Structure and Regulation of the Envoplakin Gene*

Received for publication, February 2, 2000, and in revised form, March 15, 2000
Published, JBC Papers in Press, March 29, 2000, DOI 10.1074/jbc.M001028200

Arto Määttä‡, Christiana Ruhrberg§, and Fiona M. Watt¶

From the Keratinocyte Laboratory, Imperial Cancer Research Fund, 44 Lincoln’s Inn Fields, London WC2A 3PX, United Kingdom

Envoplakin, a member of the plakin family of proteins, is a component of desmosomes and the epidermal cornified envelope. To understand how envoplakin expression is regulated, we have analyzed the structure of the mouse envoplakin gene and characterized the promoters of both the human and mouse genes. The mouse gene consists of 22 exons and maps to chromosome 11E1, syntenic to the location of the human gene on 17q25. The exon-intron structure of the mouse envoplakin gene is common to all members of the plakin family: the N-terminal protein domain is encoded by 21 small exons, and the central rod domain and the C-terminal globular domain are coded by a single large exon. The C terminus shows the highest sequence conservation between mouse and human envoplakins and between envoplakin and the other family members. The N terminus is also conserved, with sequence homology extending to Drosophila Kakapo. A region between nucleotides ~101 and 288 was necessary for promoter activity in transiently transfected primary keratinocytes. This region is highly conserved between the human and mouse genes and contains at least two different positively acting elements identified by site-directed mutagenesis and electrophoretic mobility shift assays. Mutation of a GC box binding Sp1 and Sp3 proteins or a combined E box and NF-κB element interacting with unidentified nuclear proteins virtually abolished promoter activity. 600 base pairs of the mouse upstream sequence was sufficient to drive expression of a β-galactosidase reporter gene in the suprabasal layers of epidermis, esophagus, and forestomach of transgenic mice. Thus, we have identified a regulatory region in the envoplakin gene that can account for the expression pattern of the endogenous protein in stratified squamous epithelia.

At present five members of the plakin family of proteins are well characterized. Desmoplakin (4) is an abundant component of the desmosome inner plaque that binds keratin filaments to epithelial cell-cell attachment sites (2, 5). Plectin (6, 7) is a ubiquitously expressed protein that is able to bridge intermediate filaments to microtubules and the actin cytoskeleton (8) and is found in desmosomes, hemidesmosomes, and adherens junctions (3). In epithelia BPAG1 (9) is also part of the hemidesmosome plaque (3, 5, 10), whereas the splice variants of BPAG1 found in neurons bind not only neurofilaments but also actin filaments and microtubules (11, 12).

The two newest members of the plakin family are envoplakin and periplakin (13–15), which were originally identified as components of the epidermal cornified envelope, a submembranous layer of transglutaminase cross-linked protein that contributes to the barrier properties of the outermost layers of the skin (16). Envoplakin and periplakin are also found in desmosomes, and it has been proposed that they act as an interdesmosomal scaffold on which the cornified envelope is assembled (13, 14, 17). A further contribution of envoplakin and periplakin to epidermal barrier function is the covalent attachment of ceramide lipids to these proteins (17). In addition to expression in epidermis, envoplakin and periplakin are found in other stratified squamous epithelia and in two-layered and transitional epithelia such as mammary gland and bladder (14). Although envoplakin and periplakin share the characteristic plakin domain structure, their C-terminal domains are considerably smaller than the respective domains of the other plakins, and they are unique among the plakins in having the potential to heterodimerize with each other (14).

Gene targeting of plakins in mice and characterization of certain human pathologies have underlined the importance of this protein family. Desmoplakin is crucial for the assembly or stability of desmosomes and mice without desmoplakin die at day 6.5 of embryonic development (18). Desmoplakin is haploinsufficient in man, a heterozygous null allele causing a striated palmoplantar hyperkeratosis (19). In mice lack of either plectin or BPAG1 is not embryonically lethal but causes epidermal blistering as a result of the dissociation of keratin bundles from hemidesmosomes (11, 20); in addition, the BPAG1 null animals have severe neuronal degeneration (11). Humans with autoantibodies against BPAG1 or mutations resulting in loss of plectin also have skin blistering (21, 22). All the plakins, including envoplakin and periplakin, are targets for autoantibodies in paraneoplastic pemphigus, a skin and mucosal blistering disease that develops in some patients with lymphatic malignancies (23, 24).

So far, knowledge about the regulation of plakin genes is...
limited. None of the promoters has been analyzed for tissue-specific regulatory regions in vivo, and only the BPAG1 and periplakin promoters have been characterized by reporter gene transfection in keratinocytes (25, 26). To facilitate the analysis of envoplakin function, we have cloned the mouse envoplakin gene and analyzed the envoplakin promoter in cultured keratinocytes and the epidermis of transgenic mice.

