Identification and Sequence Analysis of Two New Members of the SKALP/ELAfin and SPAI-2 Gene Family

BIOCHEMICAL PROPERTIES OF THE TRANSGLUTAMINASE SUBSTRATE MOTIF AND SUGGESTIONS FOR A NEW NOMENCLATURE

(Received for publication, March 17, 1997, and in revised form, May 20, 1997)

Patrick L. J. M. Zeeuwen†, Wiljan Hendriks‡, Wilfried W. de Jong§, and Joost Schalkwijk∥

From the Departments of Dermatology, Cell Biology and Histology, and Biochemistry, Institute of Cellular Signaling, University of Nijmegen, P. O. Box 9101, 6500 HB Nijmegen, The Netherlands

The human epithelial proteinase inhibitor SKALP/ELAfin and the porcine sodium-potassium ATPase inhibitor SPAI-2 are two highly homologous proteins that share an NH2-terminal transglutaminase substrate domain and a COOH-terminal whey acidic protein (WAP) domain. Here we describe the bovine and simian orthologs of SKALP/ELAfin as well as two new bovine family members that are designated Trappin-4 and Trappin-5 on the basis of a new nomenclature that we propose (Trappin = Transglutaminase substrate and WAP motif-containing Protein). Sequence analysis of Trappin-4 and Trappin-5 revealed a domain structure that is very similar to SPAI-2 (Trappin-1) and SKALP/ELAfin (Trappin-2). The transglutaminase substrate motifs are conserved although the number of repeats varies among species and among family members. The sequence of Trappin-4 and Trappin-5 diverges from Trappin-1 and Trappin-2 at the putative reactive site in the WAP domain. The bovine ortholog of Trappin-2 is expressed in tongue, ileum, and tongue; and Trappin-5 is expressed at low levels in tongue and snout; and Trappin-5 is expressed in tongue and snout and in mouse melanocytes. The cross-linked SKALP/ELAfin is proteolytically processed via proteases from human foreskin epidermis that SKALP/elafin is expressed in several human stratifying squamous epithelia, except for epidermis where it is expressed only as acceptor. We propose that the Trappin protein family forms a new group of enzyme inhibitors with various specificities of the WAP domain, which share transglutaminase substrate motifs that can act as an anchoring sequence.

Skin-derived antileukoproteinase (SKALP)† (1) and sodium-potassium ATPase inhibitor-2 (SPAI-2) (2) are two molecules that share an NH2-terminal domain that functions as a transglutaminase (TGase) substrate, and a COOH-terminal whey acidic protein (WAP) domain that harbors an inhibitory activity toward at least two distinct enzymes. Porcine SPAI-2 was the first of these molecules to be described and is expressed mainly in the intestine (3). Human SKALP, otherwise known as elafin (4), or elastase-specific inhibitor (5), is a potent inhibitor of the leukocyte proteinases elastase and proteinase-3 (6, 7). We found that SKALP/elafin is expressed in several human stratifying squamous epithelia, except for epidermis where it is only expressed in the context of inflammation, such as psoriasis or wound healing (8–10). We mapped the genomic localization of human SKALP/elafin to chromosome 20q12–18 (11). Interestingly, this region contains various other genes involved in TGase-mediated cross-linking processes such as the tissue TGase gene (12), epidermal TGase (13, 14), and the genes coding for semenogelin I and semenogelin II (15), which are also epithelial TGase substrates. SKALP/elafin is distinct from SPAI-2 on the basis of its inhibitory activity, the amino acid sequence of the putative active site, and the epithelial expression pattern. The cDNAs and genes for porcine SPAI-2 (2, 3), porcine SKALP/elafin (16), human SKALP/elafin (7, 17), and a new porcine family member (16) that is more similar to SPAI-2 than to SKALP/elafin, have been cloned and revealed a high degree of conservation in the gene structure and the intronic sequences but a strong sequence divergence in the second exon.

The SKALP/SPAI-2 gene family members are composed of two evolutionary building blocks that are found in other proteins as well (Fig. 1). The COOH-terminal WAP domain is homologous to the second domain of secretory leukocyte proteinase inhibitor (SLPI), which inhibits elastase and cathepsin G (18). The NH2-terminal domain, containing the TGase substrate motifs, is homologous to the guinea pig seminal vesicle protein-1 (19) and the human semenogelins (15). We showed that the NH2-terminal TGase substrate domain, for which the name "cementoin" was coined by others (20), is actually used in vivo and in vitro for cross-linking to stratum corneum proteins. Recently, it was shown by direct sequencing of cross-linked peptides from human foreskin epidermis that SKALP/elafin is cross-linked in vivo to loricrin and cytokeratin-1 (21), which are structural proteins of the terminally differentiating keratinocyte. The cross-linked SKALP/elafin is proteolytically processed further by unidentified proteinases to yield low molecular weight COOH-terminal fragments containing the antiproteinase activity, starting at amino acid positions 149, 151 and 156 (numbering according to Fig. 2), as shown by NH2-terminal sequencing of purified SKALP/elafin from epidermal scale extracts (7, 10). We have found that the COOH-terminal part of SKALP/elafin is cleared via the plasma and can be recovered from the urine (22, 23). This mechanism provides the epidermis with an anchored proteinase inhibitor,
which could protect the structural proteins of the stratum corneum against unwanted proteolysis; in addition it generates, after cleavage, a gradient of low molecular weight inhibitors from the epidermis to the dermis, thereby possibly interfering with polymorphonuclear leukocyte chemotaxis and polymorphonuclear leukocyte-induced degradation of extracellular matrix proteins such as elastin and basal membrane components.

