MOLECULAR AND CELLULAR BIOLOGY, Jan. 1993, p. 613–625
0270-7306/93/010613-13$02.00/0

Vol. 13, No. 1

Profilaggrin Is a Major Epidermal Calcium-Binding Protein

NEDIALKA G. MARKOVA, LYUBEN N. MAREKOV, CONSTANTIN C. CHIPEV, SONG-QING GAN, WILLIAM W. IDLER, AND PETER M. STEINERT*

Skin Biology Branch, National Institute of Arthritis and Musculoskeletal and Skin Diseases, Building 10, Room 9N228, Bethesda, Maryland 20892

Received 30 June 1992/Returned for modification 1 September 1992/Accepted 2 October 1992

Profilaggrin is a major highly phosphorylated protein component of the keratohyalin granules of mammalian epidermis. It contains 10 to 12 tandemly repeated filaggrin units and is processed into the intermediate filament-associated protein filaggrin by specific dephosphorylation and proteolysis during terminal differentiation of the epidermal cells. Later, filaggrin itself is degraded to free amino acids that participate in maintenance of epidermal flexibility. The present paper describes the structural organization of the 5′ region of the human profilaggrin gene as well as the amino terminus of the profilaggrin protein. The primary profilaggrin transcript consists of three exons and two introns. The first exon (exon I) is only 54 bp and is untranslated. The coding sequences are distributed between exon II (159 bp) and exon III, which contains the information for 10 to 12 filaggrin repeats (972 bp each) and the 3′ noncoding sequences. A very large intron separates exons I and II. The combination of a very short exon I with an unusually long intron 1 makes the structure of the profilaggrin gene unique among the epidermally expressed genes investigated so far. Comparison of the expression patterns revealed by primer extension and RNase protection analysis of foreskin epidermal and cultured keratinocyte RNAs suggests that alternately spliced messages, which are different from profilaggrin mRNA, are transcribed from the profilaggrin gene system at earlier stages of epidermal differentiation. The amino terminus of profilaggrin exhibits a significant homology to the small calcium-binding S100-like proteins. It contains two α-helical regions, termed EF-hands, that bind calcium in vitro. This is the first example of functional calcium-binding domains fused to a structural protein. We suggest that in addition to its role in filament aggregation and the maintenance of epidermal flexibility, profilaggrin may play an important role in the differentiation of the epidermis by autoregulating its own processing in a calcium-dependent manner or by participating in the transduction of calcium signal in epidermal cells.

Terminal differentiation of human epidermis involves the expression of sets of novel structural proteins including the keratin intermediate filaments (composed of keratins 1 and 10) and two other proteins, loricrin and profilaggrin, which are codeposited as keratohyalin granules (7, 60). Loricrin is now known to be a principal component of the insoluble cell envelope to which it is covalently cross-linked both by disulfide bonds and by N^{ε}-(γ-glutamyl)lysine cross-links formed by the action of epidermal transglutaminases (24, 25). Profilaggrin has a completely different fate. This protein initially accumulates as an insoluble phosphorylated polyprotein precursor consisting largely of filaggrin repeats (36, 44, 55). It has a half-life of about 6 h, and at very late stages of differentiation, it is dephosphorylated and proteolytically processed into individual filaggrin molecules (22, 45, 51) that presumably interact with and align the keratin filaments in a plane parallel to the surface of the epidermis (8, 57, 58). Filaggrin itself has a half-life of only about 24 h in stratum corneum cells (45) and is degraded mostly into free amino acids that are required for maintenance of epidermal osmolarity and hence flexibility and perhaps even for immunological protection of the organism (9, 43, 56). Thus, on the basis of our current understanding, the product of the profilaggrin gene is ultimately utilized in at least two important functions in normal epidermis. This series of events is clearly abnormal in a number of diseases of cornification in the epidermis, particularly the various classes of ichthyoses (65). For example, in epidermolytic hyperkeratosis and lamellar ichthyosis, there is usually a massive accumulation of profilaggrin and little apparent processing into functional filaggrin; in ichthyosis vulgaris, the production of profilaggrin seems greatly reduced from the normal level (9).

In order to better understand these abnormalities, a more-detailed characterization of profilaggrin gene organization and expression is needed. Recently, we isolated and partially characterized the human gene and confirmed that it encodes a polyprotein consisting of many tandemly arranged filaggrin repeats (18, 41). We showed that it is highly polymorphic, both with respect to the numbers of repeats (10, 11, or 12 repeats within the human population) and to the actual sequences themselves (neighboring repeats show as much as 20% sequence variation). We also identified sequences on the amino and carboxyl termini that are unlike the filaggrin repeats; however, their functions remain obscure. An initiation codon was proposed but could not be confirmed (18), and we described one intron in the putative 5′ untranslated sequence. The purpose of the present study was to extend these initial findings. Primer extension and RNase protection experiments have now enabled us to determine that the transcription initiation site (Cap site) for profilaggrin mRNA lies many kilobases upstream of the site in our earlier model and that there is another very large intron (9,713 bp) in the 5′ untranslated sequences. Moreover, the amino terminus of profilaggrin extends 221 amino acids upstream of the previously proposed initiation codon and contains two functional calcium-binding domains, which possibly afford a third important role for this protein in epidermal differentiation.

* Corresponding author.

TABLE 1. Positions and sequences of the primers

Primer	Positions	Exon	Sequence[a]	Expt[b]
1^-	9924–9897	II	TCAGGATTTGCCGAAATTCCTTT	PE, A-PCR, RNA PCR
2^-	10524–10498	III	AGTGAAGTCAATTTTCTTGTTGTGGTC	PE, A-PCR, RNA PCR
3^-	11156–11132	III	GGTACCTGGCTTGTATTTTCATGTC	PE, A-PCR, RNA PCR
4^-	9900–9883	II	gtgcacgcatgcaagcttCCAGAAGTTCCTTCAGCT	A-PCR
5^-	11128–11101	III	ACCTGTTCACTTGAGATGATGATTGGCC	A-PCR
6^-	10474–10457	III	gtgcacgcatgcaagcttCATCAACCATATCTGGGT	A-PCR
7^-	11098–11081	III	gtgcacgcatgcaagcttCAGCCTTGATCTTT	A-PCR
8^-	10601–10552	III	CCTGATATCGGTAAATTCTCTTTCTGGTAGACTCATAATATGCTTGAGC	A-PCR
1^+	2–32	I	GACCCAGCAGGCTCCTTCAGGCTACATTCT	RNA PCR
2^+	9767–9792	II	GTTCACATTTATTGCCAAAAGATGT	RNA PCR
3^+	10457–10474	III	CCAGATGACCCAGATATGGTTGATG	RNA PCR
4^+	10552–10577	III	GCTCAAGCATATTATGAGTCTACCAG	RNA PCR
AdT			gactcgagtcgacatcga(T_{16})	A-PCR

[a] The lowercase letters represent the sequences of the restriction sites that were added to the profilaggrin-specific sequences as adaptors.
[b] PE, primer extension; A-PCR, RNA-mediated anchored PCR.

MATERIALS AND METHODS

Positions and sequences of the primers. The primers used in primer extension and polymerase chain reaction (PCR) experiments are described in Table 1.

RNA isolation. Total foreskin epidermal RNA was isolated from rapidly frozen foreskin epidermis by guanidine hydrochloride extraction as described elsewhere (41). The RHEK cell RNA was a kind gift from J. H. Han (National Institute of Dental Research).

Primer extension. The primer extensions were done essentially as described elsewhere (52). Briefly, 15 μg of total foreskin epidermal or RHEK cell RNA was annealed to 10^6 dpm of 5′-end-labeled primer (10^8 dpm/μg) overnight at 37, 45, or 50°C in a buffer containing 40 mM PIPES [piperazine-*N*,*N*′-bis(2-ethanesulfonic acid); pH 6.5], 1 mM EDTA, 400 mM NaCl, and 80% formamide. The extensions were done in 1× RT buffer (Bethesda Research Laboratories) with 200 U of SuperScript or 4 U of avian myeloblastosis virus reverse transcriptase (Bethesda Research Laboratories) for 90 min at 45 or 50°C. The extended products were resolved on a 6% denaturing polyacrylamide gel and exposed to an XAR-5 film (Kodak) with a Lightning Plus intensifying screen (Dupont).

RNase protection experiments. The RNase protection experiments were done with a commercial kit, RPA II (Ambion), essentially according to the manufacturer's protocol. The template DNA fragments were subcloned into the pGEM-3z vector (Promega) by standard procedures (52). The antisense riboprobes were synthesized with the T7 or SP6 RNA polymerase to an average specific activity of 2×10^8 dpm/μg. In each case, 15 μg of RNA was annealed to 5×10^5 dpm of riboprobe overnight at 50°C in 400 mM Na^+. The duplexes were digested with a mixture of RNases A and T1 for 60 min at 30°C and resolved on a 6% denaturing polyacrylamide gel. After being dried, the gels were autoradiographed on XAR-5 film (Kodak) with an intensifying screen (Dupont).

RNA-mediated anchored PCR. The procedure of Frohman (17) was followed, with slight modifications. Briefly, 5 μg of DNase I-pretreated RNA was extended in separate reactions with 20 pmol of primer 1^-, 8^-, or 3^-. The excess primers were removed on Chroma spin 100 columns (Clontech), and the tailing reactions were carried out in 20 μl of 1× tailing buffer (Bethesda Research Laboratories) in the presence of 200 μM dATP with 25 U of terminal deoxynucleotidyltransferase (Bethesda Research Laboratories) for 1 h at 37°C. Samples were diluted twice with Tris-EDTA buffer and purified through Chroma spin 100 columns. PCR was done with the commercial DNA amplification reagent kit (Perkin-Elmer Cetus) according to the manufacturer's recommendations. In the first round, 20 μl of deoxyribosyladenine-tailed cDNA was amplified with pairs of 25 pmol of adaptor deoxyribosylthymine primer and 25 pmol of gene-specific primer, e.g., primers 1^-, 2^-, and 5^-, respectively. All PCR runs were done as step cycle reactions with a Perkin-Elmer Cetus DNA Thermal Cycler. The conditions for the first round of PCR included initial denaturation at 95°C for 5 min followed by 10 cycles of denaturation at 95°C (1 min), annealing at 45°C (1.5 min), and extension at 72°C (2.5 min) and another 20 cycles at 95°C (1 min), 55°C (1.5 min), and 72°C (2.5 min); finally, extension was completed at 72°C for 7 min. The amplification mixtures were diluted 1:1,000 with Tris-EDTA buffer, and 1 μl of each dilution was reamplified under the following conditions: initial denaturation at 95°C (5 min); 30 cycles of 95°C (1 min), 55°C (1 min), and 72°C (2.5 min); and final extension at 72°C (7 min), again with adaptor deoxyribosylthymine as a plus-strand primer and another set of nested profilaggrin-specific primers, namely, 4^-, 6^-, and 7^-, for the minus strand. The products of the second amplification were ethanol precipitated, digested with *Hin*dIII and *Xho*I (New England Biolabs), purified through Chroma spin 100 columns, and cloned into the pGEM-3z vector (Promega). A large number of clones carrying profilaggrin sequences were selected and sequenced by dideoxy chain termination reaction with the U.S. Biochemicals Sequenase version 2.0 sequencing kit.