**EXPERIMENTAL PROCEDURES**

**Genomic Cloning and Sequencing**—Three λ-phage clones were isolated and purified from a 129/Sv mouse genomic library using the p210-23 human envoplakin cDNA clone (13) as a probe in Southern mapping and Southern blotting with several different envoplakin cDNA fragments were carried out to characterize the clones. Overlapping restriction fragments were subcloned into pBluescript KS II for sequencing. The three mutations create a new plasmid p-363 Mut Klf; 5

9

(9Promega). The following oligonucleotide primers were used: 5

dIII site in the pGL3 polylinker.

10 M cholera toxin, and 10 ng/ml epidermal growth serum-free conditions for at least one passage before transfection. For transient transfections 1–2 x 10

190 to yield plasmid p-363

190 to yield plasmid p-363

190 to yield plasmid p-363

190 to yield plasmid p-363

190 to yield plasmid p-363

190 to yield plasmid p-363

190 to yield plasmid p-363

190 to yield plasmid p-363

190 to yield plasmid p-363

190 to yield plasmid p-363

190 to yield plasmid p-363

190 to yield plasmid p-363

190 to yield plasmid p-363

190 to yield plasmid p-363

190 to yield plasmid p-363

190 to yield plasmid p-363

190 to yield plasmid p-363

190 to yield plasmid p-363

190 to yield plasmid p-363

190 to yield plasmid p-363

190 to yield plasmid p-363

190 to yield plasmid p-363

190 to yield plasmid p-363

190 to yield plasmid p-363

190 to yield plasmid p-363

190 to yield plasmid p-363

190 to yield plasmid p-363

190 to yield plasmid p-363

190 to yield plasmid p-363

190 to yield plasmid p-363

190 to yield plasmid p-363

190 to yield plasmid p-363

190 to yield plasmid p-363

190 to yield plasmid p-363

190 to yield plasmid p-363

190 to yield plasmid p-363

190 to yield plasmid p-363

190 to yield plasmid p-363

190 to yield plasmid p-363

190 to yield plasmid p-363

190 to yield plasmid p-363

190 to yield plasmid p-363

190 to yield plasmid p-363

190 to yield plasmid p-363

190 to yield plasmid p-363

190 to yield plasmid p-363

190 to yield plasmid p-363

190 to yield plasmid p-363

190 to yield plasmid p-363

190 to yield plasmid p-363

190 to yield plasmid p-363

190 to yield plasmid p-363

190 to yield plasmid p-363

190 to yield plasmid p-363

190 to yield plasmid p-363

190 to yield plasmid p-363

190 to yield plasmid p-363

190 to yield plasmid p-363

190 to yield plasmid p-363

190 to yield plasmid p-363

190 to yield plasmid p-363

190 to yield plasmid p-363

190 to yield plasmid p-363

190 to yield plasmid p-363

190 to yield plasmid p-363

190 to yield plasmid p-363

190 to yield plasmid p-363

190 to yield plasmid p-363

190 to yield plasmid p-363

190 to yield plasmid p-363

190 to yield plasmid p-363

190 to yield plasmid p-363

190 to yield plasmid p-363

190 to yield plasmid p-363

190 to yield plasmid p-363

190 to yield plasmid p-363

190 to yield plasmid p-363

190 to yield plasmid p-363

190 to yield plasmid p-363

190 to yield plasmid p-363

190 to yield plasmid p-363

190 to yield plasmid p-363

190 to yield plasmid p-363

190 to yield plasmid p-363

190 to yield plasmid p-363

190 to yield plasmid p-363

190 to yield plasmid p-363

190 to yield plasmid p-363

190 to yield plasmid p-363

190 to yield plasmid p-363

190 to yield plasmid p-363

190 to yield plasmid p-363

190 to yield plasmid p-363

190 to yield plasmid p-363

190 to yield plasmid p-363

190 to yield plasmid p-363

190 to yield plasmid p-363

190 to yield plasmid p-363

190 to yield plasmid p-363

190 to yield plasmid p-363

190 to yield plasmid p-363

190 to yield plasmid p-363

190 to yield plasmid p-363

190 to yield plasmid p-363

190 to yield plasmid p-363

190 to yield plasmid p-363

190 to yield plasmid p-363

190 to yield plasmid p-363

190 to yield plasmid p-363

190 to yield plasmid p-363

190 to yield plasmid p-363

190 to yield plasmid p-363

190 to yield plasmid p-363

190 to yield plasmid p-363

190 to yield plasmid p-363

190 to yield plasmid p-363

190 to yield plasmid p-363

190 to yield plasmid p-363

190 to yield plasmid p-363

190 to yield plasmid p-363

190 to yield plasmid p-363

190 to yield plasmid p-363

190 to yield plasmid p-363

190 to yield plasmid p-363

190 to yield plasmid p-363

190 to yield plasmid p-363

