cDNA Cloning and Characterization of Sciellin, a LIM Domain Protein of the Keratinocyte Cornified Envelope*

Marie-France Champliaud, Robert E. Burgeson, William Jin, Howard P. Baden, and Pamela F. Olson‡

From the Cutaneous Biology Research Center, Massachusetts General Hospital, and the Department of Dermatology, Harvard Medical School, Charlestown, Massachusetts 02129

Sciellin is a precursor of the cornified envelopes of mammalian keratinizing tissues. We have cloned the cDNA encoding sciellin by screening a human keratinocyte expression library with a sciellin-specific monoclonal antibody. The composite cDNA of 2.35 kilobase pairs encodes a protein of 75.3 kDa with a pI of 10.09. The translated sequence has a central domain containing 16 repeats of 20 amino acids each that is rich in Gln and Lys residues, which are potential transglutaminase substrates, and a carboxyl domain, which contains a single LIM motif. Sciellin cDNA probes hybridize to bands of 3.4 and 4.4 kilobase pairs on Northern blots of cultured human keratinocyte RNA. The gene was mapped to human chromosome band 13q22 by fluorescence in situ hybridization. Radiation hybrid mapping demonstrated that sciellin is linked to the sequence tagged site marker WI-457 with a logarithm of the odds score of 7.77. In situ hybridization of human foreskin tissue sections demonstrated that sciellin is expressed in the stratum granulosum. Immunofluorescent staining with a polyclonal rabbit antibody made to a recombinant sciellin protein showed peripheral cytoplasmic localization in the upper cell layers of epidermis and in stratified squamous epithelia such as the oral cavity, esophagus, and vagina. Simple and columnar epithelia, with the exception of the amnion, showed no reaction.

The cornified envelope is a 15-nm-thick insoluble protein layer that is formed under the plasma membrane in the upper layers of epidermis and keratinizing stratified epithelium (1). It appears to play a major role in the physical barrier properties of the stratum corneum (2). Cornified envelope precursor proteins have structural roles in the stratum corneum cells, and their expression is associated with terminal differentiation. The envelope is formed from several precursor proteins by the calcium-dependent enzyme transglutaminase, which catalyzes formation of ε-(γ-glutamyl)lysine cross-links (3) that are resistant to proteolytic digestion. It has been postulated that involucrin is cross-linked at the plasma membrane as a first step in envelope assembly (4) followed by the less abundant precursors such as SPRR proteins, elafin, envoplakin, filaggrin, keratin filaments, and cystatin-α (5–11). Finally, loricrin covers the cytoplasmic side of the envelope (12).

It has been suggested that multiple components are necessary for envelope structure and function and that they have different roles. For example, involucrin acts as the framework for the attachment of other envelope components and is covalently linked to the lipids, which are important components of the barrier of stratum corneum cells (13). The SPRR/pancor- nulin proteins have been shown to act as molecular bridges and are able to cross-link with two different proteins (14). Loricrin is believed to form a meshlike scaffold, which is flexible as a result of its high glycine content and insoluble due to disulfide bonding and transglutaminase-catalyzed cross-linking to other cornified envelope components (12).

Gene knockout and mutation studies have been used to gather information on the function of epidermal proteins, with the keratins being the most well known (15, 16). Only one null mutation study of an envelope-related protein has been reported, a loricrin knockout mouse (17). Heterozygous mice were normal, whereas homozygotes have abnormal skin during the first few days, but the animals appeared structurally normal as adults. However, the mice have a defect in barrier function and respond abnormally to the application of irritants such as 12-O-tetradecanoylphorbol 13-acetate. Mutations of the loricrin gene have been observed in patients with a rare autosomal dominant palmoplantar keratoderma, Vohwinkel’s keratoderma (19), as well as in progressive, symmetric erythrokeratoderma (20). Mutations in the epidermal transglutaminase gene result in the human disease lamellar ichthyosis, which is characterized by a thickened stratum corneum, disturbed epidermal keratinization, and inflammatory changes (18). Overexpression of human loricrin in transgenic mice resulted in no detectable phenotype, even though all of the expressed protein was incorporated into the cornified envelope (21). The assembly of the CE was not interfered with, supporting the hypothesis that loricrin is deposited very late during CE assembly. However, overexpression of human involucrin in transgenic mice did result in an abnormal epidermal phenotype (22). Heterozygous mice appeared normal, although their hair stood erect. A subset of homozygous animals had reduced birth weight, a scaly epidermis, the onset of their first hair cycle was delayed