To study the evolutionary divergence of SKALP/elafin in various species and to identify potential new members of this gene family we performed reverse transcription-polymerase chain reaction (PCR) on mRNA of epithelial tissues from various mammals, using degenerate primers that encompass the TGase substrate domain and the WAP domain. In this way we identified the putative simian and bovine orthologs of SKALP/elafin, and we have identified two new family members from bovine tissue. On the basis of the family members thus far identified we propose a new unifying nomenclature. From all known SKALP/elafin gene family members in various species we derived a consensus hexapeptide sequence for the TGase substrate motif. Using synthetic peptides we show that this hexapeptide motif is an extremely efficient substrate for various TGases.

**EXPERIMENTAL PROCEDURES**

**Tissues**—Bovine and simian (rhesus monkey) tissues were obtained from the central animal laboratory, University of Nijmegen, The Netherlands. Spontaneously shed scales from psoriatic, eczema, and lamellated ichthyosis (LI) patients were collected and stored at -20 °C. The LI patients were homozygous for a splice site mutation in intron 5 of the keratin 16 (KRT16) gene. DNA from four LI patients was used for all experimental purposes. A set of 16 patients, nine with psoriasis vulgaris (PV), one with psoriatic arthritis (PA), four with eczema, and two with lamellated ichthyosis (LI), was subjected to DNA analysis. The 16 DNA samples were a subset of a larger set of 36 DNA samples derived from the central animal laboratory, University of Nijmegen, The Netherlands. Spontaneously shed scales from psoriatic, eczema, and lamellated ichthyosis (LI) patients were collected and stored at -20 °C. The LI patients were homozygous for a splice site mutation in intron 5 of the keratin 16 (KRT16) gene. DNA from four LI patients was used for all experimental purposes. A set of 16 patients, nine with psoriasis vulgaris (PV), one with psoriatic arthritis (PA), four with eczema, and two with lamellated ichthyosis (LI), was subjected to DNA analysis. The 16 DNA samples were a subset of a larger set of 36 DNA samples derived from the central animal laboratory, University of Nijmegen, The Netherlands. Spontaneously shed scales from psoriatic, eczema, and lamellated ichthyosis (LI) patients were collected and stored at -20 °C. The LI patients were homozygous for a splice site mutation in intron 5 of the keratin 16 (KRT16) gene. DNA from four LI patients was used for all experimental purposes. A set of 16 patients, nine with psoriasis vulgaris (PV), one with psoriatic arthritis (PA), four with eczema, and two with lamellated ichthyosis (LI), was subjected to DNA analysis. The 16 DNA samples were a subset of a larger set of 36 DNA samples derived from the central animal laboratory, University of Nijmegen, The Netherlands. Spontaneously shed scales from psoriatic, eczema, and lamellated ichthyosis (LI) patients were collected and stored at -20 °C. The LI patients were homozygous for a splice site mutation in intron 5 of the keratin 16 (KRT16) gene. DNA from four LI patients was used for all experimental purposes. A set of 16 patients, nine with psoriasis vulgaris (PV), one with psoriatic arthritis (PA), four with eczema, and two with lamellated ichthyosis (LI), was subjected to DNA analysis. The 16 DNA samples were a subset of a larger set of 36 DNA samples derived from the central animal laboratory, University of Nijmegen, The Netherlands. Spontaneously shed scales from psoriatic, eczema, and lamellated ichthyosis (LI) patients were collected and stored at -20 °C. The LI patients were homozygous for a splice site mutation in intron 5 of the keratin 16 (KRT16) gene. DNA from four LI patients was used for all experimental purposes.