190 to yield plasmid p-363

190 to yield plasmid p-363

190 to yield plasmid p-363

190 to yield plasmid p-363

190 to yield plasmid p-363

190 to yield plasmid p-363

190 to yield plasmid p-363

190 to yield plasmid p-363

190 to yield plasmid p-363

190 to yield plasmid p-363

190 to yield plasmid p-363

190 to yield plasmid p-363
clones that showed highest similarity to the human envoplakin cDNA (GenBank™ accession numbers AA726169, AA727101, and AA798910) and performed genomic PCR to isolate a 3-kb fragment that, as predicted, encoded the missing part of the gene.

The mouse envoplakin gene consists of 22 exons (Fig. 1). The first 21 exons, corresponding to the large N-terminal protein domain, are small (39–194 bp in length). The sizes of these exons are perfectly conserved between the human and mouse genes. In general, the pattern of exon sizes in the envoplakin gene resembles human periplakin more than the other plakin genes: 13 out of 21 N-terminal exons are identical in size in envoplakin and periplakin. Although small N-terminal exons characterize all the known plakin genes, the sizes of the exons are not conserved among family members, except for a few exons near the end of the N-terminal head domain of the proteins (not shown).

The λ clone M210-52 was used to determine the chromosomal localization of the mouse envoplakin gene by fluorescence in situ hybridization. The Evpl locus was present as a single copy residing on chromosome 11E1 in 20 metaphases analyzed on chromosome spreads of normal mouse spleen cells (Fig. 2). This region is syntenic to the human chromosomal band 17q25, where the human envoplakin gene has been localized (33).

Conservation of Coding Sequences between Mouse and Human Envoplakin and Other Plakin Family Members—The coding sequence of the mouse envoplakin gene that we compiled predicts a polypeptide of 2035 amino acid residues (Fig. 3). This is one amino acid longer than the corresponding human protein owing to an additional residue at the C terminus before the stop codon. Mouse envoplakin is characterized by the common plakin structure: N-terminal and C-terminal domains are separated by an 828-amino acid rod domain. The boundaries of the rod domain were determined by the Coils program in the OCG package that predicts regions with a high probability of forming coiled coil structures (not shown). The same boundaries were seen when using the Coils algorithm in the Predict-Protein server. The boundaries for the central rod domain of mouse envoplakin were predicted to differ from the human envoplakin protein (13). This reflects the more interrupted structure of the envoplakin rod when compared with predicted rod domains of the other plakins, and the actual borders for the protein domains must await experimental verification.

Comparison of the mouse envoplakin sequence with human envoplakin and other plakins is shown in Table I. Mouse and human envoplakins are highly conserved. The N-terminal globular domain of the protein is similar to other plakin N termini including Kakapo, a Drosophila protein carrying domains homologous to plectin and dystrophin (34, 35). The most conserved N-terminal sequence within the whole family is shown in Fig. 4A; it includes a tyrosine residue (amino acid 210) followed by closely spaced leucines and ending with a DWSD motif. Notably this structure is present in both Drosophila and Caenorhabditis elegans Kakapo proteins (GenBank™ accession number for the cosmid that contains the coding sequence is ZK1151) and its mouse homologue, the actin cross-linking protein ACF-7 (mACF-7; Ref. 36). Interestingly, the five times repeated KGS motif in the beginning of envoplakin (Fig. 3) is conserved between mouse and human but is not found in the other plakins.

The newly defined linker domain between the rod and the C-terminal globular domains (14, 15) has the highest sequence conservation between human and mouse envoplakin (Table I). The alignment of the linker domains in different plakins (Fig. 4B) indicates the conserved and similar residues between mouse envoplakin and the other plakin family members. The linker domain is lacking in ACF-7 and in Kakapo, the C terminus of which is not homologous to plakins but to dystrophin (34–36).

