Identification and dissection of an enhancer controlling epithelial gene expression in skin

Satrajit Sinha and Elaine Fuchs*

Howard Hughes Medical Institute, Department of Molecular Genetics and Cell Biology, University of Chicago, Chicago, IL 60637

Contributed by Elaine Fuchs, December 29, 2000

Keratins 14 and 5 are the structural hallmarks of the basal keratinocytes of the epidermis and outer root sheath (ORS) of the hair follicle. Their genes are controlled in a tissue-specific manner and thus serve as useful tools to elucidate the regulatory mechanisms involved in keratinocyte-specific transcription. Previously we identified several keratinocyte-specific DNase I hypersensitive sites (HSs) in the 5′ regulatory sequences of the K14 gene and showed that a 700-bp regulatory domain encompassing HSs II and III can confer epidermal and ORS-specific gene expression in transgenic mice *in vivo*. Although HS II harbored much of the transactivation activity *in vitro*, it was not sufficient to restrict expression to keratinocytes *in vivo*. We now explore the HS III regulatory element. Surprisingly, this element on its own confers gene expression to the keratinocytes of the inner root sheath (IRS) of the hair follicle, whereas a 275-bp DNA fragment containing both HSs II and III shifts the expression from the IRS to the basal keratinocytes and ORS *in vivo*. Electrophoretic mobility-shift assays and mutational studies of HSs III reveal a role for CACCC-box binding proteins, Sp1 family members, and other factors adding to the list of previously described factors that are involved in keratinocyte-specific gene expression. These studies highlight a cooperative interaction of the two HSs domains and strengthen the importance of combinatorial play of transcription factors that govern keratinocyte-specific gene regulation.

The epidermis and its notable appendage, the hair follicle, are stratified squamous epithelia in which unique sets of genes are turned on and off in a temporal and spatial pattern during growth and differentiation. The dividing basal cells of the epidermis transcribe the genes encoding keratins 5 and 14 (K5 and K14), whereas terminally differentiating cells in the spinous layers switch to the expression of K1 and K10 (1, 2). As cells enter the granular layer, they express the final proteins needed for the assembly of the cornified envelope that then serves as a scaffold for the secreted bilayers that constitute the epidermal barrier (1, 2). Similarly, hair-specific keratin genes are expressed in the precortical cells that produce the hair shaft, whereas the outer root sheath (ORS), contiguous with the epidermis, expresses K5 and K14 (1, 2). It has been suggested that the inner root sheath (IRS) expresses K1 and K10 (3).

Studies of the promoters and enhancers of genes expressed in skin epithelia have identified several regulatory transcription factors, some of which are expressed preferentially in epidermal keratinocytes (4–6). These factors include members of the AP2, Sp1, Ets, and AP-1 families, which interact with the regulatory regions of many skin epithelial genes (7–13). Other relevant factors in keratinocyte development and differentiation include members of the POU-domain proteins, homeobox proteins like Dlx3, the Kruppel-like factor Klf4, and NF-κB family members, all of which result in skin phenotypes when deficient or abnormally expressed in mice (14–17).

Despite a growing number of studies, the transcriptional control of keratinocyte-specific genes remains poorly understood. No single class of transcription factor has emerged as the key player in regulating keratinocyte gene expression, and evidence has been accumulating to suggest that keratinocyte-specific activity is controlled by the combinatorial action of more broadly expressed transcription factors (4, 10). However, functional demonstration has been complicated by redundancy between family members of these classes that are coexpressed in skin epithelium.

To address some of these questions we have begun a thorough and systematic study of the promoter/enhancer regions of the human *K14* gene, one of the most abundantly transcribed genes of mitotically active skin keratinocytes (9, 10). In our previous study we demonstrated that a 700-bp regulatory domain encompassing two keratinocyte-specific hypersensitive sites (HSs II and III) can direct keratinocyte-specific expression when combined with a heterologous promoter *in vitro* and *in vivo* (10). However, the HS II domain, despite its importance in *in vitro* assays, exhibited very low levels of activity *in vivo* and was not sufficient to restrict gene expression to keratinocytes in mice.