RNA PCR. DNase I-pretreated total foreskin epidermal RNA and RHEK cell RNA (200 ng each) were reverse transcribed and amplified in single reaction tubes by the thermostable rTth DNA polymerase with the Perkin-Elmer Cetus Thermostable rTth reverse-transcriptase RNA PCR kit, essentially according to the manufacturer's protocol. The cDNA synthesis was primed with primer 1^-, 2^-, or 3^- and was carried out for 10 min at 70°C. The upstream primers 1^+, 2^+, 3^+, and 4^+ were added to the corresponding reaction mixtures for the amplification step. The cycling conditions were adapted as follows: initial denaturation at 95°C (2 min) and 30 cycles at 95°C (1 min), 60°C (1 min), and 72°C (1.5 min). The PCR products were analyzed on a 2.5% SeaKem GTG agarose gel.

cDNA library screening and sequencing of the λ clones. A foreskin epidermis λgt11 library (41) was screened with primers 1^-, 2^-, and 3^-. Phage DNA was isolated from single

positive plaques and sequenced by a double-stranded DNA cycle sequencing system (Bethesda Research Laboratories) with the λgt11 forward and reverse primers (Bethesda Research Laboratories), as well as with profilaggrin-specific primers.

Miscellaneous procedures. DNA and RNA gel electrophoresis, end labeling of synthetic oligonucleotides with [γ-^{32}P]ATP, enzyme digestions, ligation reactions, DNA cloning, and colony and plaque hybridizations were done according to standard procedures (2, 52). DNA for sequencing was purified from miniprep cultures by the Magic Miniprep DNA purification system (Promega).

Protein extraction, electrophoresis, and Western blot (immunoblot) analysis. Proteins were extracted from human foreskin epidermis or from 2-day-old mouse epidermis by a procedure described previously (59) with 50 mM Tris–384 mM glycine buffer, supplemented with 2% sodium dodecyl sulfate (SDS)–5% 2-mercaptoethanol, for 5 min in a boiling water bath. The extract was clarified by centrifugation for 15 min at 12,000 × *g*, and profilaggrin was selectively precipitated from the supernatant with 0.25 volume of 4 M NaCl. A pellet was collected at 12,000 × *g* for 10 min and was washed three times with 75% ethanol to remove SDS. Profilaggrin was dephosphorylated by dissolving the pellet in 25 mM NaOH and incubating the solution for 30 min at 22°C (40). The solution was neutralized with 0.25 volume of 500 mM Tris-HCl (pH 6.8) and stored in aliquots at −20°C for further use. Proteins were separated on a 5 to 16% gradient polyacrylamide gel according to the method described by Laemmli (31) and were transferred to a polyvinylidene difluoride (PVDF) membrane on a semidry blotting apparatus in 20% methanol–25 mM Tris–192 mM glycine–0.1% SDS (63) at 50 V/cm for 150 min or at 10 V/cm overnight. For immunological detection, the protocol of Scheidtmann (54) was followed. The blots were blocked in 0.5% Tween 20 in phosphate-buffered saline for 2 h and were probed separately with antibodies against human filaggrin (AKH1; Biomedical Technologies Inc. [6]), mouse filaggrin (24, 68), and an antibody designated DTN (raised in rabbits; Hazleton Research Products Inc., Denver, Pa.) against a synthetic peptide (LERRNNRKGNKGRSKS) corresponding to positions 127 to 142 of human profilaggrin. A second antibody conjugated to horseradish peroxidase was used to detect the bands.

Purification of 50-kDa protein. The supernatant from the precipitation of mouse profilaggrin was loaded on a reverse-phase high-pressure liquid chromatography column (4.6 by 75 mm; Ultrapore C3 [Beckman]) and was fractionated with a 20-min linear gradient from 40 to 60% acetonitrile in 0.1% CF_3COOH. Two neighboring fractions which gave a positive reaction with the DTN antibody were pooled, concentrated under vacuum, and rechromatographed on the same column with a more shallow gradient for 30 min. The fraction containing a major protein band on SDS-polyacrylamide gel electrophoresis with a molecular mass of about 50 kDa was retained.

Calcium-binding assay. Calcium binding to proteins was assayed by the micromethod described by Kawasaki et al. (27), which permits the use of picomolar amounts of protein. The proteins were incubated for 1 h in 10 μM $CaCl_2$–50 mM Tris-HCl (pH 7.5)–100 mM NaCl–10 μCi of $^{45}CaCl_2$ per ml and filtered through Schleicher & Schuell BA85 nitrocellulose membranes. Alternatively, PVDF membranes with proteins blotted on them were washed with two changes of methanol for 5 min, soaked in 10 μM $CaCl_2$–50 mM Tris-HCl (pH 7.5)–100 mM NaCl, and incubated for 1 h with 10 μCi of $^{45}Ca^{2+}$ per ml in the same buffer. The filters were washed twice for 30 s at room temperature in a solution containing 10 μM $CaCl_2$, 50 mM NaCl, 20 mM Tris-HCl (pH 7.5), and 10% methanol, dried, and exposed to an X-ray film for 48 h with an intensifying screen. The autoradiograms were scanned on a computing densitometer with ImageQuant software version 3.0 (Molecular Dynamics).

Computer analysis of sequences. Protein sequence homologies, secondary-structure prediction, and nucleic acid sequences were analyzed with the University of Wisconsin sequence analysis software packages compiled by the University of Wisconsin Genetics Computer Group (10) and with the IBI Pustell sequence analysis software (version 3.5; International Biotechnologies, Inc.).

Nucleotide sequence accession number. The nucleic acid and protein sequences of human profilaggrin are deposited in GenBank under accession number M96943.

RESULTS

Human profilaggrin is synthesized in the late stages of epidermal differentiation from a large (more than 13-kb) mRNA (41). The level of expression is Ca^{2+} dependent, and in cultured human keratinocytes profilaggrin mRNA transcription can be induced with calcium (24, 68). In order to elucidate the mechanisms regulating the expression of this major epidermal protein, we needed to define precisely the structure of the mature mRNA and of the gene encoding it, especially the 5′ region, where most of the regulatory signals are normally located. To this aim, three independent approaches were applied. Figure 1 shows the sequences determined by the results presented below.

Primer extension. Figure 2A shows in schematic form the primer extension experiments that were performed. Three antisense oligonucleotide primers were chosen (Fig. 1B and Table 1). Primer 1^- was located at the 3′ boundary of the first exon defined by comparison of a partial cDNA and a genomic clone (exon II in Fig. 1) (18). Primers 2^- and 3^- resided 51 and 685 bp, respectively, downstream from the 570-bp intron. The primers were end labeled and hybridized in separate reactions to total foreskin epidermal RNA as described in Materials and Methods. After extension with reverse transcriptase, the products were resolved on a denaturing 6% polyacrylamide gel (Fig. 2B). The extension of primer 2^- yielded a single band of about 290 nucleotides (nt) (Fig. 2B, lane 2). The extension pattern with primers 1^- and 3^- was more complex. Only the 210- and 920-nt bands (indicated by arrowheads in Fig. 2B, lanes 1 and 3, respectively) were consistent in length with the 290-nt extension product of primer 2^-. The primer extension experiment was repeated with different RNA preparations, under different conditions (temperatures of annealing and extension) and with different enzymes (avian myeloblastosis virus reverse transcriptase and SuperScript reverse transcriptase). In each case, a similar pattern, although much fainter with the avian myeloblastosis virus than with SuperScript reverse transcriptase, was discerned (data not shown), thus indicating that the extended products were not artifacts of the experimental procedure or due to the degradation of the foreskin epidermal RNA. From our previous data (18), it was clear that the sequence of primer 2^- was contained in the profilaggrin cDNA. The 290-nt band was the only extended product of this primer. We therefore concluded that the 290-nt band in Fig. 2B, lane 2, and the corresponding 210- and 920-nt bands in lanes 1 and 3, respectively, were the major extended products of the human profilaggrin mRNA.