Conservation of the Human and Mouse Upstream Sequences Indicates Potential Regulatory Regions—The nucleotide sequence of the putative human envoplakin promoter was determined from the cosmid ICRPc105D03119 (33). Comparison of the corresponding mouse sequence revealed a considerable degree of sequence homology (Fig. 5). Both the mouse and human sequences lacked a TATA box and had instead an initiator consensus sequence only. Conservation was highest around and just upstream from the initiator consensus sequence and in a stretch of about 150 bp starting at nucleotide −137 of the human sequence (Fig. 5). The human promoter has an Alu repetitive element further upstream from the conserved
region (from bp −550). The relative locations and sequences of several putative binding sites for transcription factors were well conserved between species. These include several GC and GT boxes for the Sp1 family of transcription factors, E boxes for helix-loop-helix factors, and two binding sites for Krüppel-like transcription factors.

Two Conserved DNA Motifs Are Necessary for High Level Reporter Gene Expression in Primary Human Keratinocytes—

The high degree of conservation between the human and mouse envoplakin upstream sequences suggested that important regulatory motifs might lie within the first few hundred base pairs of that region. To test this hypothesis we assayed promoter activity in reporter gene transfections. Both human and mouse upstream sequences were cloned into luciferase reporter gene constructs. In addition, we constructed a deletion series of the potential human promoter. These plasmids were transiently transfected into human primary epidermal keratinocytes. Each plasmid was tested in at least five independent transfections.

The longest fragment of the human upstream sequence tested, extending over 1 kb from the 5’ end of the gene, consistently gave the highest luciferase activities (Fig. 6). These values were on average 500 times higher than obtained by transfection of the empty vector alone, which indicates that the fragment contains sequences capable of high promoter (or combined promoter and enhancer) activity. The mean activity of this construct (p-1068ELuc) was designated as 100%, and the mean activities of the other constructs were calculated relative to that. The mouse promoter (p-MEPLuc; up to the HindIII site at −523 of the mouse sequence) was on average slightly more active than the longest tested human fragment (Fig. 6), possibly reflecting the presence of an Alu repetitive sequence in the human promoter upstream from the highly conserved region. The shortest fragment of the human promoter tested (p-101ELuc) was only about five times more active than the empty vector alone (Fig. 6).

In comparing the activity of a series of fragments of the human promoter intermediate in length between p-1068 and p-101, the most remarkable difference was between the shortest promoter and a construct extending to nucleotide −363 (Fig. 6). To test the importance of this region, we deleted it from the full-length promoter. The resultant plasmid, p-1068SS, had as little promoter activity as the shortest (p-101) plasmid. Thus, deletion of the 260-bp fragment rendered the envoplakin promoter inactive. This region includes that part of the upstream sequence that is most highly conserved between human and mouse genes (Fig. 5). Furthermore, even though the more distal sequences (from −363 to −1068) seemingly contain additional positive regulatory elements, the activity of these elements was dependent on the presence of the 260-bp fragment. The activity of the 260-bp fragment could be divided into two additive elements: transfection of the construct p-220ELuc resulted in approximately half the activity of the construct p-363ELuc (Fig. 6).

To further characterize the 260-bp fragment, we mutated three conserved elements harboring putative transcription factor binding sites. The mutations were designed to abolish the consensus binding sites for a Klf site at nucleotide −265 (p-Mut Klf), an overlapping E box and Klf site at −240 (p-Mut E + Klf), and an Sp1 site at −190 (p-Mut Sp1) (Fig. 5). The mutations were compared with the wild type promoter for activity in luciferase reporter gene assays (Fig. 7). The first mutation did not significantly change promoter activity, and an additional deletion construct (p-288 ELuc) confirmed that the two basic helix-loop-helix protein binding sites upstream from −265 Klf site were not needed for high level activity, which further narrowed the critical region to 187 bp between −101 and −288. Mutations p-Mut E + Klf and p- Mut Sp1 very effectively reduced promoter activity, each causing a greater than 20-fold decrease in luciferase activity compared with wild type (Fig. 7).

To study transcription factor binding to these sites, we employed electrophoretic mobility shift assays (Fig. 8). The Sp1

Table I

| Mouse envoplakin domain | Human envoplakin | Mouse periplakin | Human desmoplakin | Rat plectin | Human BPAG1 | Mouse ACP7 | Drosophila Kakapo | C. elegans Kakapo |
|------------------------|----------------|----------------|------------------|------------|-------------|-----------|----------------|----------------|
| N term | 82 | 39 | 29 | 32 | 25 | 30 | 29 | 26 |
| Rod term | 80 | 42 | 36 | 30 | 32 | 40 | 44 | 50 |
| Linker | 91 | 58 | 48 | 51 | 32 | 58 | 59 | 64 |
| C term | 90 | 46 | 52 | 44 | 32 | 60 | 63 | 69 |
Regulation of Envoplakin Expression