**RNA Isolation and cDNA Synthesis**—Total RNA from bovine tongue or to 10 ug of total RNA from human psoriatic skin was extracted with RNAzol B as suggested by the supplier (Cinna/Biotex). First strand cDNA was generated from RNA using 2.5 units of avian myeloblastosis virus reverse transcriptase (Expand Reverse Transcriptase, Boehringer Mannheim) in the presence of [α-32P]UTP (Amersham). Antisense RNA probe was transcribed from the Xbal-linearized SKALP/elafin (bovine Trappin-2) template. DNA probe was degraded with RNase-free DNase I and the labeled RNA probe was purified using Chroma spin plus-TE 100 columns (CLONTECH), acidic phenol/chloroform extraction, and ethanol precipitation. Using the RNA labeled antisense probe, the 360-base antisense transcript was hybridized to a denatured radiolabeled SKALP/elafin template. DNA bands were detected by autoradiography. The DNA bands were subjected to Southern analysis on a GeneScreen membrane (Boehringer Mannheim) using a DNA fragment encoding bovine Trappin-2. Autoradiography was performed on Biomax-MR films (Kodak) at -80 °C with intensifying screens. Processing of the autoradiographs was performed using the ImageMaster™ data image system (Pharmacia Biotech Inc.).

**Northern Blot Analysis**—For Northern blot analysis, 10 μg of total RNA from different bovine tissues was electrophoretically separated in a 1% agarose gel (dissolved in 10 mM sodium phosphate buffer, pH 7.0), and blotted onto GeneScreen Plus membranes (27). The blot was hybridized using a radiolabeled 1.5-kb bovine TGase cDNA probe and washed with G418 obtained by reverse transcription-PCR. The blot was washed twice for 15 min at 60 °C using 125 mM sodium chloride, 0.15 mM sodium citrate. After transfer, RNA was fixed to the membrane by ultraviolet irradiation (312 nm, 0.2 J/cm²). The membrane was incubated with ethidium bromide (1 μg/ml) for 15 min prior to photography. Hybridization was performed in 250 mM phosphate buffer at 60 °C according to Church and Gilbert (28), using 32P-labeled random primed probes encoding bovine TGase. The blot was washed in 0.1 × SSC (1.5 mM sodium chloride, 0.15 mM sodium citrate). The blot was washed twice for 15 min at 60 °C using 125 mM sodium phosphate buffer. Autoradiography was performed on Biomax-MR films at -80 °C with intensifying screens. Processing of the autoradiographs was performed using the ImageMaster™ data image system.

**Biotinylated Peptides**—Six hexapeptides with an NH₃-terminal biotin followed by a Cys spacer were synthesized: GQDPVK, GQDPVR, GQDPVPK, GQDPVPR, GQDPVPPK, and -5 obtained by reverse transcription-PCR. The peptide was biotinylated by coupling reagent onto positively charged nylon membranes (Boehringer Mannheim) using 10 × SSC (1.5 mM sodium chloride, 0.15 mM sodium citrate). After transfer, RNA was fixed to the membrane by ultraviolet irradiation (312 nm, 0.2 J/cm²). The membrane was incubated with ethidium bromide (1 μg/ml) for 15 min prior to photography. Hybridization was performed in 250 mM phosphate buffer at 60 °C according to Church and Gilbert (28), using 32P-labeled random primed probes encoding bovine TGase. The blot was washed in 0.1 × SSC (1.5 mM sodium chloride, 0.15 mM sodium citrate). The blot was washed twice for 15 min at 60 °C using 125 mM sodium phosphate buffer. Autoradiography was performed on Biomax-MR films at -80 °C with intensifying screens. Processing of the autoradiographs was performed using the ImageMaster™ data image system.

**Cross-linking of the Biotinylated Peptides to Stratum Corneum Proteins**—Epidermal scales from a psoriasis patient (200 mg) were homogenized in 50 mM Tris-HCl, pH 7.8, 100 mM sodium chloride, and 1 mM phenylmethylsulfonyl fluoride (Sigma), and centrifuged for 30 min at 25,000 × g. The supernatant was used for scales from patients with eczema and LI. The supernatants were stored at -20 °C until further use. For cross-linking experiments 10 μl of scale extract was used with 5 μl (100 μg/ml) of biotinylated peptide. The following buffer conditions were used: 50 mM Tris-HCl, pH 7.8, 100 mM sodium chloride, and calcium chloride at a concentration of 2 mM, in a final reaction volume of 50 μl. In some experiments 2 μl (0.0313 unit/ml) of guinea pig liver TGase (Sigma) was added. After 60 min at 37 °C, the reaction was stopped by the addition of 20 μl of 1 M Tris, pH 8.0 (30).
Fig. 2. Alignment of amino acid sequences of members from the Trappin protein family. The amino acid sequences (single letter code) of the novel bovine Trappin (bTrappin-4 and -5) and the bovine and ovine orthologs of SKALP/elafin (hTrappin-2 and sTrappin-2) are deduced from the partial cDNA sequences obtained by reverse transcription-PCR (primer sequences are excluded). Comparison with human SKALP/elafin (hTrappin-2), porcine SKALP/elafin (pTrappin-2), the porcine sodium-potassium ATPase inhibitor SPAI-2 (pTrappin-1), and a new porcine family member (pTrappin-3) are shown. The boxed and italicized residues correspond to the known active site of SLPI which is a protein that consists of two WAP motifs (18, 57). The scoring matrix used is Needleman and Wunsch (15). The hydrophobic signal sequences of hTrappin-2 (7) and pTrappin-1 (2) are indicated in bold. A phase-1 intron (\(>\)) and a phase-2 intron (\(\ast\)) between the first and second nucleotides of the codon for the amino acid at position 42 separates the exon coding for the signal sequence from the Tgase substrate domain. The hexapeptide repeats in the Tgases are grouped in the second and third blocks of the alignment. As pointed out by Tamechika et al. (16), expansion and contraction of the number of hexapeptides in this gene family can readily, presumably as a result of gene conversions and DNA polymerase slippage. Some sequence similarities suggestive of conversions among the three porcine Trappin genes are indicated (*). The positions of the oligonucleotide forward primers (skal1 and skal2) and reverse primers (skal3 and oligo-dT) used in the PCR are indicated.