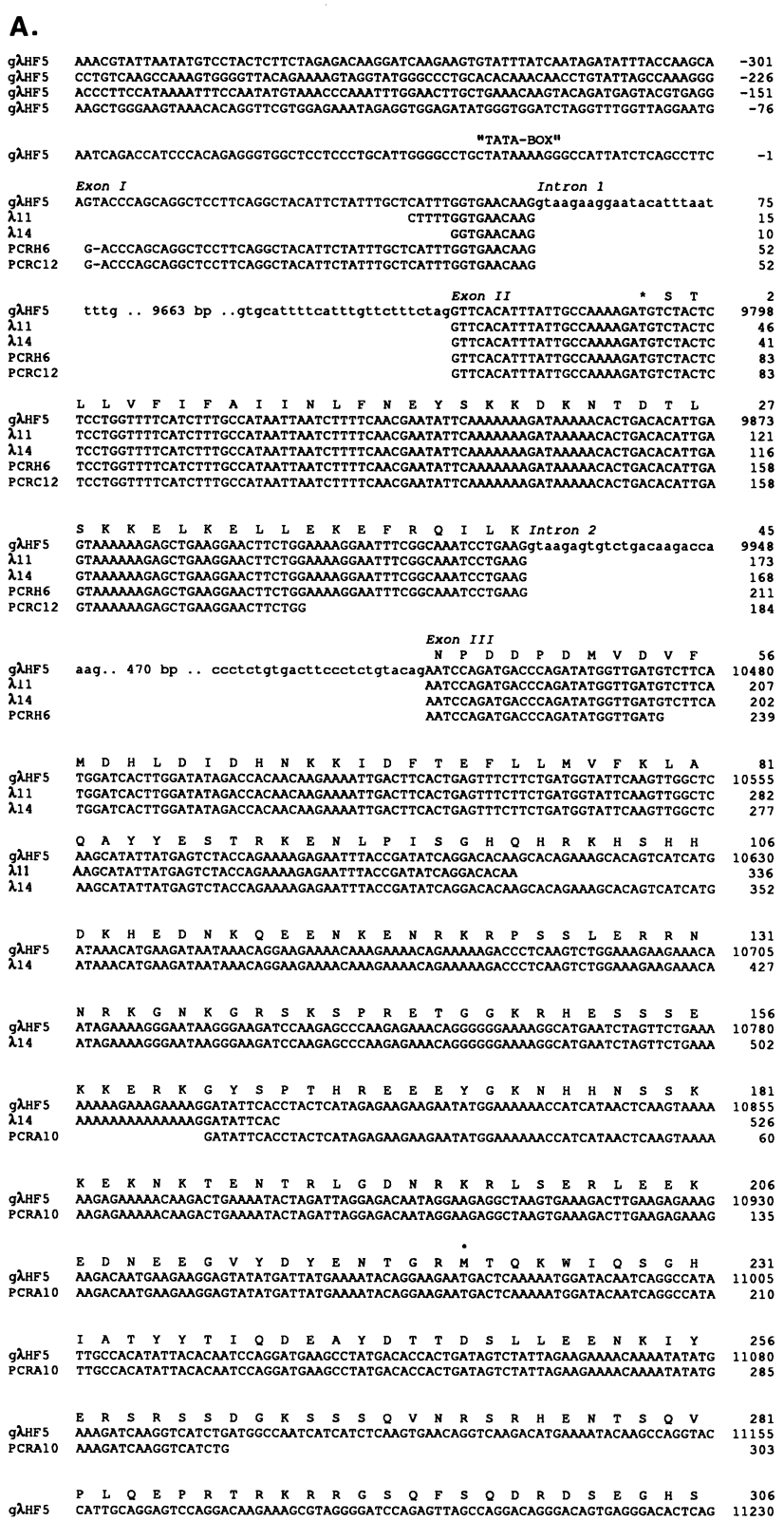

FIG. 1. Sequence of the 5′ end of the human profilaggrin gene and the amino terminus of human profilaggrin. (A) The nucleotide sequence of a 5′ portion of the genomic clone gλHF5 is aligned with the entire nucleotide sequence of two cDNA clones, λ11 and λ14, and the nucleotide sequence of two clones, PCR6 and PCR12, obtained by RNA-mediated anchored PCR of total human foreskin epidermal RNA. Only sequences around the splice junctions of the introns in gλHF5 are shown in lowercase letters; the numbers in the introns correspond to the lengths (in base pairs) of the sequences that are not shown. The deduced amino acid sequence is presented in single-letter code and is numbered from the codon next to the initiation ATG (marked by an asterisk). A black dot on amino acid 222 marks the previously proposed initiation codon. (B) Sequence of the genomic clone gλHF5. Arrows underline the sequences and show the directions of the primers used in primer extension and PCR experiments.

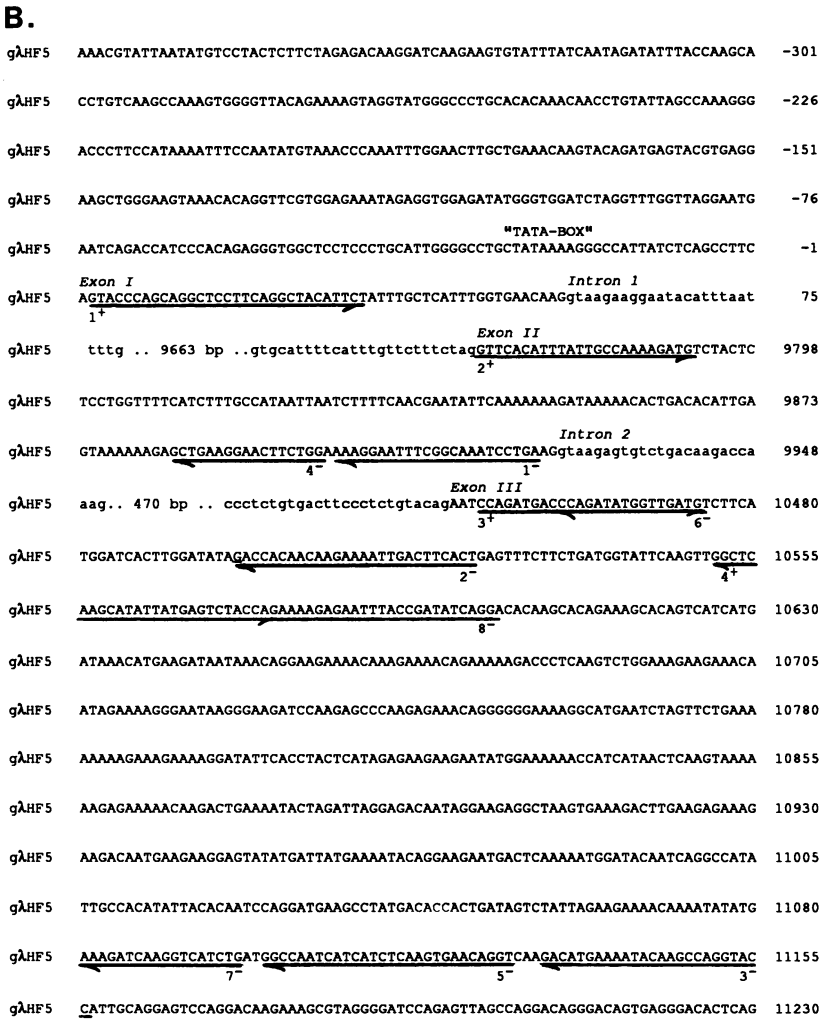

FIG. 1—*Continued.*

In an attempt to elucidate the nature of the additional bands in Fig. 2B, lanes 1 and 3, we repeated the primer extensions with total RNA from RHEK cells (Fig. 2C). RHEK cells (46) are an established simian virus 40-transformed human epidermal keratinocyte cell line that do not synthesize profilaggrin mRNA when grown in monolayers, even in the presence of 1.2 mM Ca^{2+} (39). Comparison of the corresponding lanes in Fig. 2B and C shows that the profilaggrin-characteristic 210-, 290-, and 920-nt extension products were absent in RHEK RNA. However, the profiles in Fig. 2C, lanes 1 and 3, matched all of the nonprofilaggrin bands seen in the corresponding lanes of Fig. 2B, although some of the bands had different intensities. The extension of RHEK RNA with primer 2^- produced a strong band of about 800 nt that was not observed with foreskin RNA.

Collectively, these results suggest that the 210-nt band in Fig. 2B, lane 1, the 290-nt band in lane 2, and the 920-nt band in lane 3 define the same transcription initiation site and originate from the primer extension of the profilaggrin mRNA. The comparatively lower intensity of the 920-nt band may be attributed to the lower efficiency with which reverse transcriptase can extend such long fragments. We interpreted the additional bands in lanes 1 and 3 of Fig. 2B and the 800-nt band in Fig. 2C as most probably representing alternately spliced messages of the profilaggrin primary transcript and/or copies of homologous but different mRNAs expressed during earlier stages of keratinocyte differentiation.

RNase protection. Sequence analysis of the region of the human profilaggrin gene immediately preceding the splice donor site of the 570-bp intron did not reveal the presence of consensus sequence Cap sites that could account for the observed primer-extended fragments. We therefore assumed that the initiation site of profilaggrin mRNA transcription must be located in another exon further upstream. To identify this exon and to elucidate the exon-intron organization in the 5′ end of the profilaggrin gene, a series of RNase protection experiments was carried out. Selected fragments of the human profilaggrin gene spanning the region from 1.2 kb downstream of the *Sac*I site in intron 2 to more than 10 kb upstream of it were subcloned and used as templates for highly labeled antisense riboprobes (Fig. 3A). Each of these was hybridized to total foreskin epidermal RNA. The resulting RNase-resistant duplexes were resolved on a 6% denaturing polyacrylamide gel (Fig. 3B). A single fragment of about 160 nt was protected by the antisense RNA transcribed from the region between the *Sac*I site in intron 2 and the *Hin*dIII site at position 8349 in intron 1 (Fig. 3B, lane 2).