*This work was supported in part by a generous gift from Barbara Jordan, by United States Public Health Service Grant AR35689 (to R. E. B.), and by the Cutaneous Biology Research Center through the Massachusetts General Hospital/Shiseido Co. Ltd. Agreement. The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked “advertisement” in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

The nucleotide sequence(s) reported in this paper has been submitted to the GenBank™/EBI Data Bank with accession number(s) AB045941.

‡To whom correspondence should be addressed: MGH-East CBRC, Bldg. 149, 13th St., Charlestown, MA 02129. Fax: 617-726-4453; E-mail: pamela Olson@cbrc.mgh.harvard.edu.

1 The abbreviations used are: CE, cornified envelope; Ab, antibody; bp, base pair(s); FISH, fluorescence in situ hybridization; HED, hidrotic ectodermal dysplasia; kb, kilobase pair(s); LOD, logarithm of the odds; mAb, monoclonal antibody; pAb, polyclonal antibody; PAGE, polyacrylamide gel electrophoresis; PCR, polymerase chain reaction; RACE, rapid amplification of cDNA ends; RT-PCR, reverse transcription-polymerase chain reaction; SHGC, Stanford Human Genome Center; STS, sequence tagged site.
Sciellin

Northern Blotting—Total RNA was isolated using the RNeasy kit (Qiagen) from human foreskin keratinocytes cultured in high calcium medium and also from postpartum human placenta, carefully dissected into amniot, chorion, villous, and non-villous layers. A total of 25 μg/lane was electrophoresed through a 1% agarose gel, blotted onto Duralon (Stratagene, La Jolla, CA) and hybridized using the NorthernMax kit (Ambion, Austin, TX) according to the enclosed protocols. \[^{[2]}\] PdCTP-labeled probes were generated from gel-purified DNA using the Rediprime random primer labeling kit (Amersham Pharmacia Biotech). The probe for the keratinocyte and placenta blot (Fig. 4A) was generated from the EcoRI insert of cDNA clone SC4. The probe for the 4.4 kb sciellin message was generated by 3' RACE from clone 86A using primer MF175 (GCAACCGGCTGATGATTGTG). Human Normal Tissue Blots I and II (Northern Territory total RNA blots; Invitrogen, Carlsbad, CA) were hybridized with a \[^{[2]}\] UTP-labeled antisense RNA probe made from pPO325 (sciellin cDNA nucleotides 830–2310 subcloned into bluScript II KS') using the RNA transcription kit (Stratagene). The blots were hybridized and washed under stringent conditions using the NorthernMax buffer system (Ambion), according to the supplied protocols.

Protein Expression and Polyclonal Antibody Generation—Sciellin cDNA, corresponding to cDNA clone SC4 (Fig. 1), nucleotides 1120–1736, was amplified by RT-PCR from human differentiated keratinocyte RNA using adaptor primers PFO235 (GTTCATATGGAAAATACCGATATCGAAAGAA) and PFO238 (TCCCGCGGTACTCTGAGCTGTGGTACGGAGC) and subcloned into the NdeI and SacII sites of the T7 expression vector pET-15b (Novagen), which had been modified by replacement of the BarnHI site with a SacI site (a gift of Manuel Koch). The resulting fusion protein construct, pPO324, was transformed into Novabact(DE3) competent cells (Novagen), and the fusion protein SC4 was expressed and purified using the His-Bind resin and His-Bind buffer system (Novagen, Madison, WI) according to manufacturer's protocols. Polyclonal antiserum was raised in two rabbits by Charles River Pharmaservices using 0.1 mg of SC4 fusion protein/injection.