The mouse envoplakin gene, *Evpl*, lies on chromosome 11E1, syntenic to human chromosome 17q25, where the corresponding human gene, *EVPL*, resides (33). In both species, the envoplakin gene is proximal to the acidic keratin gene clusters that localize to 11D in mouse and 17q21–23 in man (38, 39). There are several mouse mutations with skin phenotypes, such as *bareskin* (*Bsk*) and *rex* (*Re*), that segregate with chromosome 11 (40). One mouse mutation, *Rim3*, affecting skin and hair follicles, has been mapped more accurately to the distal region of chromosome 11 (41) and envoplakin can thus be considered as a potential candidate gene for this disorder. No human diseases have so far been linked to the envoplakin gene. A skin disorder, focal nonepidermolytic palmoplantar keratoderma with esophageal cancer, is mapped to 17q25, but high resolution mapping and sequencing have excluded envoplakin as a candidate gene (32).

The mouse envoplakin gene has 22 exons, like its human counterpart (32) and the human periplakin gene (26). As evaluated by the sizes of the exons and homology at the DNA level,
envoplakin and periplakin form a pair of closely related genes that are to some extent divergent from the other plakins. This is further emphasized by the fact that at the protein level mouse and human envoplakin are most similar to periplakin. The linker domain immediately C-terminal to the rod domain is the most conserved part of the plakin family (14, 15, 23), and some paraneoplastic pemphigus autoantibodies cross-react with C-terminal fusion proteins of envoplakin, periplakin, and desmoplakin (23). The function of this part of the plakin proteins has not been demonstrated conclusively but is likely to be involved, together with the C-terminal repeats, in the interaction of plakins with intermediate filaments (42, 43). The presence of a potential protein kinase C phosphorylation site in the linker further suggests a regulatory role in protein interactions for this domain.

The putative promoters of both the human and mouse envoplakin genes lack a TATA box. As originally described for housekeeping genes, they have a possible initiator element preceded by several binding sites for transcription factors of the Sp1 family. It is of interest that the envoplakin upstream sequence shares several features with the periplakin promoter (26). They both are TATA-less, harbor several Sp-1 sites, and need sequences distal to the basal promoter for optimal expression in cultured keratinocytes. Because envoplakin and periplakin are usually co-expressed and possibly form heterodimers (14), it is conceivable that they are targets for the same signaling pathways and transcription factors.

Using transgenic mice we were able to show that only 600 bp of the mouse envoplakin promoter are needed for gene expression in suprabasal keratinocytes. Moreover, transient transfections into human primary keratinocytes demonstrated that a 187-bp element of the human upstream sequence, which is highly conserved in the mouse, is necessary for high level reporter gene expression. The 187-bp region harbors several fully conserved binding sites for transcription factors, such as Krüppel-like factors and basic helix-loop-helix factors that interact with E box sequences. Notably, no conserved activator protein-1 sites were found in this region, although members of the Fos and Jun protein families are involved in the regulation of several promoters of keratinocyte-specific genes (43–45).

The importance of the conserved Sp1-binding site in the promoter activity of the envoplakin gene is interesting in the light of recent reports on the function of Sp1 in regulating other epithelial genes. Usually, Sp1 does not determine the tissue specificity of a gene on its own but acts co-operatively with other transcription factors such as Ets (in the human transglutaminase 3 promoter; Ref. 47) or AP-2 (in the keratin K3 promoter; Ref. 48). The Sp1 family comprises four proteins, one of which, Sp3, can antagonize the action of Sp1. It has been suggested that the ratio of Sp1 to Sp3 regulates the papilloma

FIG. 5. Sequence comparison of human and mouse envoplakin promoters. A sequence comparison between the human (top) and mouse (bottom) envoplakin promoters was produced by the GAP program in the GCG package. ATG translation start site is in bold type. The beginning of the human envoplakin cDNA is in bold italics; this nucleotide was designated as +1. The putative initiator element is underlined, and some consensus binding sites for transcription factors are boxed.

FIG. 6. Deletion analysis of envoplakin gene promoter activity in human keratinocytes. Envoplakin promoter-luciferase constructs were transiently transfected in human foreskin keratinocytes, and luciferase activity was measured 24 h after transfection. p-MEPLuc is the mouse promoter; all the other constructs are human. The raw luciferase values were normalized against co-transfected β-galactosidase reference plasmid. The values are presented as percentages of the most active human promoter fragment, which was assigned the value of 100%. Each bar shows the mean and standard error of five independent transfections (three transfections in the case of pMEPLuc). The positions of putative transcription factor binding sites are shown above the graph.
virus 16 promoter during epithelial differentiation (49). Sp3 levels are higher in basal keratinocytes, and Sp1 is up-regulated in differentiating keratinocytes. These observations support the conclusion that Sp1 plays a role in activating differentiation-specific genes such as envoplakin (the present study), distal-less Dlx3 (50), and involucrin (51).