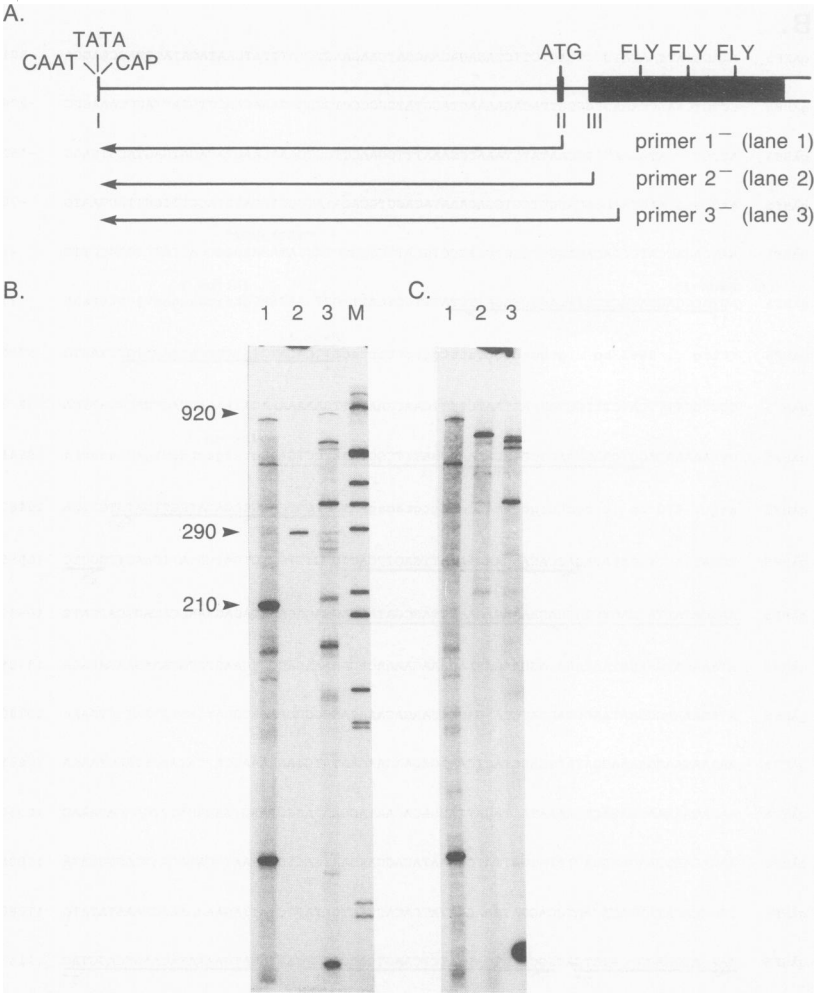

FIG. 2. A primer extension experiment (schematically represented in panel A) was carried out with total human foreskin epidermal RNA (B) and RHEK cell RNA (C). DNase I-pretreated total RNAs isolated from human foreskin epidermis and RHEK cells were annealed to primers 1^- (lanes 1), 2^- (lanes 2), and 3^- (lanes 3) and extended with reverse transcriptase as described in Materials and Methods. The arrowheads point to the three profilaggrin-characteristic extension products. A ^{32}P-end-labeled 1-kb DNA ladder (Bethesda Research Laboratories) was used as a marker.

The same fragment was protected by the antisense transcript extending from the *Sac*I site up to an *Ava*I site at position 7592 (Fig. 3B, lane 3). This fragment represents the protected sequences of exon II, whose splice donor site has been defined by a partial cDNA clone (18). The 5′ boundary of the exon was found to be a consensus sequence splice acceptor site situated 159 bp upstream of the 3′ junction.

The primer extension data in Fig. 2B, lane 1, indicated that the initiation site of the profilaggrin mRNA must lie about 210 bp preceding the 3′ boundary of exon II. Since exon II was found to be 159 bp long, another exon of about 50 bp should exist upstream of exon II. As the results of Fig. 3B, lanes 3 to 8, show, no protected fragment with this length was detected in the region extending 8.3 kb above the *Hin*dIII site. The missing exon I sequences were found as 54 protected nt in the 1.5-kb-long *Xba*I fragment located 10.1 kb upstream of exon II (Fig. 3B [lane 9] and D).

However, exon I and exon II were not the only protected sequences in the 5′ region of the human profilaggrin gene. A number of fragments with various intensities were protected upon hybridization of total foreskin epidermal RNA to antisense transcripts spanning the whole sequence of intron 1 upstream of the *Hin*dIII site (Fig. 3B, lanes 3 to 8) as well as about 3.6 kb of the sequences upstream of the protected exon I (lanes 10 and 11). More information regarding the nature of these sequences was obtained when an RNase protection assay with the same antisense probes was carried out with RHEK RNA. Consistent with the primer extension results, the sequences of exon I and exon II were not protected by RHEK RNA (Fig. 3C [lanes 2, 3, and 9] and D). However, the RNase protection profiles in lanes 4 to 8, 10, and 11 were identical with both RNA species (compare Fig. 3B and C), indicating that sequences in the profilaggrin gene are already used as transcription templates in less-differentiated keratinocytes.

The position of the 5′ boundary of exon III was determined previously by comparison of the sequences of a genomic clone and a cDNA clone (18). When the RNase protection assay was carried out with an antisense transcript of the *Sac*I-*Kpn*I region (Fig. 3A) in accordance with the primer extension data, several fragments were detected (Fig. 3B, lane 1). The longest fragment, about 700 bp, corre-

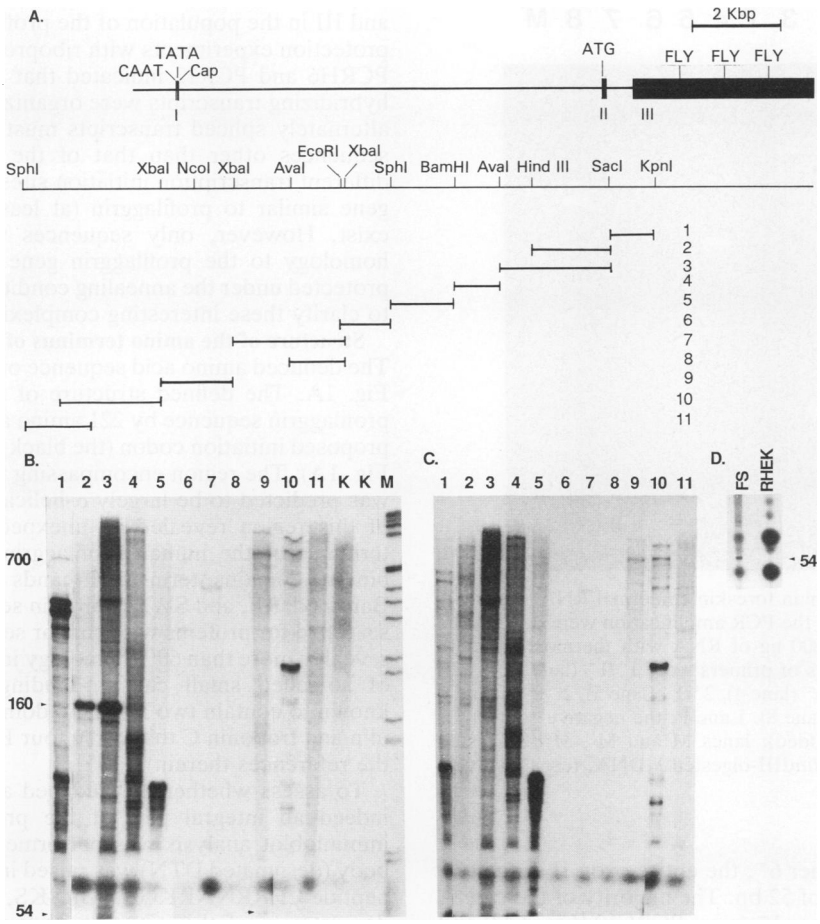

FIG. 3. RNase protection assay. (A) Schematic representation of the 5′ portion of the human profilaggrin gene. Shown are the DNA templates for the highly labeled riboprobes used in the assay shown in the corresponding lanes in panels B and C as follows: *Kpn*I-*Sac*I, lanes 1; *Sac*I-*Hin*dIII, lanes 2; *Sac*I-*Ava*I, lanes 3; *Ava*I-*Bam*HI, lanes 4; *Bam*HI-*Sph*I, lanes 5; *Sph*I-*Eco*RI, lanes 6; *Xba*I-*Xba*I, lanes 7; *Xba*I-*Ava*I, lanes 8; *Xba*I-*Xba*I, lanes 9; *Xba*I-*Sph*I, lanes 10 and 11. (B and C) RNase protection profiles of total human foreskin epidermal RNA (15 μg each) and total RHEK cell RNA (15 μg each) upon annealing to the riboprobes denoted in panel A. (D) Overexposure of lanes 9 in panels B and C, run in parallel, showing that the 54-nt fragment is protected only in the foreskin (FS) RNA.

sponded to the exon III sequence between the splice acceptor site at position 10447 and the *Kpn*I site. The other three strong bands (about 380, 340, and 115 nt, respectively) were highly reproducible with different foreskin epidermal RNA preparations but could not be explained by the combinations of putative Cap sites and splice junctions in this region of the profilaggrin gene predicted by computer sequence analysis with Cap and splice site consensus sequences. Therefore, the RNase protection experiment was repeated with RHEK RNA. With the exception of the 700-nt fragment which was unique to the foreskin RNA, the rest of the signals were again identical with both RNAs (compare lanes 1 of Fig. 3B and C). Presumably they resulted from protection of the RNA transcripts that generated the complex profiles of the primer extension experiments in lanes 1 and 3 of Fig. 2.

cDNA library screening. To confirm this new exon-intron organization, we looked for cDNA clones that might contain exon I sequences. The oligonucleotide primers for the primer extension experiments were used as probes for screening a foreskin cDNA library (41). Two new cDNA clones were isolated (Fig. 1A). Both of them contained the 159 bp of the exon II and 161 and 358 bp of the exon III sequences, respectively. In addition, they contained 10 bp of the new exon I sequence. As we have shown previously (41), in the foreskin cDNA library there are very few cDNA clones to the 5′ end of the profilaggrin gene, probably because of the higher level of RNase activity in the terminally differentiated epidermal cells committed to death and/or the rapid processing of the profilaggrin transcripts. Nevertheless, the sequences of the cDNA clones λ11 and λ14 are in full agreement with the primer extension and the RNase protection results.