Radiation Hybrid Mapping—The Stanford G3 Radiation Hybrid Mapping Panel (Research Genetics, Huntsville, AL) was screened by PCR using two primer pairs derived from the sciellin genomic DNA sequence. MF122 (TTTGGGAGCTGTTTGTGCTTT) and MF124 (GCTCTCTGGAGACTAAAAATATGCTT), and MF79 (TTCCAGGGGCTACCTTGATGAGCT) and PFO2980 (AGGGCGCAATAGGCTGAGAC). The G3 panel was also screened using two STS markers that were linked to Clouston's hidrotic ectodermal dysplasia (HED), D13S141 and D13S175 (Research Genetics). These markers were ordered relative to the SHGC framework markers of the G3 RH map v2.0 using the statistical analysis program RHMAP.\[^{3}\]

In Situ Hybridization to Tissue Sections—Sense and antisense digeRNA probes were transcribed from the plasmids pPOSC5 and pPOK5 using the RNA transcription kit (Stratagene) and DIG RNA labeling mix (Boehringer Mannheim). The RT-PCR amplification product of the keratin K5 mRNA from nucleotides 73–1076 (33) was subcloned into bluScript KS+ to create pPOK5, and clone SC5 was subcloned into pluScript KS+ to create pPOSC5. Sections (10 μm) were cut from human foreskin that had been fixed in 4% paraformaldehyde and frozen in OCT compound (TissueTek, Miles Inc., Elkhart, IN). Tissue sections were hybridized overnight at 60 °C in a 50% formamide hybridization buffer. Digoxigenin was detected using anti-digoxigenin alkaline phosphatase Fab fragments (Boehringer Mannheim) according to manufacturer’s protocols.

Protein Purification—The purification of sciellin was done as follows. Human placenta was dissected to collect only the amniotic membrane. These were then frozen in liquid nitrogen, ground in a Waring blender (Waring Product Division, New Hartford, CT), and resuspended in 0.1 M citric acid monohydrate (100 ml for 50 g of tissue) containing 625 mg/liter N-ethylmaleimide and 150 mg/liter phenylmethylsulfonyl fluoride. The suspension was incubated at 4 °C with stirring overnight. The soluble fraction was collected following centrifugation (30,000 g, 30 min). Precipitated protein was redissolved to buffer A (150 mM NaCl, 2 mM EDTA, 10 mM Tris-HCl, pH 7.8) containing 1% SDS. At this point, the protein was dialyzed directly by cycle sequencing after excision from a TAE-ETBr agarose gel and purification using the QIAquick gel extraction kit (Qiagen, Chatsworth, CA).

The sequencing project was maintained and analyzed using the Wisconsin Package version 9.1 (Genetics Computer Group (GCG), Madison, WI). Data base comparisons were made using BLAST v2.0 (31). The protein sequence was analyzed using the program PSORTII (32).

by 5 days, and their hair also stood erect. Overexpression of involucrin caused abnormalities in the structure and function of the CE and hair follicles.

The protein sciellin was identified in earlier studies by a monoclonal antibody generated from mice immunized with cornified envelopes purified from cultured human keratinocytes (23, 24). Sciellin was extracted from keratinocytes with a neutral buffer when a chaotrophic agent such as urea or SDS was added, but not when non-ionic detergents were present (24). The protein remained in solution when a urea extract was dialyzed against a neutral buffer with 0.3 M NaCl, but precipitated when the salt was omitted. The mobility of the protein in SDS-PAGE was identical with or without a reducing agent in the lysis buffer, indicating little if any intermolecular disulfide bonding of the molecule. The protein in human keratinocytes and epidermis had a molecular weight of 82,000 and pl of about 8.0, but with charge heterogeneity. The antibody identified proteins of similar molecular weight in other animals by Western blot. The antibody stained the periphery of keratinocytes in the granular layer, and evidence was presented that the protein was a substrate for transglutaminase and cross-linked into the CE. The protein was present in stratified squamous epithelium but not simple epithelium. In the hair follicle, it was found in the upper part of the outer root sheath in the inner most layer, as well as in the periphery of cells of the inner root sheath. In the nail it was demonstrated in the nail fold, matrix, and bed (25). Involutin and SPRR proteins were present in the inner root sheath, but had a cytoplasmic location (23, 26), and loricrin was absent (27, 28). Their distribution in the nail was limited to the matrix.

