efficiently secreted protein: multiple intracellular and extracellular oligomeric and conformational forms. *J. Cell Biol.* **110**, 999-1011.

Pauli, B. U., Alroy, J. and Weinstein, R. S. (1983). The ultrastructure and pathobiology of urinary bladder cancer. In *The Pathology of Bladder Cancer* (ed. G. T. Bryan and S. M. Cohen), pp. 42-140. CRC Press, Boca Raton, Florida.

Pearson, W. R. and Lipman, D. J. (1988). Improved tools for biological sequence comparison. *Proc. Nat. Acad. Sci. USA* **85**, 2444-2448.

Porter, K. R. and Bonneville, M. A. (1963). *An Introduction to the Fine Structure of Cells and Tissues*. Lea & Febiger, New York.

Rapoport, T. and Wiedmann, M. (1985). Application of the signal hypothesis to the incorporation of integral membrane proteins. *Curr. Top. Membr. Trans.* **24**, 1-63.

Roach, P. J. (1991). Multisite and hierarchal protein phosphorylation. *J. Biol. Chem.* **266**, 14139-14142.

Roberts, D. D., Leighton, J., Abaza, N. A. and Troll, W. (1974). Heterotopic urinary bladders in rats produced by an isograft inoculum of bladder fragments and air. *Cancer Res.* **34**, 2773-2778.

Robertson, J. D. and Vergara, J. (1980). Analysis of the structure of intramembrane particles of the mammalian urinary bladder. *J. Cell Biol.* **86**, 514-528.

Ruggieri, M. R., Balagani, R. K., Rajter, J. J. and Hanno, P. M. (1992). Characterization of bovine bladder mucin fractions that inhibit Escherichia coli adherence to the mucin deficient rabbit bladder. *J. Urol.* **148**, 173-178.

Ruoslahti, E. and Pierschbacher, M. D. (1987). New perspectives in cell adhesion: RGD and integrins. *Science* **238**, 491-497.

Sanger, F., Nicklen, S. and Coulson, A. R. (1977). DNA sequencing with chain-terminating inhibitors. *Proc. Nat. Acad. Sci. USA* **74**, 5463-5467.

Sarikas, S. N. and Chlapowski, F. J. (1986). Effect of ATP inhibitors on the translocation of luminal membrane between cytoplasm and cell surface of transitional epithelial cells during the expansion-contraction cycle of the rat urinary bladder. *Cell Tiss. Res.* **246**, 109-117.

Sarikas, S. N. and Chlapowski, F. J. (1989). The effect of thioglycolate on intermediate filaments and membrane translocation in rat urothelium during the expansion-contraction cycle. *Cell Tiss. Res.* **258**, 393-401.

Saunders, S., Jalknen, M., O'Farell, S. and Bernfield, M. (1989). Molecular cloning of syndecan, an integral membrane proteoglycan. *J. Cell Biol.* **108,** 1547-1556.

Schaafsma, H. E., Ramaekers, F. C. S., van Muijen, G. N. P., Ooms, E. C. M. and Ruiter, D. J. (1989). Distribution of cytokeratin polypeptides in epithelia of the adult human urinary tract. *Histochemistry* **91**, 151-159.

Schagger, H. and von Jagow, G. (1987). Tricine-sodium dodecyl sulfate-polyacrylamide gel electrophoresis for the separation of proteins in the range from 1 to 100 kDa. *Anal. Biochem.* **166**, 368-379.

Severs, N. J. and Hicks, R. M. (1979). Analysis of membrane structure in the transitional epithelium of rat urinary bladder: 2. The discoidal vesicles and Golgi apparatus: their role in luminal membrane biogenesis. *J. Ultrastruct. Res.* **69**, 279-296.

Staehelin, L. A., Chlapowski, F. J. and Bonneville, M. A. (1972). Luminal plasma membrane of the urinary bladder. I. Three-dimensional reconstruction from freeze-etch images. *J. Cell Biol.* **53**, 73-91.

Streuli, M., Krueger, N. X., Ariniello, P. D., Tang, M., Munro, J. M., Blattler, W. A., Adler, D. A., Disteche, C. M. and Saito, H. (1992). Expression of the receptor-linked protein tyrosine phosphatase LAR: proteolytic cleavage and shedding of the CAM-like extracellular region. *EMBO J.* **11**, 897-907.

Struck, D. K. and Lennarz, W. J. (1980). In *The Biochemistry of Glycoproteins and Proteoglycans* (ed. W. J. Lennarz), pp. 35-84. Plenum Publishing Corp., New York.

Sun, T.-T. and Green, H. (1978). Immunofluorescent staining of keratin fibers in cultured cells. *Cell* **14**, 468-476.

Surya, B., Yu, J., Manabe, M. and Sun, T.-T. (1990). Assessing the differentiation state of cultured bovine urothelial cells: elevated synthesis of stratification-related K5 and K6 keratins and persistent expression of uroplakin I. *J. Cell Sci.* **97**, 419-432.

Towbin, H., Staehelin, T. and Gordon, J. (1979). Electrophoretic transfer of proteins from polyacrylamide gels to nitrocellulose sheets: procedure and some applications. *Proc. Nat. Acad. Sci. USA* **76**, 4350-4354.

Vergara, J. A. and Chesnut, D. B. (1983). A spin label study of the urinary bladder luminal membrane. *Biochim. Biophys. Acta* **734**, 18-24.

Vergara, J. A., Longley, W. and Robertson, J. D. (1969). A hexagonal arrangement of subunits in membrane of mouse urinary bladder. *J. Mol. Biol.* **46**, 593-596.

Vergara, J., Zambrano, F., Robertson, J. D. and Elrod, H. (1974). Isolation and characterization of luminal membranes from urinary bladder. *J. Cell Biol.* **61**, 83-94.

von Heijne, G. (1986). A new method for predicting signal sequence cleavage sites. *Nucl. Acids Res.* **14**, 4683-4690.

von Heijne, G. (1990). The signal peptide. *J. Membr. Biol.* **115**, 195-201.

von Heijne, G. and Gavel, Y. (1988). Topogenic signals in integral membrane proteins. *Eur. J. Biochem.* **174**, 671-678.

Woodcock-Mitchell, J., Eichner, R., Nelson, W. G. and Sun, T.-T. (1982). Immunolocalization of keratin polypeptides in human epidermis using monoclonal antibodies. *J. Cell Biol.* **95**, 580-588.

Wu, X.-R., Manabe, M., Yu, J. and Sun, T.-T. (1990). Large scale purification and immunolocolization of bovine uroplakins I, II, and III. *J. Biol. Chem.* **265**, 19170-19179.

Yamada, K. (1991). Adhesive recognition sequences. *J. Biol. Chem.* **266**, 12809-12812.

Yu, J., Manabe, M. and Sun, T.-T. (1992). Identification of a 85-100-kD glycoprotein as a cell surface marker for an advanced stage of urothelial differentiation: association with interplaque ("hinge") area. *Epith. Cell Biol.* **1**, 4-12.

Yu, J., Manabe, M., Wu, X.-R., Xu, C., Surya, B. and Sun, T.-T. (1990). Uroplakin I: a 27-kD protein associated with the asymmetric unit membrane of mammlian urothelium. *J. Cell Biol.* **111**, 1207-1216.

(Received 30 April 1993 - Accepted 8 June 1993)