RNA-mediated anchored PCR. In order to determine the exact sequence of the first exon and to define precisely the transcription initiation site of the profilaggrin mRNA, RNA-mediated anchored PCR (17) was performed. Total human foreskin epidermal RNA was primed with primers 1^-, 8^-, and 3^- (Fig. 1B; Table 1). The extended cDNAs were deoxyribosyladenine tailed and amplified in two consecutive rounds of PCR (see Materials and Methods for details). The products of the second PCR reaction were introduced into the pGEM-3z vector, and the profilaggrin-specific clones were sequenced. In accordance with the primer extension data, the nested primers 8^-, 2^-, and 6^- eventually generated a single species of cDNA molecules that were 239 bp in length (PCRH6; Fig. 1A) encompassing the sequence of

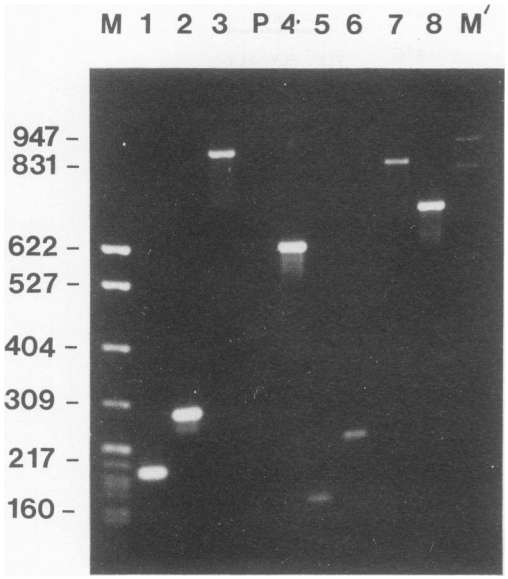

FIG. 4. PCR of total human foreskin epidermal RNA. The first-strand cDNA synthesis and the PCR amplification were done in the same reaction tubes with 200 ng of RNA with thermostable rTth DNA polymerase. The pairs of primers were $1^+/1^-$ (lane 1), $1^+/2^-$ (lane 2), $1^+/3^-$ (lane 3), $4^+/3^-$ (lane 4), $2^+/1^-$ (lane 5), $2^+/2^-$ (lane 6), $2^+/3^-$ (lane 7), and $3^+/3^-$ (lane 8). Lane P, the negative control for primers $1^+/1^-$ (no RNA added); lanes M and M', *Msp*I-digested pBR322 DNA and *Eco*RI-*Hin*dIII-digested λ DNA, respectively.

exon III 5′ from the primer 6^-, the entire exon II sequence, and an exon I sequence of 52 bp. The majority of the cDNA clones generated by primer 1^- (e.g., PCRC12 [Fig. 1A]) had the same structure. In addition, another cDNA clone was found which contained exon II linked to a sequence of 10 bp that had the characteristics of a splice donor site but was different from exon I (data not shown). This is consistent with the existence of a population of alternately spliced transcripts all of which contain exon II.

PCR of total epidermal foreskin RNA. Although our results strongly indicated the juxtaposition of exons I and II together with exon III in the same mature profilaggrin mRNA, we did not possess definitive proof that they resided on the same transcript. To obtain this proof, total human foreskin epidermal RNA was amplified by PCR with several different pairs of primers (Fig. 1B and Materials and Methods). The results are shown in Fig. 4. Since equal amounts of RNA were used in all lanes, the different intensities of the bands could be attributed most probably to the different efficiencies of the PCR amplification with each particular set of primers. The very faint bands on the gel presumably represented PCR artifacts, since none of them was consistent with the observed primer extension and RNase protection profiles. As seen, each pair of primers containing the exon I-located primer 1^+, e.g., $1^+/1^-$, $1^+/2^-$, and $1^+/3^-$, generated virtually a single product of the expected length (Fig. 4, lanes 1, 2, and 3). Similarly, no multiple bands in the region between the 5′ end of exon II and the *Kpn*I site in exon III were discerned (Fig. 1B and 4, lanes 4 to 8). PCR of total RHEK RNA under the same conditions and with the same pairs of primers failed to produce any fragments (data not shown).

Taken together, these results suggest that the observed exon-intron organization is unique to the profilaggrin transcripts. There are no alternate combinations of exons I, II, and III in the population of the profilaggrin mRNAs. RNase protection experiments with riboprobes derived from clones PCRH6 and PCR12 indicated that about 60 to 70% of the hybridizing transcripts were organized in this way (39). Any alternately spliced transcripts must therefore involve exon sequences other than that of the present exon I and/or different transcription initiation sites. Alternatively, another gene similar to profilaggrin (at least in its 5′ region) may exist. However, only sequences with greater than 90% homology to the profilaggrin gene would be extended or protected under the annealing conditions used. Further work to clarify these interesting complexities is in progress.

Structure of the amino terminus of the profilaggrin protein. The deduced amino acid sequence of profilaggrin is shown in Fig. 1A. The defined structure of the mRNA extends the profilaggrin sequence by 221 amino acids past the previously proposed initiation codon (the black dot on amino acid 222 in Fig. 1A). The region encompassing the first 100 amino acids was predicted to be largely α-helical (10). Detailed analysis of this region revealed an unexpected feature: the amino terminus of the human profilaggrin contains two calcium-binding domains, termed EF-hands (29, 61, 64). When GenBank, NBRF, and SWISS protein sequence data bases were searched for proteins with similar sequence, the comparison revealed more than 60% homology in this region, with a class of so-called small calcium-binding proteins (see below) known to contain two EF-hand domains, as well as calmodulin and troponin C that carry four EF-hands (28, 29, 64 and the references therein).

To assess whether the deduced amino-terminal region is indeed an integral part of the profilaggrin molecule, an immunoblot analysis was performed. As a probe, an antibody (designated DTN) was raised in rabbits to the synthetic peptide LERRNNRKGNKGRSKS, located 127 amino acids downstream of the initiation codon and immediately adjacent to the α-helical region. Total human foreskin epidermal protein extract was prepared, resolved on a 5 to 16% gradient polyacrylamide gel (Fig. 5A, lane 1), and blotted onto a PVDF membrane as described in Materials and Methods. The proteins were reacted with anti-human filaggrin antibody (Fig. 5A, lane 2) and with the DTN antibody (lane 3). Whereas the filaggrin antibody recognized both profilaggrin and filaggrin as well as all of the intermediary products of the profilaggrin processing, the DTN antibody detected only profilaggrin and the highest-molecular-weight products of its initial processing (45, 55). The protein transfer for Fig. 5A, lane 3, was done under conditions maximizing the blotting of high-molecular-weight fractions (e.g., 10 V/cm, overnight) but greatly reducing the amounts of relatively low-molecular-weight proteins that are retained on the membrane. When the blotting was performed at 50 V/cm for 150 min, the DTN antibody detected one prominent band with a molecular mass of about 90 kDa and three fainter products with masses of about 70, 50, and 20 kDa (Fig. 5A, lane 4). The nature of these polypeptides is not yet clear. However, the results show that the DTN antibody epitope resides in the profilaggrin molecule and suggest that the amino terminus of the profilaggrin contains the EF-hand domains.

The immunoblot analysis was repeated with newborn-mouse epidermal extracts. It was found that the anti-human DTN antibody cross-reacts with mouse profilaggrin and, similarly to the human extract, with two bands with molecular masses of about 70 and 50 kDa (Fig. 5B, lane 3). Thus, we concluded that mouse profilaggrin also contains the DTN epitope and that its amino-terminal sequence organization is

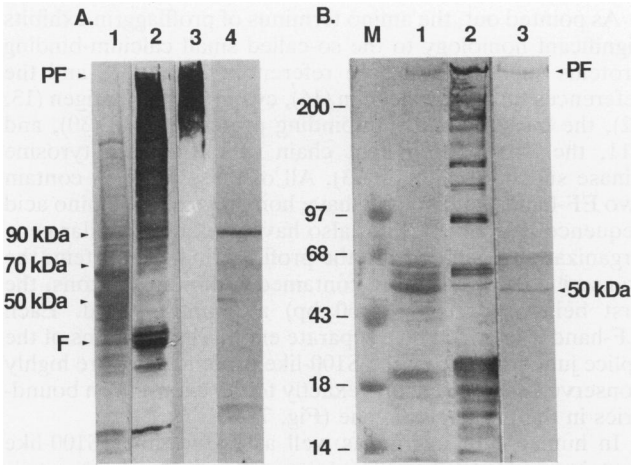

FIG. 5. Western blot analysis of total human foreskin epidermal SDS extracts (A) and newborn-mouse epidermal SDS extracts (B). Extracts of human foreskin epidermis and newborn-mouse epidermis were made in boiling SDS, resolved on 5 to 16% gradient polyacrylamide gels, and blotted onto PVDF membranes as described in Materials and Methods. Lanes 1, Coomassie blue staining of the PVDF-transferred extracts; lanes 2, blot reacted with anti-human filaggrin antibody (A) and with anti-mouse filaggrin antibody (B); lanes 3, human foreskin epidermal extract and newborn-mouse epidermal extract blotted at 10 V/cm overnight and reacted with DTN antibody; lane 4, human foreskin epidermal extract blotted at 50 V/cm for 150 min and reacted with DTN antibody. The arrowheads in panels A and B mark the antibody-positive polypeptides and the DTN antibody-positive profilaggrin (PF) and 50-kDa protein, respectively. Protein molecular mass standards in the range of 14.3 to 200.0 kDa (Bethesda Research Laboratories) were used to calibrate the gels (lane M in panel B). F, filaggrin.

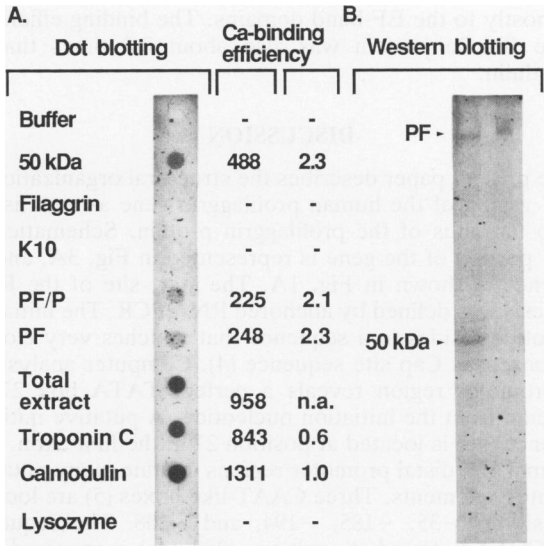

FIG. 6. Calcium-binding assay. (A) Column 1, proteins incubated for 1 h in reaction buffer containing 10 μCi of $^{45}Ca^{2+}$ per ml and filtered through nitrocellulose membrane. Rows: 1, reaction buffer; 2, purified mouse 50-kDa protein (1 μg); 3, purified mouse filaggrin (1 μg); 4, purified mouse K10 (1 μg); 5, dephosphorylated mouse profilaggrin (PF/P; 5 μg); 6, purified nondephosphorylated mouse profilaggrin (PF; 5 μg); 7, total newborn-mouse epidermal extract; 8, troponin C (1 μg); 9, calmodulin (1 μg); 10, lysozyme (1 μg). The autoradiogram in column 1 was scanned with a computing densitometer. The numbers in column 2 are the values above the background for each dot; the numbers in column 3 are the respective efficiencies of calcium binding in comparison with that of calmodulin. The data shown in panel A are from one experiment. The values in column 2 varied between different experiments, but those in column 3 were highly reproducible. n.a., not applicable. (B) Total newborn-mouse epidermal extract (left lane) and purified mouse profilaggrin (right lane) were resolved on a 5 to 16% gradient polyacrylamide gel, transferred to PVDF membrane, and overlaid with $^{45}Ca^{2+}$ as described in Materials and Methods. The arrowheads mark the positions of profilaggrin (PF) and the 50-kDa protein.

therefore most probably very similar to that of human profilaggrin.