Here we report the cloning and characterization of the human sciellin cDNA. The cDNA encoded a unique protein of 75.3 kDa, which is characterized by internal repeat units and a man sciellin cDNA. The cDNA encoded a unique protein of 75.3 kDa, which is characterized by internal repeat units and a man sciellin cDNA. The cDNA encoded a unique protein of 75.3 kDa, which is characterized by internal repeat units and a man sciellin cDNA. The cDNA encoded a unique protein of 75.3 kDa, which is characterized by internal repeat units and a man sciellin cDNA. The cDNA encoded a unique protein of 75.3 kDa, which is characterized by internal repeat units and a man sciellin cDNA. The cDNA encoded a unique protein of 75.3 kDa, which is characterized by internal repeat units and a man sciellin cDNA. The cDNA encoded a unique protein of 75.3 kDa, which is characterized by internal repeat units and a man sciellin cDNA. The cDNA encoded a unique protein of 75.3 kDa, which is characterized by internal repeat units and a man sciellin cDNA. The cDNA encoded a unique protein of 75.3 kDa, which is characterized by internal repeat units and a man sciellin cDNA. The cDNA encoded a unique protein of 75.3 kDa, which is characterized by internal repeat units and a man sciellin cDNA.
Sciellin was 2347 bp, including 86 bp of 5'-untranslated sequence and 257 bp of 3'-untranslated sequence. The nucleotide sequence surrounding the AUG initiator codon fits the consensus sequence for the initiation of translation by eukaryotic ribosomes (37). The 2004-bp open reading frame encoded a protein with a sequence for the initiation of translation by eukaryotic ribosomes surrounding the AUG initiator codon fits the consensus sequence (37). The 2004-bp open reading frame encoded a protein with a sequence for the initiation of translation by eukaryotic ribosomes. The predicted isoelectric point of the unmodified protein as calculated from sequence information was 10.09.

The amino acid sequence of sciellin, which was translated from the cDNA sequence, is shown in Fig. 2 and is 668 residues long. Comparison of the sciellin protein sequence with the GenBank data base using the program BLAST showed that sciellin was a unique protein. The protein was generally hydrophilic and did not contain any regions likely to form either a signal peptide or a transmembrane domain, consistent with its cytoplasmic localization by immunohistochemistry.

The sciellin protein may be divided into three domains based upon secondary structure predictions such as those generated by the program Plotstructure (GGG): an amino-terminal domain, a central domain composed of repeating units, and a carboxyl domain containing a single LIM motif. Pro was evenly distributed throughout the molecule, except for an enriched "hinge" region before the first repeat unit, starting at residue 165. There were 16 inexact repeat units from residues 231 to 543, of which 12 were 20 amino acids in length. The 4th, 7th, 8th, and 10th amino acid residues in the repeats were mainly hydrophilic; the 5th, 9th, 13th, and 16th charged; and the 2nd, 9th, and 16th rich in Gln and Lys. The overall percent content of the transglutaminase target residues Gln and Lys in the repeat region of the molecule was identical to that of the remainder of the sequence. The 12th position was rich in Pro, which would allow for flexibility in the approximate center of the repeats. In plots of relative hydrophobicity, the repeats appear as alternating patterns of hydrophobic and hydrophilic stretches (Fig. 3). The hydrophobic stretches are predicted to form β structures of 5–7 residues, terminating at the Pro residue.