<table>
<thead>
<tr>
<th>Protein</th>
<th>Amino Acid Sequence</th>
</tr>
</thead>
</table>
| hTrappin-2| MRASSFLI----             ----     ----       ----       VVVFIA       TGLVLEAAVTV       GV--       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----       ----�
The Trappin Gene Family

RESULTS AND DISCUSSION

Cloning and Sequence Analysis of Bovine and Simian SKALP/Elafin Orthologs and Two New Bovine Members of the SKALP/Elafin Gene Family; Suggestions for a New Nomenclature—To isolate partial cDNAs of bovine and simian SKALP/elafin orthologs, three oligonucleotide primers (Fig. 2) were designed based on conserved sequences within human SKALP/elafin and porcine SPAI-2. Using total RNA derived from bovine tongue and rhesus monkey skin, first strand cDNA was generated in a reverse transcriptase reaction and amplified by PCR with the designed primers. PCR products were cloned and sequenced. Four clones were identified as members of the SKALP/elafin gene family based on the presence of sequences encoding an NH₂-terminal TGase substrate domain and a COOH-terminal WAP-domain. Computer-assisted comparison with published sequences of this gene family was performed to reveal the possible identity of the clones. Fig. 2 shows the alignment of the deduced amino acid sequences of the four clones with the currently known family members. Two sequences can be regarded as the simian and bovine orthologs of SKALP/elafin. Overall, simian SKALP/elafin is 93% identical to human SKALP/elafin, and bovine SKALP/elafin is 71% identical to porcine SKALP/elafin. Within a part of the WAP domain (amino acid residues at positions 163–194 in Table I) these percentages are 97% and 81%, respectively (Table I). The sequences of the putative active sites (residues italicized in Table I) of both simian and bovine SKALP/elafin closely correspond to the known protease binding site of SKALP/elafin (36) and the protease binding site of SLPI (36) which also belongs to the WAP protein superfamily. The homology in the putative active site of both simian and bovine SKALP/elafin suggests that these molecules could be elastase inhibitors, although we have no data from functional studies to substantiate this contention. It was, however, shown recently that the porcine SKALP/elafin ortholog is indeed an inhibitor of at least porcine pancreatic elastase (16).

At present, the nomenclature of the proteins containing an NH₂-terminal TGase substrate domain and a COOH-terminal WAP-domain is very confusing. Until now porcine SPAI-2 (2) and the proteinase inhibitor SKALP/elafin/elastase-specific inhibitor were described in the literature (1, 7, 17, 37). For the NH₂-terminal TGase substrate domain a separate name, elastatin, was coined by Nara et al. (20). Here we propose to give these proteins the acronym Trappin (Trappin = TRansglutaminase substrate and WAP motif-containing ProteIN) as a new nomenclature for this protein family. Since SPAI-2 and SKALP/elafin/elastase-specific inhibitor were the first members to be described, these are designated Trappin-1 and Trappin-2, respectively. Trappin-3 is a new porcine member of this gene family which was called pWAP-3 by Tamechika et al. (18). The two other sequences from bovine tongue we identified (see Fig. 2) are new members of the SKALP/elafin gene family and are designated Trappin-4 and Trappin-5 on the basis of the new nomenclature that we proposed. Sequence analysis of these new members of the Trappin gene family revealed a domain structure that is very similar to porcine SPAI-2 (Trappin-1) and human SKALP/elafin/elastase-specific inhibitor (Trappin-2). The sequence of Trappin-4 and Trappin-5 diverges from SKALP/elafin and SLPI at the putative reactive site; however, the cysteine residues of the four-disulfide core structure obtained from the partial cDNA sequence are conserved. The biological function of these new bovine Trappin gene family members is not known. The putative reactive site region in Trappin-4 and Trappin-5 is distinct from SLPI and SKALP/elafin at the Met181 and Leu182 residues (numbering according to Fig. 2) and could possibly lead to dissimilar functions. Diversity in biological functions among SKALP/elafin, SLPI, and SPAI-2 could be the result of the differences in residues around the protease cleavage site within the WAP domain (see Fig. 2).