The 50-kDa protein was found to be unreactive with anti-mouse filaggrin antibody (Fig. 5B, lane 2) and therefore should not contain sequences characteristic of mature filaggrin repeats. Whether it is a specific product of the profilaggrin processing or an independently translated protein is still unknown, but preliminary immunofluorescence data (38) indicate that a protein different from profilaggrin is recognized by the DTN antibody in the basal layer of normal human and newborn-mouse epidermis as well as in normal human keratinocytes grown in 0.15 mM Ca^{2+}, conditions that are nonpermissive for profilaggrin expression (24, 68).

Profilaggrin is a functional calcium-binding protein. Although the Western blot data suggested that both human and mouse profilaggrins contain the EF-hands, there was no conclusive evidence that this really was the case and, if so, that the EF-hands were functional calcium-binding domains. Since mouse profilaggrin was more stable and easier to purify, it was used in subsequent experiments to test these questions. Total newborn-mouse epidermal SDS extract, purified profilaggrin (5 μg), purified and dephosphorylated profilaggrin (5 μg), purified filaggrin (1 μg) and purified 50-kDa protein (1 μg) were dot blotted on a nitrocellulose filter. Calmodulin and troponin C (1 μg each) were used as positive controls. Reaction buffer, purified mouse keratin 10, and lysozyme were the negative controls. The membrane was overlaid with 10 μM $^{45}Ca^{2+}$, treated as described in Materials and Methods, and exposed to an X-ray film for 48 h (27). The autoradiogram is shown in Fig. 6A. Whereas buffer, filaggrin, keratin 10, and lysozyme did not bind calcium at all, the total protein extract as well as the purified profilaggrin and the purified 50-kDa protein clearly bound Ca^{2+}. That profilaggrin is indeed a calcium-binding protein was further confirmed by the same technique on a PVDF-blotted SDS-polyacrylamide gel of purified profilaggrin (Fig. 6B). The data were quantified by scanning the autoradiogram shown in Fig. 6A in a computing densitometer. The numbers in the second column in Fig. 6A are the values above background for each dot; the numbers in the third column show the respective efficiencies of calcium binding compared with that of calmodulin. Considering that the profilaggrin has molecular weight at least 30 times higher than that of calmodulin, that it contains two instead of four EF-hands, and that 5 μg of profilaggrin versus 1 μg of calmodulin was loaded, it was estimated that profilaggrin binds calcium ions 2.3 times more efficiently than calmodulin. It is known that in the epidermis, the profilaggrin is heavily phosphorylated on serine residues (see reference 43). In order to exclude a possible contribution of the phosphate groups to calcium binding, the purified profilaggrin was dephosphorylated with NaOH (40). This dephosphorylation did not cause any significant decrease in the signal (2.3 versus 2.1), thus indicating that the calcium-binding activity of profilaggrin is

due mostly to the EF-hand domains. The binding efficiency of the 50-kDa protein was also about 2.3 times that of calmodulin.

DISCUSSION

The present paper describes the structural organization of the 5′ region of the human profilaggrin gene as well as the amino terminus of the profilaggrin protein. Schematically, the 5′ portion of the gene is represented in Fig. 3A, and its sequence is shown in Fig. 1A. The start site of the RNA synthesis was defined by anchored RNA PCR. The initiation nucleotide resides in a sequence that matches very closely the consensus Cap site sequence (4). Computer analysis of the promoter region reveals a perfect TATA box 27 bp upstream from the initiation nucleotide. A putative initiator sequence (48) is located at position 27 in the first exon. The proximal and distal promoter regions contain many putative regulatory elements. Three CAAT-like boxes (5) are located at positions −35, −185, −194, and −208. The sequence AAGCCAAA, found at position −288, was recognized previously by Blessing et al. (3) to be present in the 5′-upstream region of a number of genes expressed primarily in keratinocytes. An imperfect direct repeat, CCAGCAGGCTCCT TCAGGCT, of the palindromic KER1 recognition site, found in keratin 14, keratin 5, and keratin 1 genes and implicated in the epithelial specificity of their expression (33, 34), resides 6 bp downstream of the transcription initiation site in the human profilaggrin gene. A perfect AP1/TRE 0/A recognition sequence, ATGAATCAG (47), is located 78 bp upstream of the transcription initiation site. Interestingly, Rothnagel at al. (50) recently identified an AP1 site as the calcium-sensitive regulatory element in the human keratin 1 gene. The functional significance of these regulatory motifs is now under detailed study.

Profilaggrin mRNA is transcribed from three exons. A large intron (9,713 bp) separates exons I and II; intron 2 is only 570 bp long. Exon I is 54 bp long and is untranslated. The protein-coding sequences are distributed between exons II and III. The combination of a very short exon I and an unusually large intron 1 makes the structure of the human profilaggrin gene unique among the epidermally expressed genes so far investigated.

The primer extension and the RNase protection data point out that profilaggrin mRNA is not the only transcriptional product of the profilaggrin gene. Several fragments residing in the large intron 1 and in the 5′-flanking region were protected by both foreskin epidermal and RHEK cell RNAs (Fig. 3). Similarly, the 5′ portion of exon III contains sequences protected by both RNA species. It is conceivable that sequences in the profilaggrin gene are transcribed and undergo alternate splicing to produce mature mRNAs different from that of profilaggrin. The data with the cultured RHEK cell RNA indicate that these transcripts are expressed in relatively undifferentiated basal epidermal cells. Alternatively, these messages could be transcribed from a different but highly homologous gene.

As already mentioned, the first ATG codon in frame with the filaggrin-like amino acid sequence is situated 75 nt downstream from the transcription start site in exon II. No in-frame stop codons are found in the 5′-untranslated region. Interestingly, in view of our preliminary data suggesting that a portion of the primary transcripts are alternately spliced, this permits the open reading frame to be extended further upstream should another Cap site and/or different exon I sequences be utilized.

As pointed out, the amino terminus of profilaggrin exhibits significant homology to the so-called small calcium-binding proteins such as S100 (see references 26 and 28 and the references therein), calcyclin (16), cystic fibrosis antigen (13, 32), the intestinal calcium-binding protein ICaBP (30), and p11, the S100-related light chain of the protein tyrosine kinase substrate p36 (19, 53). All of these proteins contain two EF-hands and not only have homology at the amino acid sequence level (Fig. 7) but also have strikingly similar gene organizations. Similarly to the profilaggrin gene system, the transcribed sequences are contained within three exons, the first being short (50 to 60 bp) and untranslated. Each EF-hand is encoded by a separate exon. The locations of the splice junctions within the S100-like protein genes are highly conserved and correspond exactly to the exon-intron boundaries in the profilaggrin gene (Fig. 7).

In human profilaggrin, as well as in the other S100-like proteins, each EF-hand is composed of two α-helices surrounding a loop of oxygen-containing amino acids that form the actual calcium coordination site (Fig. 7) (28, 32, 61). The helix-loop-helix motifs are flanked by stretches of hydrophobic amino acids. In contrast to the EF-hand sequence, the C terminus and the hinge region between the two EF-hands are much less conserved among the different proteins. These regions are believed to participate in the specific interactions between the small calcium-binding proteins and the target effector proteins in the presence of Ca^{2+} (28).

The Western blots in Fig. 5 show that the amino acid sequence immediately adjacent to the EF-hand motifs is an integral part of the human as well as of the mouse profilaggrin molecule. By a calcium overlay technique, we were able to show that purified mouse profilaggrin binds calcium (Fig. 6A and B). Together, these results strongly indicate that profilaggrin is indeed a calcium-binding protein. However, unlike the small calcium-binding proteins, in human profilaggrin the EF-hand domains are adjacent to 10 to 12 repeats of the mature filaggrin sequence (18). The filaggrin itself is an intermediate filament-associated protein that plays a role in keratin filament aggregation (37, 43, 58). Thus, profilaggrin seems to be a "fused" protein that harbors two entirely different functional domains. There are three examples in the literature so far in which calcium-binding EF-hands have been found linked to sequences of Ca^{2+}-regulated enzymes. One is a serine-threonine protein kinase from soybean that is found in association with the actin microfilament system (23) and the two others are a calcium-dependent protease, calpain, which is present in most mammalian tissues (62), and the highly homologous neutral proteinase from the parasitic trematode *Schistosoma mansoni* (1). Calpain plays a role in the activation of protein kinase C, in the degradation of cytoskeletal proteins, and in the modification of neurofilaments (42, 62). To our knowledge, no enzymatic activity with profilaggrin has ever been detected. Therefore, it is the first example of a protein that exhibits both calcium-binding and nonenzymatic functions. Recently, the partial amino acid sequence of another epidermally expressed gene, trichohyalin, was published (49). Although the most 5′ end of the gene has not yet been characterized, it is clear that the deduced amino terminus of the protein contains two EF-hands and that the core of the protein is composed of numerous tandem repeats, an organization remarkably similar to that of profilaggrin. Not only are profilaggrin and trichohyalin organized in the same way and expressed in the same stage of differentiation of the epidermis (20, 41), but they are both mapped to the chromosome region 1q21.1 (35, 41). The functional implications of this finding remain unelu-