Transglutaminase catalyzes isopeptide bond formation between the residues Gln and Lys, and ε-(γ-glutamyl)lysine-cross-linked peptides have been identified from several cornified envelope proteins. The sciellin amino acid sequence was examined for the presence of previously reported transglutaminase target sequences. The repeat GQDPVK found as a transglutaminase substrate in elaphin (8) and the repeats AQEPVK and GQDKVK (8) found to link elaphin to loricrin were not observed in sciellin. Additionally, the preferred transglutaminase sites in involucrin at Gln165 (EEQV) (38) and at Gln165 and Lys168 in the context of QEEK (39) were not present in sciellin. The sequences surrounding the sites of transglutaminase cross-linking, the amine donor at Lys3 and the amine acceptor at Gln102, that were utilized in the oligomerization of the calcium-binding CE protein S100A11 were also not observed in sciellin (40, 41). However, short sequences including Lys or Gln were found several times in the repeat region of sciellin, as shown in Fig. 2. The peptide sequence LIKV was repeated seven times, NQG was found three times, and GQS was found twice.

The carboxyl domain of sciellin contained a LIM motif consensus sequence (Fig. 4). The LIM domain has a characteristic arrangement of Cys and His residues, which coordinate zinc ions into structures that function as protein-binding interfaces (42). All nine Cys residues in sciellin were found in this carboxyl domain, starting at residue 601. These were not expected to form interchain cross-links because sciellin has the same mobility in SDS-PAGE with and without treatment with a disulfide-bond reducing agent.

Comparison of the encoded protein with the PROSITE data base (43) revealed that sciellin has 16 potential N-glycosylation sites, 5 potential myristoylation sites, and 35 potential phosphorylation sites (including 3 cAMP- and cGMP-dependent protein kinase phosphorylation sites, 22 protein kinase C phosphorylation sites, 9 casein kinase II phosphorylation sites, and 9 protein kinase C phosphorylation sites).
1 tyrosine kinase phosphorylation site). Western blots of large
two-dimensional gels of cultured keratinocyte extracts using
34D11 showed bands of 85.7 kDa (pI 8.32), 86.1 kDa (pI 8.72),
86.1 kDa (pI 8.86), and 86.2 kDa (pI 9.08). Post-translational
phosphorylation of sciellin may explain the difference between
the observed and calculated pI and
$\text{Mr}$ values of the protein.

Analysis of the sciellin protein sequence with the subcellular
localization site prediction program PSORTII (32) suggested
that sciellin may be a nuclear protein, based upon the presence
of four potential nuclear localization signals. However, we have
never found sciellin located in the nucleus by immunological
staining of keratinocyte cell cultures, which presumably con-

The sequence reported in Fig. 2 is believed to represent the
short message. When PCR primers made to the sequences
surrounding the initiator methionine and stop codons were
used to amplify the complete coding region of the sciellin cDNA,
only a single band was ever found. Both monoclonal and poly-
clonal antibodies recognized a single band on a Western blot,
suggesting the difference in message sizes is not in the coding
region. Multiple attempts at 5'$\text{RACE}$, using nested PCR prim-
ers that were derived from several different regions within the
coding sequence always gave a single band corresponding in
size to the reported 5'$\text{RACE}$, using nested PCR prim-
ers that were derived from several different regions, resulted in two

---

3 The Human Keratinocytes-NEPHGE Database (Danish Center for
Human Genome Research) is available via the World Wide Web
(http://biobase.dk/cgi-bin/celis).
bands corresponding to the reported 3’ end and one that was 1 kb longer. The larger band, cDNA 86A (Fig. 1), was difficult to subclone and, when sequenced directly from the PCR product, gave a sequence identical to the short form clone 86C, including a polyadenylation signal and a poly(A) tail in the same position, and unreadable sequence 3′ of this tail. It is not known if clone 86A has an additional poly(A) tail at its 3′-most end, as this region remains unsequenced, but the utilization of alternative transcription termination signals is a well documented phenomenon (46). When a portion of cDNA 86A (2.0–3.35 kb, Fig. 1) was used as a Northern probe, it hybridized to both the 3.4- and 4.4-kb bands (Fig. 5C, lane 2), although the 4.4-kb band signal intensity was much stronger than the 3.4-kb band intensity. Assuming that the size difference in mRNA bands is due to an additional 1 kb at the 3′ end, one would expect that only 0.35 kb of the 86A probe could hybridize with the 3.4-kb band and all 1.35 kb of the probe could hybridize with the 4.4-kb band, resulting in significantly different hybridization intensities. This is consistent with the observed 86A Northern band intensities. Northern probes derived from anywhere else in the sciellin cDNA hybridized to the two bands with roughly equal intensity (Fig. 5C, lane 1). Therefore, it is probable that the two message sizes observed for sciellin can be explained by the presence of two transcription termination signals that are 1 kb apart in the 3′-untranslated region.