Amino acid residues of substrates numbered P₁, P₂, etc., are toward the NH₂-terminal direction, and P₁', P₂', etc., are toward the COOH-terminal direction from the scissile bond, as in the nomenclature of Schechter and Berger (38). The preferable residues at P₁ and P₂ in α-lytic protease, which demonstrates an elastase-like primary specificity, are, respectively, methionine and leucine (39). As the scissile peptide bond in SKALP/elafin and SLPI is Ala180(P₁)-Met181(P₁') and Leu182(P₂'), suggesting a specificity like α-lytic protease (35). Although porcine SKALP/elafin varies from human SKALP/elafin and SLPI at the P₁ and P₂ residue, this protein shows the ability to inhibit porcine pancreatic elastase (16), suggesting that the conserved Met181 residue is necessary for elastase specificity. An intact methionine is required for human SKALP/elafin to act as a high affinity inhibitor of leukocyte elastase (38). The putative reactive site in the new bovine Trappin gene family members Trappin-4 and Trappin-5 lacks this methionine residue, which makes it unlikely that these proteins are elastase inhibitors. Recently, Trappin-3 was described (16) which is 59% identical to Trappin-1 within a part of the WAP domain (amino acid residues 163–194 according to Fig. 2). Comparison of our new family members Trappin-4 and -5 with porcine Trappin-1 and Trappin-3 revealed less than 48% identity. Trappin-4 and -5 share a 93% identity. On the basis of sequence identity in the active site, the Trappin family members can be divided in three

Table I

Sequence comparison between the WAP domains as present in the various Trappin protein family members

<table>
<thead>
<tr>
<th></th>
<th>bTrappin-2</th>
<th>sTrappin-2</th>
<th>pTrappin-2</th>
<th>bTrappin-3</th>
<th>bTrappin-4</th>
<th>bTrappin-5</th>
<th>pTrappin-1</th>
<th>pTrappin-3</th>
</tr>
</thead>
<tbody>
<tr>
<td>bTrappin-2</td>
<td>97</td>
<td>75</td>
<td>75</td>
<td>75</td>
<td>52</td>
<td>52</td>
<td>47</td>
<td>44</td>
</tr>
<tr>
<td>sTrappin-2</td>
<td>78</td>
<td>97</td>
<td>97</td>
<td>75</td>
<td>55</td>
<td>55</td>
<td>47</td>
<td>44</td>
</tr>
<tr>
<td>pTrappin-2</td>
<td>91</td>
<td>91</td>
<td>88</td>
<td>74</td>
<td>68</td>
<td>68</td>
<td>53</td>
<td>47</td>
</tr>
<tr>
<td>bTrappin-3</td>
<td>77</td>
<td>77</td>
<td>87</td>
<td>88</td>
<td>93</td>
<td>93</td>
<td>48</td>
<td>42</td>
</tr>
<tr>
<td>bTrappin-4</td>
<td>74</td>
<td>74</td>
<td>81</td>
<td>81</td>
<td>93</td>
<td>93</td>
<td>48</td>
<td>39</td>
</tr>
<tr>
<td>pTrappin-1</td>
<td>73</td>
<td>73</td>
<td>69</td>
<td>69</td>
<td>68</td>
<td>68</td>
<td>48</td>
<td>59</td>
</tr>
<tr>
<td>pTrappin-3</td>
<td>68</td>
<td>68</td>
<td>72</td>
<td>72</td>
<td>68</td>
<td>68</td>
<td>59</td>
<td>59</td>
</tr>
</tbody>
</table>

Upper right, identities in percentage; lower left, similarities in percentage. For comparison 32 amino acid residues were used (residues at positions 163 to 194 in Fig. 2).
subgroups: Trappin-1 and -3, Trappin-4 and -5, and Trappin-2 (Table I).