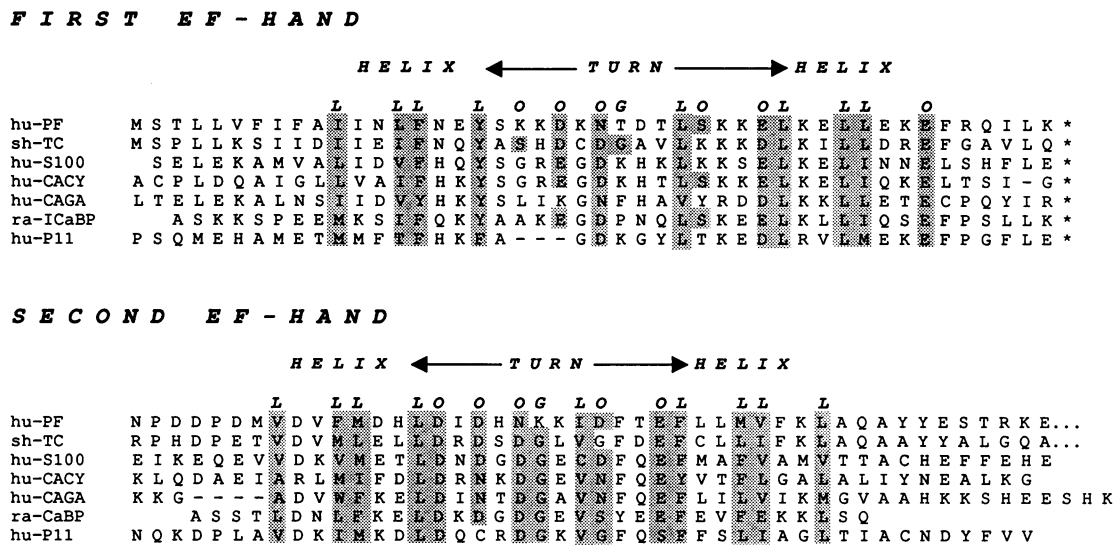

FIG. 7. BESTFIT alignment of S100-like calcium-binding proteins with two homologous EF-hand motifs. Asterisks indicate the positions of the intron sequences separating the two EF-hand-coding exons. The α-helices and turns of the EF-hands are denoted according to Kretsinger (29). L, hydrophobic amino acids (L, V, I, F, and M); O, oxygen-containing residues (E, D, Q, N, S, and T); G, glycine. The conserved residues of the EF-hands are shadowed. Abbreviations: hu, human; ra, rat; sh, sheep; PF, profilaggrin; TC, trichohyalin; CACY, calcyclin; CAGA, cystic fibrosis antigen subunit A; ICaBP, intestinal calcium-binding protein; p11, the S100-related light chain of the protein-tyrosine kinase substrate p36.

cidated. Interestingly, however, several small calcium-binding proteins have also been mapped at 1q21 (12, 21) and some of them, like calcyclin, have been shown to be expressed in the epidermis (66). At the same time, the genes for another two epidermal-specific intermediate filament-associated proteins, involucrin (14) and loricrin (67), have been mapped to 1q21. Similarly to profilaggrin and trichohyalin, both of these genes are organized as multiple tandem repeats. Thus, the 1q21 locus appears to harbor the information for two entirely distinct cellular functions: calcium binding on the one hand and interaction with cytoskeleton on the other. In the cases of profilaggrin and trichohyalin, these two functions reside on one and the same gene product. It is possible that both the profilaggrin and trichohyalin genes arose during evolution by fusion of an ancestral epidermal structural gene composed of multiple tandem repeats with a gene for a calcium-binding protein that resided in the same chromosome locus.

What is the biological function of the calcium-binding domain in profilaggrin? Although limited experimental data are available, there are several Ca-dependent processes in the epidermis in which profilaggrin may play a crucial role. It has been known for some time that the terminal differentiation of keratinocytes is tightly regulated by extra- and intracellular calcium concentrations (15, 24, 68). Recently, it was established that induction of protein kinase C activity in mouse keratinocytes enhances the expression of profilaggrin mRNA only at permissive Ca^{2+} concentrations, i.e., above 0.2 mM (11). It is possible that profilaggrin has a role in the transduction of the Ca signal during the coordinate proliferation and differentiation of keratinocytes. Alternatively, the interaction of filaggrin with the keratin intermediate filaments must be tightly regulated in the keratinocytes to prevent premature aggregation. The different steps of this regulation include phosphorylation and dephosphorylation of profilaggrin and its proteolysis to mature filaggrin repeats. Preliminary evidence suggests that the activity of at least one of the phosphatases involved and the proteolytic processing from intermediates to filaggrin is influenced by calcium (43). The profilaggrin calcium-binding domains could be the modulators of the activity of the responsible enzymes. In this way, profilaggrin could effectively autoregulate its own processing. Finally, profilaggrin may exert its role as a calcium-binding protein by interacting in a Ca-dependent manner with different components of the cytoskeleton and/or of the plasma membrane. In support of these speculations, there are several instances in which small calcium-binding proteins have been observed to modulate the activities of key enzymes or to change cell shape and morphology during differentiation (for a review, see reference 28).

In conclusion, by means of various molecular biology techniques we have elucidated the exon-intron organization of the human profilaggrin gene. We have shown that profilaggrin mRNA consists of three exons, the first of which is predicted to be untranslated. The mRNA encodes a large protein that contains two calcium-binding domains. Our new data suggest that the profilaggrin gene system subserves a third important function in epidermal differentiation. In addition to production of filaggrin that binds and organizes the keratin intermediate filaments and then production of free amino acids ensuring the flexibility of cornified cells, it is very likely that the calcium-binding domains of profilaggrin play an important regulatory role in the epidermis. An understanding of how these functions are exerted in concert and finely tuned during different stages of epidermal differentiation could provide a clue to the defects underlying several skin disorders in which the profilaggrin-filaggrin system has been implicated (9, 43).

ACKNOWLEDGMENTS

We thank John Compton for numerous helpful discussions during the course of this work. We are grateful to Stuart Yuspa and colleagues, who provided unlimited access to mouse epidermal

tissue and shared their unpublished results. We also thank Jung-Ho Han for RHEK cell RNA.

REFERENCES

1. **Andresen, K., T. D. Tom, and M. Strand.** 1991. Characterization of cDNA clones encoding a novel calcium-activated neutral proteinase from *Schistosoma mansoni*. J. Biol. Chem. **266:** 15085–15090.
2. **Ausubel, F. M., R. Brent, R. E. Kingston, D. D. Moore, J. G. Seidman, J. A. Smith, and K. Struhl.** 1991. Current protocols in molecular biology. John Wiley & Sons, New York.
3. **Blessing, M., J. L. Jorcano, and W. W. Franke.** 1989. Enhancer elements directing cell-type-specific expression of cytokeratin genes and changes of the epithelial cytoskeleton by transfections of hybrid cytokeratin genes. EMBO J. **8:**117–126.
4. **Bucher, P., and E. Trifonov.** 1986. Compilation and analysis of eukaryotic POL II promoter sequences. Nucleic Acids Res. **14:**10009–10026.
5. **Chodosh, L. A., A. S. Baldwin, R. W. Carthew, and P. A. Sharp.** 1988. Human CCAAT-binding proteins have heterologous subunits. Cell **53:**11–24.
6. **Dale, B. A., A. M. Gown, P. Fleckman, J. R. Kimball, and K. A. Resing.** 1987. Characterization of two monoclonal antibodies to human epidermal keratohyalin: reactivity with filaggrin and related proteins. J. Invest. Dermat. **88:**306–313.
7. **Dale, B. A., K. A. Holbrook, J. R. Kimball, M. Hoff, and T.-T. Sun.** 1983. Expression of epidermal keratins and filaggrin during human fetal skin development. J. Cell Biol. **101:**1257–1269.
8. **Dale, B. A., K. A. Holbrook, and P. M. Steinert.** 1978. Assembly of stratum corneum basic protein and keratin filaments in macrofibrils. Nature (London) **276:**729–731.
9. **Dale, B. A., K. A. Resing, and P. V. Haydock.** 1990. Filaggrins, p. 393–412. *In* R. D. Goldman and P. M. Steinert (ed.), Cellular and molecular biology of intermediate filaments. Plenum Press, New York.
10. **Deveroux, J., P. Haeberli, and O. Smithies.** 1984. A comprehensive set of sequence analysis programs for the VAX. Nucleic Acids Res. **12:**387–395.
11. **Dlugosz, A., and S. Yuspa.** Coordinate changes in gene expression which mark the spinous to granular cell transition in epidermis are regulated by protein kinase C. J. Cell Biol., in press.
12. **Dorin, J. R., E. Emslie, and V. van Heyningen.** 1990. Related calcium-binding proteins map to the same subregion of chromosome 1q and to an extended region of synteny on mouse chromosome 3. Genomics **8:**420–426.
13. **Dorin, J. R., M. Novak, R. E. Hill, D. J. H. Brock, D. D. Secher, and V. van Heyningen.** 1987. A clue to the basic defect in cystic fibrosis from cloning the CF antigen. Nature (London) **326:**614–617.
14. **Eckert, R., and H. Green.** 1986. Structure and evolution of the human involucrin gene. Cell **46:**583–589.
15. **Fairley, J.** 1991. Calcium: a second messenger, p. 314–328. *In* L. A. Goldsmith (ed.), Physiology, biochemistry, and molecular biology of the skin. Oxford University Press, New York.
16. **Ferrari, S., B. Calabretta, J. DeRiel, R. Battini, F. Ghezzo, E. Lauret, C. Griffin, B. Emanuel, F. Gurrieri, and R. Beserga.** 1987. Structural and functional analysis of a growth related gene, the human calcyclin. J. Biol. Chem. **262:**8325–8332.
17. **Frohman, M. A.** 1990. RACE: rapid amplification of cDNA ends, p. 28–38. *In* M. A. Innis, D. H. Gelfand, J. J. Sninski, and T. J. White (ed.), PCR protocols: a guide to methods and applications. Academic Press, New York.
18. **Gan, S.-Q., O. W. McBride, N. Markova, W. W. Idler, and P. M. Steinert.** 1990. Structure, organization and polymorphisms of the human profilaggrin gene. Biochemistry **29:**9432–9440.
19. **Gerke, V., and K. Weber.** 1985. The regulatory chain in the p36-Kd substrate complex of viral tyrosine-specific protein kinases is related in sequence to the S100 protein of glial cells. EMBO J. **4:**2917–2920.
20. **Hamilton, E. H., R. E. Payne, Jr., and E. J. O'Keefe.** 1991. Trichohyalin: presence in the granular layer and stratum corneum of normal human epidermis. J. Invest. Dermatol. **96:**666–672.
21. **Hardas, B. D., and J. T. Elder.** 1992. Physical linkage of CRABP-II and calcyclin on chromosome 1q21 by yeast artificial chromosome cloning, abstr. 105. J. Invest. Dermat. **58:**569.
22. **Harding, C. R., and I. R. Scott.** 1983. Histidine-rich proteins (filaggrins): structural and functional heterogeneity during epidermal differentiation. J. Mol. Biol. **170:**651–673.
23. **Harper, J. F., M. R. Sussman, G. E. Schaller, C. Putnam-Evans, H. Charbonneau, and A. C. Harmon.** 1991. A calcium-dependent protein kinase with a regulatory domain similar to calmodulin. Science **252:**951–954.
24. **Hohl, D., U. Lichti, D. Breitkreutz, P. Steinert, and D. Roop.** 1991. Transcription of the human loricrin gene in vitro is induced by calcium and cell density and suppressed by retinoic acid. J. Invest. Dermatol. **96:**414–418.
25. **Hohl, D., T. Mehrel, U. Lichti, M. L. Turner, D. R. Roop, and P. M. Steinert.** 1991. Characterization of human loricrin. Structure and function of a new class of cell envelope proteins. J. Biol. Chem. **266:**6626–6636.
26. **Jensen, R., D. R. Marshak, C. Anderson, T. J. Lukas, and D. M. Watterson.** 1985. Characterization of human grain S100 protein fraction: amino acid sequence of S100β. J. Neurochem. **45:**700–705.
27. **Kawasaki, H., H. Kasai, and T. Okuyama.** 1985. Protein analysis and reagents: microscale assay of calcium-binding activity of proteins and peptides using a nitrocellulose membrane. Anal. Biochem. **148:**297–302.
28. **Kligman, D., and D. Hilt.** 1988. The S100 protein family. Trends Biochem. Sci. **13:**437–443.
29. **Kretsinger, R. H.** 1980. Structure and evolution of calcium-modulated proteins. Crit. Rev. Biochem. **8:**119–174.
30. **Krisinger, J., H. Darwish, N. Maeda, and H. F. De Luca.** 1988. Structure and nucleotide sequence of the rat intestinal vitamin D-dependent calcium-binding protein gene. Proc. Natl. Acad. Sci. USA **85:**8988–8992.
31. **Laemmli, U. K.** 1970. Cleavage of structural proteins during the assembly of the head of bacteriophage T4. Nature (London) **227:**680–685.
32. **Lagasse, E., and R. Clerc.** 1988. Cloning and expression of two human genes encoding calcium-binding proteins that are regulated during myeloid differentiation. Mol. Cell. Biol. **8:**2402–2410.
33. **Leask, A., C. Byrne, and E. Fuchs.** 1991. Transcription factor AP2 and its role in epidermal-specific gene expression. Proc. Natl. Acad. Sci. USA **88:**7948–7952.
34. **Leask, A., M. Rosenberg, R. Vassar, and E. Fuchs.** 1990. Regulation of a human epidermal keratin gene: sequences and nuclear factors involved in keratinocyte-specific transcription. Genes Dev. **4:**1985–1998.
35. **Lee, S.-C., I.-G. Kim, O. W. McBride, J. G. Compton, E. J. O'Keefe, and P. M. Steinert.** 1992. The human trichohyalin gene, abstr. 448. J. Invest. Dermatol. **98:**626.
36. **Lonsdale-Eccles, J. D., J. A. Haugen, and B. A. Dale.** 1980. A phosphorylated keratohyalin-derived precursor of epidermal stratum corneum basic protein. J. Biol. Chem. **255:**2235–2238.
37. **Mack, J. W., and P. M. Steinert.** The mechanism of interaction of filaggrin with intermediate filaments: the ionic zipper hypothesis. Submitted for publication.
38. **Markova, N. G., S.-Q. Gan, and P. M. Steinert.** Unpublished data.
39. **Markova, N. G., and P. M. Steinert.** Unpublished data.
40. **Martensen, T. M.** 1984. Chemical properties, isolation and analysis of O-phosphates in proteins. Methods Enzymol. **107:** 3–23.
41. **McKinley-Grant, L. G., W. W. Idler, I. A. Bernstein, D. A. D. Parry, L. Cannizzaro, C. M. Croce, K. Huebner, S. R. Lessin, and P. M. Steinert.** 1990. Characterization of a cDNA clone encoding human filaggrin and localization of the gene to chromosome region 1q21. Proc. Natl. Acad. Sci. USA **86:**4848–4852.
42. **Pontremoli, S., and E. Melloni.** 1986. Extralysosomal protein degradation. Annu. Rev. Biochem. **55:**455–481.
43. **Resing, K. A., and B. A. Dale.** 1991. Proteins of keratohyalin, p.