In Situ Hybridization—In situ hybridization of human foreskin sections with a sciellin antisense probe made from clone SC5 showed that sciellin mRNA is expressed in the upper stratum spinosum and the stratum granulosum, whereas the sense control probe did not (B). For comparison, the antisense probe made from keratin K5 (C) hybridized to the basal keratinocytes of the stratum germinativum, whereas the sense probe did not (D).

Immunohistochemical Localization—Two rabbits were immunized with a sciellin fusion protein expressed from the equivalent of cDNA clone SC4 using the pET system (Novagen). The polyclonal sera of both rabbits identified a band of 84 kDa by Western blotting of lysis buffer extracts of cultured human keratinocytes, whereas the preimmune sera did not react. This band was identical in mobility to the sciellin band identified by monoclonal antibody 34D11 (Fig. 7).

Immunofluorescent staining of normal human body and plantar skin using the SC4 antibody showed peripheral staining of keratinocytes in the granular and upper spinous layers of epidermis, identical to the pattern observed with 34D11 (Fig. 8, A and B). Bovine and newborn mouse skin gave a similar pattern of staining. The stratified epithelium of adult mouse tongue, oral cavity, esophagus (Fig. 8, C and D), and vagina reacted strongly with the antibody, whereas the epithelium of bladder and ureter reacted very weakly, as compared with preimmune serum staining. The simple epithelium of mouse and bovine trachea, lung, kidney, and liver, and of mouse small and large intestine did not stain with the antibody. Human, mouse, and bovine non-epithelial tissues also did not react. These findings are identical to the previously reported tissue distribution of sciellin using 34D11 (24). An exception to the lack of staining of simple epithelium was the reaction of both SC4 and 34D11 antibodies with bovine and human amnion (Fig. 8, E–H), in which sciellin presumably may play a role in barrier function.

Chromosomal Localization and Radiation Hybrid Mapping—A 1.45-kb cDNA probe, pPO325, was used to localize the sciellin gene to chromosome 13, band q22 (SecDNA Biotech, Inc., Ontario, Canada). Under the conditions used, probe pPO325 hybridized with 76% efficiency (among 100 mitotic

---

**Fig. 5.** A, human sciellin is encoded by messages of 3.4 and 4.4 kb. A sciellin cDNA probe was hybridized to a Northern blot containing 25 μg of RNA per lane. Lane 1, amnion; lane 2, smooth chorion; lane 3, cultured keratinocytes; lane 4, villous chorion; lane 5, decidua basalis. B, sciellin shows limited expression in human tissues. A sciellin cDNA probe was hybridized to a multiple tissue Northern blot (Invitrogen) containing total RNA from various adult human tissues. Lane 1, esophagus; lane 2, stomach; lane 3, intestine; lane 4, colon; lane 5, uterus; lane 6, placenta; lane 7, bladder; lane 8, adipose tissue. C, a human keratinocyte Northern blot was reprobed to demonstrate that the difference in size between the 3.4- and 4.4-kb messages is due to the addition of 1 kb to the 3′ untranslated region. Lane 1, a probe from plasmid pPO325 which recognized both long and short forms equally; lane 2, a probe representing the 3′-most 1.35 kb of cDNA 86A, which recognized predominantly the long form. The arrowheads indicate the positions of the 28 and 18 S rRNA bands.