Tissue Distribution of Trappin Family Members—The Trappin gene family members exhibit distinct tissue distributions. Porcine Trappin-1 (SPA1-2) was previously found to be abundantly expressed in the intestine, whereas porcine Trappin-3 is found in the intestine at relatively low levels (16). The mRNA for porcine Trappin-2 (SKALP/elafin) was mainly found in trachea and large intestine. This is different from human Trappin-2 which is expressed in several human stratifying squamous epithelia (9), except for epidermis where it is only expressed in the context of inflammation, such as psoriasis and wound healing (8, 10). The observed differences could be the result of different sensitivity of the detection methods or could represent real species differences. To define tissue distribution patterns of bovine Trappin-2, expression of the Trappin-2 mRNA was studied by RNase protection analysis. Protected bands with the expected length (360 base pairs) were detected in the tongue and in the epidermis of snout (Fig. 3A). As a negative control we used RNA from porcine skin, which contains human Trappin-2 mRNA. Because human Trappin-2 has several mismatches compared with the bovine ortholog it will be degraded by RNase and is hence not visible on the gel. The expression of Trappin-4 and -5 mRNA in various bovine tissues was studied by conventional Northern blot analysis. Trappin-4 is abundantly expressed in trachea, and a faint signal was found in ileum and tongue (Fig. 3B), whereas Trappin-5 was expressed at relatively low levels in trachea (Fig. 3C). Trappin-4 an -5 are very similar at the DNA level (77%), and therefore it cannot be totally excluded that the signal of the Trappin-5 probe on the Northern blot is the result of cross-hybridization. Since the Trappin family members are all expressed in tissues that are exposed continuously to microbial stimuli (oral cavity, trachea, intestine) it raises the possibility that some of them are directed against bacterial proteinases rather than exclusively against self-proteinases (as Trappin-2, which is directed against leukocyte elastase and proteinase-3). This contention is supported by the recent finding that the distant family member SLPI, which is homologous to Trappins in its COOH-terminal domain, possesses antiviral and antibacterial activity (43, 44). The high substitution rates in the reactive center region of the Trappins could conceivably provide the host with a defense system against pathogens and parasites and give them the capacity to deal with an increasing number of attacking proteinases.

The TGase substrate motifs in the NH2-terminal part of the Trappins are conserved, although the number of amino acid repeats varies among species and among the Trappin gene family members. Porcine and bovine Trappin-2 display 11 and 12 repeats, respectively, whereas six repeats were found in human and simian Trappin-2. The deduced amino acid sequences of the novel members derived from bovine tongue, Trappin-4 and Trappin-5, both contained five repeats. Combined with 14 repeats found in the published porcine Trappin-1 protein and eight repeats in the new porcine family member Trappin-3, the frequency of amino acids at each position was calculated by comparing the total of 67 repeats in the putative TGase substrate domains of the currently known Trappin protein family members from four different mammalian species (Table II). In this way a consensus hexapeptide sequence of GQDPVK could be deduced, in agreement with earlier studies (2, 7).

Conservation of a TGase Substrate Motif in Members of the Trappin Gene Family—To characterize further the biochemical properties of the above mentioned TGase consensus substrate motif, the biotinylated hexapeptide GQDPVK was synthesized, and enzyme kinetic experiments were performed using a TGase assay (29) to determine optimal reaction conditions. The TGase cross-linking reaction is based on a Ca2+-dependent exchange of primary amines for ammonia at the γ-carboxamide group of glutamine residues. Peptide-bound lysine residues or polyamines serve as the primary amines to form either ε-(γ-glutamyl)lysine or (γ-glutamyl)polyamine bonds between proteins (45), which are highly resistant to chemical and enzymatic degradation (46). As human epidermis is known to contain both TGase activity (47, 48) and various substrate proteins (e.g. involucrin, loricrin, small proline-rich proteins) (21), an extract of porcine strata was used to study the incorporation of the biotinylated hexapeptide. A time course incubation at 37 °C of the biotinylated hexapeptide with scale extract showed that the reaction rate as measured by chemiluminescence detection was linear up to 3 h (data not shown). For further experiments a reaction time of 60 min was used for practical convenience. The effect of the peptide concentration on the rate of cross-linking to epidermal proteins by endogenous TGase is shown in Fig. 4. An apparent Km of 0.46 μM was found.

The protective callus layer resulting from terminal differentiation of the squamous epithelium is thought to be cross-linked by different TGase activities present in mammalian epidermis. These TGases are probably involved in the formation of the cornified cell envelope of terminally differentiating epidermis and of other stratified squamous epithelia. Three TGases are expressed in the epidermis, a ubiquitous tissue type TGase (TGc or TGase 2), a membrane-associated keratinoctye TGase (TGk or TGase 1) present in cultured epidermal keratinocytes and in many epithelial and nonepithelial tissues, and the zymogen epidermal TGase (TGz or TGase 3) known to
The Trappin Gene Family

TABLE II

Consensus amino acid sequence of the TGase substrate motif in various Trappin family members

The established hexapeptide is derived from the frequency of amino acids at each position by comparing 67 repeats in the putative TGase substrate domain within eight members of the Trappin family from four different species. Hexapeptide sequences that contain fewer than three residues identical to GQDPVK are excluded from the frequency calculation.