Klf4, a member of the Krüppel family of transcription factors, has recently been shown to be crucial for the barrier function of the epidermis (52). Targeted inactivation of Klf4 leads to neonatal death because of transepidermal water loss and genes regulated by Klf4 in keratinocytes encode cornified envelope proteins (52). We were able to detect a weak binding of Klf proteins to the site at -265 (data not shown) that did not affect envoplakin promoter activity in cultured keratinocytes.

In contrast, the element containing overlapping sites for basic helix-loop-helix proteins and Klf proteins that was essential for promoter activity did not interact with any of the candidate factors tested by electrophoretic mobility shift assays. These included upstream stimulatory factor-1, which has been shown to interact with the envoplakin promoter (49).

**FIG. 7.** Site-directed mutagenesis of the envoplakin upstream sequence indicates two conserved elements necessary for promoter activity. Point mutations abolishing transcription factor consensus binding sites were introduced into plasmid p-363ELuc. p-Mut Klfi removes a consensus Klfi site, p-MutE+Klf is at an combined E box and Klfi site, and p-MutSp1 is at an Sp1 site. A deletion construct p-288ELuc removed the two most distal E box sequences in p-363ELuc. The constructs were transfected into primary human keratinocytes and luciferase activities were measured 24 h later. The luciferase values were normalized against a β-galactosidase reference plasmid. The means and standard errors of four independent transfections are shown.

**FIG. 8.** Electrophoretic mobility shift assays of the two DNA elements critical for envoplakin promoter activity. Nuclear extracts from primary mouse keratinocytes were incubated with radiolabeled double-stranded oligonucleotides corresponding to DNA elements critical for promoter activity. A, the Sp1 site (Mut Sp1 at bp -190). Specific complexes were competed by 100-fold excess of unlabeled probe but not by 100-fold excess of a nonspecific probe (N.S.). Preincubation of the nuclear extracts with antibodies against Sp1 or Sp3 perturbed complex formation while preincubation with a c-Fos antibody did not affect complex formation. B, the site containing overlapping basic helix-loop-helix and Klfi consensus motifs (at -240) bound specific proteins that were not competed by excess of a nonspecific probe (N.S.). Moreover, the complexes were not affected by antibodies against upstream stimulatory factor-1 (Usf1), c-Myc, or Klfi.

**FIG. 9.** The promoter of the mouse envoplakin gene drives reporter gene expression in the epidermis of transgenic mice. Epithelial tissues of DNA-positive animals were stained for β-galactosidase activity (blue cytoplasmic staining) and counterstained with hematoxylin. A, 0.6-kb promoter. B–F, 3.4-kb promoter. A and B, back skin. Note that both promoters are active in the differentiated keratinocytes of the skin. In the hair follicles, the staining is restricted to the inner root sheath. C, esophagus. D, forestomach. E, bladder. F, mammary gland. Original magnification was 250× in all panels.
to be crucial for syndecan-1 enhancer activity in a keratinocyte cell line (29, 53), c-Myc, which is known to regulate epidermal differentiation (54), and any of the several Klf family members recognized by polyclonal antibody M-19 against KIf4 (SantaCruz). Future work is needed to determine the exact nature of the transcription factors interacting with the MutE + Klf site in the envoplakin promoter. It is possible that other Klf sites in the envoplakin promoter are used in later stages of epidermal differentiation than could be studied in our transient transfection assays.

Different regulatory regions appear to be responsible for envoplakin expression in stratified squamous epithelia compared with transitional or simple epithelia. Even the 3.4-kb envoplakin promoter failed to express β-galactosidase in mammary gland, in bladder, or in the simple epithelium of gastric mucosa. It is thus likely that other, as yet uncharacterized, regulatory regions exist in the gene. At least in the case of bladder urothelium there is evidence for highly tissue-specific regulatory regions because the promoter of the uroplakin II gene is not active in any other epithelium studied (55).