**Fig. 6.** In situ hybridization of sciellin to newborn human foreskin. The antisense probe made from sciellin clone SC5 (A) hybridized to the upper stratum spinosum and the stratum granulosum, whereas the sense control probe did not (B). For comparison, the antisense probe made from keratin K5 (C) hybridized to the basal keratinocytes of the stratum germinativum, whereas the sense probe did not (D).
Sciellin

Fig. 7. Western blot of an SDS lysis buffer extract of human cultured keratinocytes demonstrating that the pAb and mAb react with the identical band. Lane 1 was reacted with pAb SC4, rabbit A, lane 2 with pAb SC4, rabbit B, lane 3 with mAb 34D11, lane 4 with preimmune serum, rabbit A, and lane 5 with preimmune serum, rabbit B. The less intense low molecular mass bands are degradation products of sciellin. The arrowheads indicate the positions of the 112-, 84-, and 53-kDa molecular mass standards.

We report here the cDNA cloning and characterization of sciellin, a novel cornified envelope precursor isolated from human keratinocytes. The cDNA was initially identified by screening a human foreskin keratinocyte expression library with a sciellin monoclonal antibody. Polyclonal antisera raised against a fusion protein expressed from the sciellin cDNA stained tissue sections by indirect immunofluorescence in a pattern identical to the pattern generated using the sciellin monoclonal antibody. Additionally, the monoclonal antibody and the fusion protein polyclonal antisera recognized bands of the same mobility on a Western blot. The amino acid sequence of tryptic peptides generated from affinity-purified sciellin protein matched the protein sequence translated from the cDNA.

The full-length cDNA encodes a unique protein of 668 amino acids and hybridizes with messages of 3.4 and 4.4 kb on Northern blots of poly(A)+ RNA from stratified squamous epithelial tissues and amnion. The deduced sciellin protein sequence can be divided into three domains based upon potential structure and function: an amino domain, a central domain comprising 16 repeats, and a carboxyl domain containing a single LIM motif.

The sciellin repeats are 20 residues long and are characterized by a striking pattern of alternating stretches of hydrophobicity and hydrophilicity on hydropathy plots. The hydrophobic stretches are 5–7 residues long and are predicted to form β sheets. Interestingly, these hydrophobic stretches are the most highly conserved region in the consensus sequence created from the alignment of all 16 sciellin repeats. A data base search with the sciellin repeat region sequence using the program BLAST revealed only a bacterial protein with similar internal repeats. The acidic repeat protein from Treponema pallidum (accession number AF015824) has 14 nearly identical repeats, which are also 20 residues long and contain an internal hydrophobic β sheet. Although there is only modest sequence similarity between the two repeat regions, their similar hydrophathy profiles suggest a similar structure (data not shown).

This arrangement of hydrophobic β sheets within repeat modules is very similar to the parallel β helix structure that has been described for the P22 tailspike protein (49) and pectate lyase C (50). The crystal structure of the P22 tailspike protein demonstrates that each subunit of the homotrimer contains a large parallel β helix. The β helix of each strand is formed by short parallel β sheets coiled into a large right-handed helix, similar to a rope coiled into a tidy cylinder. Each turn of the β helix comprises between 16 and 22 residues. The hydrophobic side chains stack into the helix interior so the β strands are arranged in stacks like runs on a ladder. The charged and polar residues form a hydrophilic interface between adjacent β helices in the homotrimer. The carboxyl domain of P22 tailspike protein is important for the association of monomers. By analogy, sciellin could form homotrimeric β helices, which become cross-linked by transglutaminase into very rigid protein girders, lending structural support to the cornified envelopes of stratum corneum cells. The presence of a PEST sequence could ensure that monomeric sciellin is rapidly degraded in the absence of cross-linking transglutaminases.

The carboxyl domain of sciellin contains a single LIM motif. LIM motifs, which have the consensus sequence CX_{2}CX_{16–21}HX_{2}CX_{2}CX_{16–21}CX_{2}CX_{2}(C,H,D), have been identified in a number of proteins with diverse functions and subcellular locations, including components of adhesion plaques and the actin-based cytoskeleton, and transcription factors. A single LIM domain of roughly 56 amino acids folds into two independent, zinc finger-like structural domains, with the conserved Cys and His residues coordinating two zinc ions (51). LIM domains are believed to mediate protein-protein interactions (38, 51, 52). The carboxyl LIM domain of sciellin may mediate the binding of sciellin to other cytoskeletal or adherens junction proteins, which are subsequently cross-linked to form the cornified envelope.