We have now probed more deeply into the analysis of this 700-bp critical enhancer region of the human *K14* gene. Specifically, we now have (*i*) defined the precise location of HS III by fine DNase I and restriction enzyme accessibility assays; (*ii*) determined that the 150 bp encompassing HS III are conserved between human and mouse and contain functional sites for factors that include, but are not limited to, CACCC-box binding factors, Sp1 family members, and a previously undefined repressor; (*iii*) found that on its own, HS III targets transgene expression to a restricted subset of skin keratinocytes in the IRS of the hair follicle, in which the complete K14 promoter/enhancer is not active normally; and (*iv*) discovered that when combined with HS II (125 bp), HS III (150 bp) shifts its activity from IRS-specific to epidermal-specific expression. These data reveal that tissue- and differentiation-specific transcription of skin epithelial genes is governed in a combinatorial fashion by complex regulatory elements, each of which on their own possess keratinocyte-preferred but not proper differentiation-specific activity. Thus, a complex network of transcription factors is required for strict keratinocyte- and differentiation-stage-specific expression of the skin epithelial genes.

Materials and Methods

DNase I HSs and Restriction-Enzyme Accessibility Assays. The existence of keratinocyte-specific DNase I HSs was identified previously (10). Restriction enzyme accessibility assays were performed as for DNase I assays except that nuclei from human keratinocytes were digested with 0, 2, 5, 10, or 20 units of *Pst*I

Abbreviations: K, keratins; ORS, outer root sheath; IRS, inner root sheath; HS, hypersensitive site; β-gal, β-galactosidase; EMSA, electrophoretic mobility-shift assay; TK, thymidine kinase; MT, mutation.

*To whom reprint requests should be addressed at: Howard Hughes Medical Institute, Department of Molecular Genetics and Cell Biology, University of Chicago, 5841 South Maryland Avenue, Room N314, Chicago, IL 60637. E-mail: nliptak@midway.uchicago.edu.

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. §1734 solely to indicate this fact.

and *Apa*I for 30 min at 37°C before genomic DNA was purified and digested subsequently with *Hin*dIII and *Eco*RV. Southern hybridizations were performed with Zeta-Probe GT membrane by using the manufacturer's protocol (Bio-Rad).

Generation of Transgenic Mice, β-Galactosidase (β-gal) Staining, and Immunohistochemistry. Transgenic mice were generated as described (18). Transgenic embryos from foster mothers or from F_1-generation females were isolated at embryonic day 15.5 or 16.5, washed in PBS, and fixed in tissue-fixation buffer (Specialty Media, Lavallette, NJ) for 20–30 min. After several washes with PBS, embryos were stained with 5-bromo-4-chloro-3-indolyl-β-D-galactoside for 12 h. After staining, embryos were washed in PBS several times and embedded in paraffin. Tissue samples from tails and back skin of adult transgenic mice were embedded in optimal cutting temperature compound and frozen sections (10 μm) were fixed in 0.5% glutaraldehyde for 2 min before assaying for β-gal activity. Sections then were counterstained briefly with eosin and/or hematoxylin. Immunofluorescence microscopy was performed as described (19) by using either AE15 antibody (19) or anti-β-gal antibody (Harlan Sera-Lab, Loughborough, England).

DNA Constructs for Transfection and Transgenics. Recombinant plasmids were constructed by using DNA fragments generated by restriction enzyme digestion or PCR and verified by sequencing. Constructs for transgenic mice were cloned into the modified pNASSβ vector (10). The LTK vector for transient transfections has been described (10).

Nuclear Extracts and Electrophoretic Mobility-Shift Assays (EMSAs). Nuclear extracts from human and mouse keratinocytes were prepared, and EMSAs were performed as described (10). Competition assays were performed by incubating the nuclear extracts with 20- or 100-fold excess cold oligonucleotides before addition of radiolabeled probe; 1 μg of poly(dI-dC) or poly(dA-dT) was added to each reaction as nonspecific DNA. For supershift assays, nuclear extracts were incubated with 1 μl of the antibody for 15 min before radiolabeled oligonucleotides were added to the reaction. Antibodies for Sp1 and Sp3 were purchased from Santa Cruz Biotechnology and Geneka Biotechnology (Montreal, Canada), respectively.

Cell Culture and Transient-Transfection Assays. Human primary keratinocytes were grown in serum-free keratinocyte growth medium (Clonetics, Walkersville, MD), and mouse keratinocytes were grown in low Ca^{2+} medium (10). HepG2 cells (American Type Culture Collection) and primary human fibroblasts were grown under standard conditions (10). Transient transfections were carried out by using FuGENE6 transfection reagent (Roche Molecular Biochemicals) according to the manufacturer's instructions. Cells plated in 6-well plates were transfected with 2 μg of the luciferase reporter plasmid and 1 μg of a cytomegalovirus-LacZ plasmid (internal control). Cells were harvested 36 h after transfection, washed with PBS, and lysed with luciferase reporter lysis buffer (Promega). Luciferase assays were performed with luciferase assay system (Promega) by using a luminometer. β-gal values were obtained by using a chemiluminescent reporter gene assay system (Galactolight kit, Tropix, Bedford, MA).