148–167. *In* L. A. Goldsmith (ed.), Physiology, biochemistry, and molecular biology of the skin. Oxford University Press, New York.

44. **Resing, K. A., B. A. Dale, and K. A. Walsh.** 1985. Multiple copies of phosphorylated filaggrin in epidermal profilaggrin demonstrated by analysis of tryptic peptides. Biochemistry **24:**4167–4175.
45. **Resing, K. A., K. A. Walsh, and B. A. Dale.** 1984. Identification of two intermediates during processing of profilaggrin to filaggrin in neonatal mouse epidermis. J. Cell Biol. **99:**1372–1378.
46. **Rhim, J. S., J. Jay, P. Arnstein, F. M. Price, K. K. Sanford, and S. A. Aaronson.** 1985. Neoplastic transformation of human epidermal keratinocytes by AD12-SV40 and Kirsten sarcoma virus. Science (Washington, D.C.) **227:**1250–1252.
47. **Risse, G., K. Jooss, M. Neuberg, H. J. Bruller, and R. Muller.** 1989. Asymmetrical recognition of the palindromic AP1 binding site (TRE) by Fos protein complexes. EMBO J. **8:**3825–3832.
48. **Roeder, R. G.** 1991. The complexities of eukaryotic transcription initiation: regulation of preinitiation complex assembly. Trends Biochem. Sci. **16:**402–408.
49. **Rogers, G. E., M. J. Fietz, and A. Fratini.** 1991. Trichohyalin and matrix proteins. Ann. N.Y. Acad. Sci. **642:**64–81.
50. **Rothnagel, J. A., D. A. Greenhalgh, T. Gagne, M. A. Longley, D. L. Horak, B. Lu, and D. R. Roop.** 1992. Identification of the calcium-sensitive regulatory unit of the gene encoding human keratin I, abstr. 097. J. Invest. Dermatol. **98:**568.
51. **Rothnagel, J. A., and P. M. Steinert.** 1990. The structure of the gene for mouse filaggrin and a comparison of the repeating units. J. Biol. Chem. **265:**1862–1865.
52. **Sambrook, J., E. F. Fritsch, and T. Maniatis.** 1989. Molecular cloning: a laboratory manual, 2nd ed. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y.
53. **Saris, C. J. M., T. Kristensen, P. D'Eustachio, L. J. Hicks, D. J. Noonan, T. Hunter, and B. F. Tack.** 1987. cDNA sequence and tissue distribution of the mRNA for bovine and murine p11, the S100-related light chain of the protein-tyrosine kinase substrate p36 (calpactin). J. Biol. Chem. **262:**10663–10671.
54. **Scheidtmann, K. H.** 1989. Immunological detection of proteins of known sequence, p. 93–115. *In* J. Creighton (ed.), Protein structure. IRL Press, Oxford.
55. **Scott, I. R., and C. R. Harding.** 1981. Studies on the synthesis and degradation of a high molecular weight, histidine-rich phosphoprotein from mammalian epidermis. Biochim. Biophys. Acta **669:**65–78.
56. **Scott, I. R., and C. R. Harding.** 1986. Filaggrin breakdown to water binding compounds during development of the rat stratum corneum is controlled by the water environment. Dev. Biol. **115:**84–92.
57. **Steinert, P. M., J. S. Cantieri, D. C. Teller, J. D. Lonsdale-Eccles, and B. A. Dale.** 1981. Characterization of a class of cationic proteins that specifically interact with intermediate filaments. Proc. Natl. Acad. Sci. USA **78:**4097–4101.
58. **Steinert, P. M., and D. R. Roop.** 1988. Molecular and cellular biology of intermediate filaments. Annu. Rev. Biochem. **57:**593–625.
59. **Steinert, P. M., R. Zackroff, M. Aynardi-Whitman, and R. Goldman.** 1982. Isolation and characterization of intermediate filaments. Methods Cell Biol. **24:**399–419.
60. **Steven, A. C., M. E. Bisher, D. R. Roop, and P. M. Steinert.** 1990. Biosynthetic pathways of filaggrin and loricrin—two major proteins expressed by terminally differentiated mouse keratinocytes. J. Struct. Biol. **104:**150–162.
61. **Strynadka, N., and M. James.** 1989. Crystal structures of the helix-loop-helix calcium-binding proteins. Annu. Rev. Biochem. **58:**951–998.
62. **Suzuki, K.** 1987. Calcium-activated neutral proteases: domain structure and activity regulation. Trends Biochem. Sci. **12:**103–105.
63. **Towbin, H., T. Staehelin, and J. Gordon.** 1979. Electrophoretic transfer of proteins from polyacrylamide gels to nitrocellulose sheets: procedure and some applications. Proc. Natl. Acad. Sci. USA **76:**4350–4354.
64. **Van Eldik, L. J., J. G. Zendegui, D. R. Marshak, and D. M. Watterson.** 1982. Calcium-binding proteins and the molecular basis of calcium action. Int. Rev. Cytol. **77:**1–61.
65. **Williams, M. L., and P. M. Elias.** 1987. Genetically transmitted, generalized disorders of cornification. The ichthyoses, p. 155–178. *In* J. C. Alper (ed.), Dermatologic clinics, vol. 5. The genodermatoses. The W. B. Saunders Co., Philadelphia.
66. **Wood, L., D. Carter, M. Mills, N. Hatzenbuhler, and G. Vogeli.** 1991. Expression of calcyclin, a calcium-binding protein, in the keratogenous region of growing hair follicles. J. Invest. Dermat. **56:**383–387.
67. **Yoneda, K., D. Hohl, O. W. McBride, M. Wang, K. Cehrs, W. W. Idler, and P. M. Steinert.** 1992. The human loricrin gene. J. Biol. Chem. **267:**18060–18066.
68. **Yuspa, S. H., A. E. Kilkenny, P. M. Steinert, and D. R. Roop.** 1989. Expression of murine epidermal differentiation markers is tightly regulated by restricted extracellular calcium concentrations in vitro. J. Cell Biol. **109:**1207–1217.