Proteins containing LIM domains have been classified ac-
According to the number and position of LIM domains and the presence of other functional motifs, such as homeodomains (53). Sciellin is a group 3 LIM protein as it has a single LIM domain in its carboxyl end. Other group 3 LIM proteins are known to interact with the cytoskeleton and might function in cell adhesiveness, in intracellular signaling, and in defining cell shape. This group includes zyxin, a low abundance phosphoprotein concentrated at adhesion plaques and associated with actin filament arrays (54). Zyxin contains three LIM domains in its carboxyl end, and these have been shown to mediate zyxin binding to CRP (55), which also contains two LIM domains. Zyxin binds \( \alpha \)-actinin through a different functional domain. AbLIM (56) is an actin-binding LIM protein that localizes to adherens junctions in the retinal outer limiting membrane. Enigma is a LIM protein that binds to insulin receptors via a LIM domain recognizing a Tyr-containing tight turn structure on the receptor (57).

Individual LIM domains can display distinct binding specificities. The conserved Cys and His residues coordinate zinc ions, folding the LIM domain into two loops. The sequences of the intervening loops may confer binding specificity. This distinct partner preference of a single LIM domain has been demonstrated in the protein zyxin. One of the three zyxin LIM domains is necessary and sufficient to direct specific binding of zyxin with CRP, but not with other LIM domain-containing proteins (38). It has also been demonstrated that a single LIM domain can interact with two other LIM domains as an adaptor element to promote the assembly and targeting of multiprotein complexes (52). The LIM domain of sciellin may be involved in homotypic or heterotypic associations and may function to localize sciellin to the cornified envelope, to mediate the assembly of multiprotein structural complexes, or to regulate the activity of its protein partners.

Acknowledgments—We thank Dr. Juergen Engel for assistance with the protein structure analysis and Dr. Jerome Gross for critical reading of the manuscript.

REFERENCES
1. Reichert, U., Michel, S., and Schmidt, R. (1993) in Molecular Biology of the Skin (Darmon, M., and Blumberg, M., eds) pp. 107–150, Academic Press Inc, San Diego
2. Elias, P. M., and Friend, D. S. (1975) J. Cell Biol. 65, 180–191
3. Polakowska, R. R., and Goldsmith, L. A. (1991) in Physiology, Biochemistry and Molecular Biology of the Skin (Goldsmith, L. A., ed) 2nd Ed., Vol. 1, pp. 168–201, Oxford University Press, New York
4. Ishida-Yamamoto, A., Kartasova, T., Matsuo, S., Kuroki, T., and Iizuka, H. (1997) J. Invest. Dermatol. 108, 729–731
5. Sun, T. T., Tseng, S. C., Huang, A. J., Cooper, D., Schermer, A., Lynch, M. H., Weiss, R., and Eichner, R. (1985) Ann. N. Y. Acad. Sci. 455, 307–329
6. Baden, H. P., and Bonar, L. C. (1968) J. Invest. Dermatol. 51, 476–483
7. Dale, B. A., Holbrook, K. A., and Steiner, P. M. (1978) Nature 276, 729–731
8. Steinert, P. M., and Marekov, L. N. (1995) J. Biol. Chem. 270, 17702–17711
9. Takahashi, H., Kinouchi, M., Wuepper, K. D., and Iizuka, H. (1997) J. Invest. Dermatol. 108, 845–847
10. Ruhberg, C., Nasser Hajibagheri, M. A., Simon, M., Dooley, T. P., and Watt, F. M. (1996) J. Cell Biol. 134, 715–729
11. Takahashi, M., Tezuka, T., and Katunuma, N. (1996) Arch. Biochem. Biophys.