<table>
<thead>
<tr>
<th>Position</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
</tr>
</thead>
<tbody>
<tr>
<td>Amino acid (frequency, in %)</td>
<td>G (78)</td>
<td>Q (87)</td>
<td>D (67)</td>
<td>P (75)</td>
<td>V (84)</td>
<td>K (81)</td>
</tr>
<tr>
<td></td>
<td>A (7)</td>
<td>G (3)</td>
<td>G (10)</td>
<td>L (6)</td>
<td>D (12)</td>
<td>E (9)</td>
</tr>
<tr>
<td></td>
<td>V (2)</td>
<td>R (3)</td>
<td>V (6)</td>
<td>S (4)</td>
<td>G (3)</td>
<td>N (4)</td>
</tr>
<tr>
<td></td>
<td>E (3)</td>
<td>K (3)</td>
<td>E (6)</td>
<td>Q (4)</td>
<td>F (1)</td>
<td>R (3)</td>
</tr>
<tr>
<td></td>
<td>T (3)</td>
<td>E (1)</td>
<td>L (6)</td>
<td>V (4)</td>
<td>V (3)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>S (3)</td>
<td>P (1)</td>
<td>S (3)</td>
<td>T (3)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>D (1)</td>
<td>H (1)</td>
<td>N (1)</td>
<td>K (1)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>L (1)</td>
<td></td>
<td>R (1)</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Fig. 4. Effect of peptide concentration on the rate of cross-linking to epidermal proteins by endogenous TGase. The biotinylated GQDPVK hexapeptide was incubated with scale extract from a psoriatic patient for 60 min at 37 °C. The reaction mixture was blotted directly onto polyvinylidene difluoride membrane, and the biotinylated proteins were detected by chemiluminescence that was recorded on x-ray film and quantified by image analysis. An apparent \( K_m \) of 0.46 μM was found.

be expressed in differentiated epidermal cells and hair follicles, but not in cultured epidermal keratinocytes (49–53). Incubation of the biotinylated GQDPVK hexapeptide with psoriatic scale extract leads to its cross-linking to stratum corneum proteins by Ca\(^{2+}\)-activated endogenous TGase (Fig. 5A, sample 1). Human epidermis and probably also stratum corneum are known to harbor TGases 1, 2, and 3 (21, 47, 48). From our data, however, it cannot be concluded which type is responsible for cross-linking the biotinylated GQDPVK hexapeptide to substrate proteins. Since TGase 1 and 3 are not available in purified form we only tested guinea pig liver TGase (type 2 TGase), and this was found to cause incorporation of the biotinylated hexapeptide into stratum corneum proteins of a native scale extract (sample 2) or scale extract that had been heat-inactivated to eliminate endogenous TGase activity (sample 3); so both the TGases present in scale extract and type 2 TGase catalyze the incorporation of biotinylated peptide into stratum corneum proteins. Specificity of the reaction was checked by the omission of guinea pig liver TGase (sample 4) or the addition of excess EDTA (not shown).

The biotinylated GQDPVK hexapeptide appeared to be an extremely efficient TGase substrate that acts both as an acyl donor and as an acyl acceptor probe (Fig. 5B). Substitution of the acyl acceptor residue lysine (K) for arginine (R) (sample 2) or substitution of the acyl donor residue glutamine (Q) for asparagine (N) (sample 3) showed no influence on cross-linking to stratum corneum proteins. Substitution of both the lysine and the glutamine residue for, respectively, arginine and asparagine, totally abolished cross-linking of the biotinylated hexapeptide (sample 4). The efficiency of cross-linking is sequence dependent as a hexapeptide in the reverse order (KVPDQG, sample 5) or the exchange of lysine and glutamine (GKDVPQ, sample 6) virtually eliminated cross-linking to stratum corneum proteins. Whether these negative effects on substrate reactivity are the result of changes in structural conformation of the peptide or are a consequence of changing the chemical nature of the side chains surrounding the substrate lysine and glutamine is not clear. Grootjans et al. (29) showed that some residues directly preceding the substrate lysines have a negative effect on TGase activity. These residues, like Asp, Gly, Pro, His, and Trp, appeared to be largely avoided in a total of 30 characterized acyl acceptor (lysine) substrates. We would speculate that degenerate hexapeptide motifs could also be used in vivo. We have shown previously that a synthetic peptide comprising the NH\(_2\)-terminal 14 amino acids of processed SKALP/elafin, which contains the degenerate motif GQDPVK, is also incorporated efficiently by TGases, suggesting that slight variations in the surrounding amino acids are tolerated. In the same study we used full-length SKALP/elafin purified from human keratinocytes and showed that the protein is cross-linked to an acyl acceptor probe by the action of type 2 TGase.

A recent study by Steinert and Marekov (21) showed that the degenerate AQDPVK and GQDKVK sequences were used for cross-linking to loricrin and cytokeratin 1 in vivo, as deter-
mined by amino acid sequencing of purified peptides from human foreskin.