Epidermal gene expression has been the subject of considerable and continuing research (44, 56). A number of promoters of genes up-regulated during keratinocyte terminal differentiation have been characterized previously. These include cornified envelope precursors such as involucrin (51, 57), small proline-rich proteins (46), and loricrin (58, 59) and desmosomal proteins such as desmoglein-1 (60). The regions needed for epidermal expression of these genes have been elucidated in many cases. Usually, a few kilobases of the promoter are sufficient for correct expression in skin. For example, a 2.5-kb fragment of the human transglutaminase-1 gene (61) and the so-called distal regulatory region (from 1.95 to 2.5 kb upstream) of the human involucrin promoter (51) direct β-galactosidase expression correctly. Likewise, a 4.2-kb fragment of the desmoglein-1 promoter controls epidermal expression, even though it fails to act position-independently and is not sufficient for expression in other stratified epithelia (60). In the involucrin gene, on the contrary, far-upstream sequences between 6.5 and 14 kb are needed for correct expression (59). Interestingly, a short 90-bp promoter fragment of the keratin-5 gene misdirects β-galactosidase expression to suprabasal keratinocytes, even though longer far-upstream sequences between 6.5 and 14 kb are needed for correct expression (59). Armstrong, D. K., McKenna, E. K., Parkins, P. E., Green, K. J., Edery, H. A., Leigh, I. M., and Hughes, A. M. (1999) Hum. Mol. Genet. 8, 143–148.

Andris, K., Lassmann, H., Bittner, R., Shorny, S., Fassler, R., Propst, F., and Wiche, G. (1997) J. Invest. Dermatol. 111, 5143–5156.

Watt, H. M., Pulkkinen, L., Smith, F. J., Rugg, E. L., Lane, E. B., Bulnich, F., Burgess, R. E., Amano, S., Rugg, E. L., Kelsell, D. P., Bryant, S. P., Spurr, N. K., Geddes, J. F., Kirtschig, G., Milana, G., de Bono, A. G., Owaribe, K., Mcgrath, J. A., McMillan, R. J., Edery, R. A., Leigh, I. M., Christiansen, A. M., and Uitto, J. (1999) J. Invest. Dermatol. 10, 1724–1735.

Smith, F. J., Edery, R. A., Leigh, I. M., McMillan, J. R., Rugg, E. L., Kelsell, D. P., Bryant, S. P., Spurr, N. K., Geddes, J. F., Kirtschig, G., Milana, G., de Bono, A. G., Owaribe, K., Whee, G., Pulkkinen, L., Uitto, J., Mclean, W. H., and Lane, E. B. (1996) Nat. Genet. 13, 450–457.

Mahoney, M. G., Aho, S., Uitto, J., and Stanley, J. R. (1998) J. Invest. Dermatol. 111, 308–313.

Kiyokawa, C., Ruhrberg, C., Nie, Z., Karashima, T., Mori, O., Nishikawa, T., Green, K. J., Anhalt, G. J., DCirolandrea, T., Watt, F. M., and Hashimoto, K. (1998) J. Invest. Dermatol. 112, 1236–1238.

Tamai, K., Sawamura, S., Do, H. C., Tamaei, Y., Li, K., and Uitto, J. (1993) Clin. Invest. 92, 114–122.

Aho, S., Rothenberger, K., Tan, E. M., Ryoo, Y. W., Cho, B. H., Mclean, W. H., and Uitto, J. (1999) Genomics. 56, 160–168.

Watt, F. M. (1998) in Cell Biology: A Laboratory Handbook (Celis, J. E., ed) 2nd Ed., Vol. 1, pp. 113–118, Academic Press, Orlando, FL.

Jaakkola, P., Vihinen, T., Maatta, A., and Jalkanen, M. (1997) Mol. Cell. Biol. 17, 3210–3219.

Maatta, A., Jaakkola, P., and Jalkanen, M. (1999) J. Biol. Chem. 274, 9891–9896.

Echelard, Y., Vassileva, G., and McMahon, A. P. (1994) Development 120, 2213–2224.

Jaakkola, P., Kontsaari, S., Kauppi, T., Maatta, A., and Jalkanen, M. (1998) FASEB J. 12, 59–69.

Riek, J. M., Ruhrberg, C., Hennies, H., Mills, H. S., Di Ciollandrea, T., Evans, K. E., Ellis, A., Watt, F. M., Bishop, D. T., Spurr, N. K., Stevens, H. P., Leigh, I. M., Reis, A., Kelsell, D. P., and Field, J. K. (1999) Genomics 59, 234–242.

Ruhrberg, C., Williamson, J. A., Sheer, D., and Watt, F. M. (1996) Genomics 37, 381–385.

Gregory, S. L., and Brown, N. H. (1998) J. Cell Biol. 143, 1271–1282.

Strumpf, D., and Volk, T. (1988) J. Cell Biol. 143, 1299–1307.

Leung, C. L., Sun, D., Zheng, M., Knowles, D. R., and Liem, R. K. H. (1999) J. Cell Biol. 147, 1275–1285.

Ichida-Yamamoto, A., and Izuka, H. (1998) Exp. Dermatol. 7, 1–10.