In addition to cross-linking of the GQDPVK hexapeptide to its natural substrate proteins in stratum corneum, we also used purified control proteins to investigate the substrate specificity of GQDPVK for these proteins. The biotinylated GQDPVK hexapeptide was found to be cross-linked efficiently by exogenous type 2 TGase to αβ-crystallin and β- low-crystallin, structural proteins of the vertebrate eye lens (Fig. 5C). These proteins are known acyl acceptor substrates for type 2 TGase, as the COOH-terminal lysine residue of αβ-crystallin was identified as the site of linkage (30), and a lysine residue in the NH2-terminal extension acts as the sole acyl acceptor substrate in βA3-crystallin, which is a component of a β-low-crystallin preparation (31). Two glutamine residues acting as acyl donor were characterized in the NH2-terminal region of βA3-crystallin by Berbers et al. (54). No cross-linking capacity was found using αA-crystallin and γ-crystallin as a substrate, which is in accordance with experiments described previously by Groenen et al. (30, 31). Recombinant SKALP/elafin, which only contains the 57 COOH-terminal amino acids of the full-length molecule (and thus lacking the TGase substrate domain), and bovine serum albumin did not show appreciable cross-linking to GQDPVK. To determine the specificity of the GQDPVK hexapeptide for cross-linking to stratum corneum proteins by endogenous TGase in psoriatic scale extracts, we introduced two other acyl donor probes: a biotinylated TVQQEL hexapeptide that is patterned on the NH2-terminal extension of bovine βA3-crystallin (30), and a biotinylated PGGQQIV heptapeptide, patterned on the amine acceptor sequence in fibronectin (55). Fig. 5D shows that the GQDPVK hexapeptide reacts to a greater extent with stratum corneum proteins than the other acyl donor probes (samples 1–3), whereas they all show the same reactivity toward β-low-crystallin by the action of exogenous type 2 TGase (samples 4–6). This suggests that the GQDPVK motif is a preferred substrate for cross-linking to stratum corneum proteins by epidermal TGases.

TGase activity in stratum corneum extracts could be derived from TGase types 1, 2, and 3. Previous studies, however, have suggested that the bulk of the soluble TGase activity comes from TGase type 3 (49). To investigate the relative contribution of the various TGase types in stratum corneum extracts we used scale extracts from three psoriatic patients (Fig. 5E, samples 1–3), a patient with LI (sample 4), and a patient with eczema (sample 5). Surprisingly, the TGase activity measured in these five scale extracts was very similar. Assuming that TGase 1 would contribute significantly, a decreased TGase activity was expected in scale extract from the patient with LI, since this patient was homozygous for a splice mutation in intron 5 of the TGase type 1 gene as determined by single strand conformation polymorphism analysis and sequence analysis. This mutation causes a deficiency for TGase type 1 as described recently (24) and is probably the origin for disturbed formation of the cornified envelopes which may explain the phenotype of LI. Despite a defect TGase 1 gene in LI, our experiments showed that the GQDPVK hexapeptide is still cross-linked to stratum corneum proteins by endogenous TGases in scale extract from a LI patient. It is most likely that TGase type 2 or 3 is responsible for this phenomenon. In a recent publication TGases 1, 2, and 3 were shown to utilize loricin in vitro as a complete substrate, but the types of cross-linking were different (56). TGase 1 mostly formed oligomeric complexes by interchain cross-links, whereas TGase 3 reactivity is involved in the formation of intrachain cross-links. The participation of TGase 2 in loricin cross-linking was quite weak. It is therefore likely that TGase 3 is the active enzyme in the scale extract of the LI patient. Evidence that the GQDPVK motif can be used for cross-linking by TGase 1 was obtained by using cultured normal epidermal keratinocytes, which do not express TGase 3 (49), as a source of TGase type 1 and substrate (Fig. 6). Subsequent reaction with the biotinylated GQDPVK hexapeptide and analysis by SDS-PAGE show incorporation in proteins predominantly between 30 and 80 kDa (Fig. 6, lane 2). For comparison the incorporation pattern in stratum corneum proteins from psoriatic epidermis is shown (Fig. 6, lane 4). Cross-linking of GQDPVK to αβ-crystallin by TGase 2 is demonstrated in lanes 5 and 6.

To conclude, we have identified novel Trappin family members, and we have characterized some of the biochemical properties of the GQDPVK motif with respect to TGase cross-linking in vitro. In addition to the TGase substrate domain, Trappin family members possess a COOH-terminal WAP motif that harbors putative proteinase inhibitory activity. The constitutive expression of Trappin gene family members in a number of normal epithelia which are subjected to continuous mechanical and microbial stress or inflammatory stimuli (e.g., oral epithelia, esophagus, trachea, ileum) is in line with a role for these molecules as proteinase inhibitors with different substrate specificities for self and (possibly) non-self proteinases. Future research will be directed at investigation of the role of other Trappin family members in epithelial homeostasis and human diseases.

Acknowledgments—We acknowledge Dr. Shigehisa Hirose (Tokyo Institute of Technology, Japan), whose group has previously cloned and described two other family members, for cooperation in coming to a consensus nomenclature and for the suggestion to use the name Trappin gene family.

\[ ^2 \text{H. Kremer, D. Melchers, A. Pol, P. C. M. van de Kerkhof, and P. M. Steijlen, manuscript in preparation.} \]