Montgomery, J. C., Silberman, K. A., and Buchberg, A. M. (1997) Mamm. Genome 8, 530–539.

Sato, H., Koide, T., Sagai, T., Ishiguro, S., Imai, M., Saitou, N., and Shirosi, T. (1999) Genomics 56, 303–309.

Lyon, M. F., and Zentzon, J. F. (1986) Mouse News Lett. 74, 96.

Sato, H., Koide, T., Maeya, H., Wakana, S., Sagai, T., Umezawa, A., Ishiguro, S., Tamai, M., and Shirosi, T. (1998) Mamm. Genome 9, 26–25.

Acknowledgments—We thank Jill Williamson for chromosome FISH, the Imperial Cancer Research Fund Transgenic Unit for DNA microinjection and animal care, and David Shima for help with screening the mouse genomic library.
42. Stappenbeck, T. S., Bornslaeger, E. A., Corcoran, C. M., Lau, H. H., Virata, M. L., and Green, K. J. (1993) J. Cell Biol. 123, 691–705
43. Nikolic, B., MacNulty, E., Mir, B., and Wiche, G. (1996) J. Cell Biol. 134, 1455–1467
44. Eckert, R. L., Crish, J. F., Banks, E. B., and Welter, J. F. (1997) J. Invest. Dermatol. 109, 501–509
45. Sark, M. W., Fischer, D. F., de Meijer, E., van de Putte, P., and Backendorf C. (1998) J. Biol. Chem. 273, 24683–24692
46. Rossi, A., Jang, S. I., Ceci, R., Steinitz, P. M., and Markova, N. G. (1998) J. Invest. Dermatol. 110, 34–40
47. Lee, J. H., Jang, S.-I., Yang, J. M., Markova, N. G., and Steinert, P. M. (1996) J. Biol. Chem. 271, 4561–4568
48. Chen, T. T., Wu, R. L., Castro-Munozledo, F., and Sun, T. T. (1997) Mol. Cell. Biol. 17, 3056–3064
49. Apt, D., Watts, R. M., Suske, G., and Bernard, H. U. (1996) Virology 224, 281–291
50. Park, G. T., and Morasso, M. I. (1999) J. Biol. Chem. 274, 26599–26608
51. Crish, J. F., Zaim, T. M., and Eckert, R. L. (1998) J. Biol. Chem. 273, 30460–30465
52. Serge, J. A., Bauer, C., and Fuchs, E. (1999) Nat. Genet. 22, 356–360
53. Jaakkola, P., Maittia, A., and Jalkanen, M. (1998) Oncogene 17, 1279–1286
54. Gandarillas, A., and Watt, F. M. (1997) Genes Dev. 11, 2869–2882
55. Lin, J. H., Zhao, H., and Sun, T. T. (1995) Proc. Natl. Acad. Sci. U. S. A. 92, 679–83
56. Fuchs, E., and Byrne, C. (1994) Curr. Opin. Genet. Dev. 4, 725–736
57. Carroll, J. M., Albers, K. M., Garlick, J. A., Harrington, R., and Taichman, L. B. (1993) Proc. Natl. Acad. Sci. U. S. A. 90, 10270–10274
58. DiSepio, D., Jones, A., Longley, M. A., Bundman, D., Rothnagel, J. A., and Roop, D. R. (1995) J. Biol. Chem. 270, 10792–10799
59. DiSepio, D., Bickenbach, J. R., Longley, M. A., Bundman, D. S., Rothnagel, J. A., and Roop, D. R. (1999) Differentiation 64, 225–235
60. Adams, M. J., Reichel, M. B., King, I. A., Marston, M. D., Greenwood, M. D., Thirlwell, H., Arinamin, J., Buxton, R. S., and Ali, H. R. (1998) Biochem. J. 329, 165–174
61. Yamada, K., Matsuoka, M., Morishima, Y., Ueda, K., Tahata, K., Yasuno, H., Suzuki, M., and Yamanishi, K. (1997) Hum. Mol. Genet. 6, 2223–2231
62. Byrne, C., and Fuchs, E. (1993) Mol. Cell. Biol. 13, 3176–3190
63. Mills, A. A., Zheng, B., Wang, X. J., Vogel, H., Roop, D. R., and Bradley, A. (1999) Nature 398, 708–713
64. Yang, A., Schweitzer, R., Sun, D., Kaghad, M., Walker, N., Bronson, R. T., Tabin, C., Sharpe, A., Caput, D., Crum, C., and McKeon, F. (1999) Nature 398, 714–718
65. Morasso, M. I., Markova, N. G., and Sargent, T. D. (1996) J. Cell Biol. 135, 1879–1887