Results

Fine Mapping of Keratinocyte-Specific DNase I HSs II and III of the Human K14 Upstream Region and Restriction Enzyme Accessibility Assays. Previously we identified four DNase I HSs within the 5′ upstream sequence of the human *K14* gene (10). Because of their close proximity to each other, additional mapping of HSs II and III using DNase I hypersensitivity assays and Southern blot analysis was necessary to define the location of HS III precisely. As in our previous study, we compared the HSs of primary human keratinocytes with two non-K14-expressing cell types: HepG2, a liver cell line, and primary fibroblasts. Nuclei from these three cell types were incubated with increasing amounts of DNase I, and the extracted genomic DNA was digested with *Eco*RV (−1,000) and *Hin*dIII (−2,200). By using a probe at the 3′ end of this fragment (−1,300 to −1,000), we refined the positions of HSs II and III at −1,300 to −1,400 and −1,600 to −1,700, respectively, as shown in Fig. 1. To confirm these positions, we also took advantage of two restriction enzyme sites within the HS III region, *Pst*I at −1,710 and *Apa*I at −1,660 and performed *in vivo* restriction enzyme accessibility analysis on nuclei isolated from human keratinocytes. As shown in Fig. 1*B*, these two enzymes digested the genomic DNA to generate the predicted fragment, suggesting that these recognition sequences within this region of the chromatin were in an open accessible state. These results define the cis-regulatory region corresponding to keratinocyte-specific HS III at ≈−1,700 to −1,600.

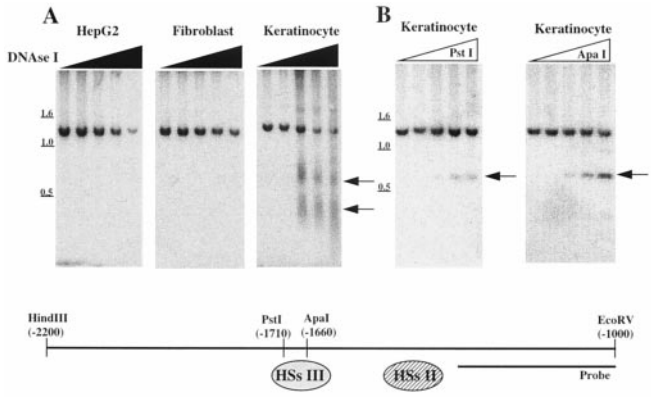

Fig. 1. DNase I HSs and restriction enzyme accessibility in the chromatin of the 5′ flanking region of the *K14* gene. Nuclei from fibroblasts, HepG2, and keratinocytes were treated with increasing amounts of DNase I (closed ascending triangle), and isolated DNAs were digested with *Hin*dIII and *Eco*RV and subjected to Southern blotting with a probe as indicated (*Lower*) in *C*. The arrowheads in *A* and *B* indicate the relative position of the HSs. The molecular weight markers (in kilobases) are indicated on the left.

HS III Targets Gene Expression to a Subset of Keratinocytes, Whereas Both HSs II and III Are Sufficient to Confer Epidermal-Specific Expression *in Vivo*. Previously we showed that the HS II segment (−1,450 to −1,325) on its own gave rise to weak expression in the epidermis of transgenic mice and some ectopic expression in dermis and bone (10). In the present study, we focused on the role of HS III and began by engineering transgenic mice harboring HS III sequences coupled to the heterologous thymidine kinase (TK) promoter to drive a LacZ reporter gene in transgenic mice. Initially, two different constructs were tested, 195 × 4TKLacZ, containing four copies of −1,752 to −1,557, and 150 × 4TKLacZ, containing four copies of −1,707 to −1,557 of the upstream K14 sequences. Five transgenic animals were generated and analyzed at 15.5–16.5 days of gestation. Irrespective of the construct used, all the transgenic animals analyzed exhibited a similar and distinctive transgene-expression pattern. Staining was pronounced and restricted largely to the whisker follicles (Fig. 2 *A* and *B*). A cross section of the nose region of the embryo revealed strong β-gal staining that was confined exclusively to the IRS cells (Fig. 2*C*). This result was surprising, because the expression of the endogenous *K14* gene in hair follicle is limited to the ORS cells. To explore this unusual phenomenon further, we created four lines of the 150 × 4TKLacZ transgenic mice. In all these adult mice, β-gal staining

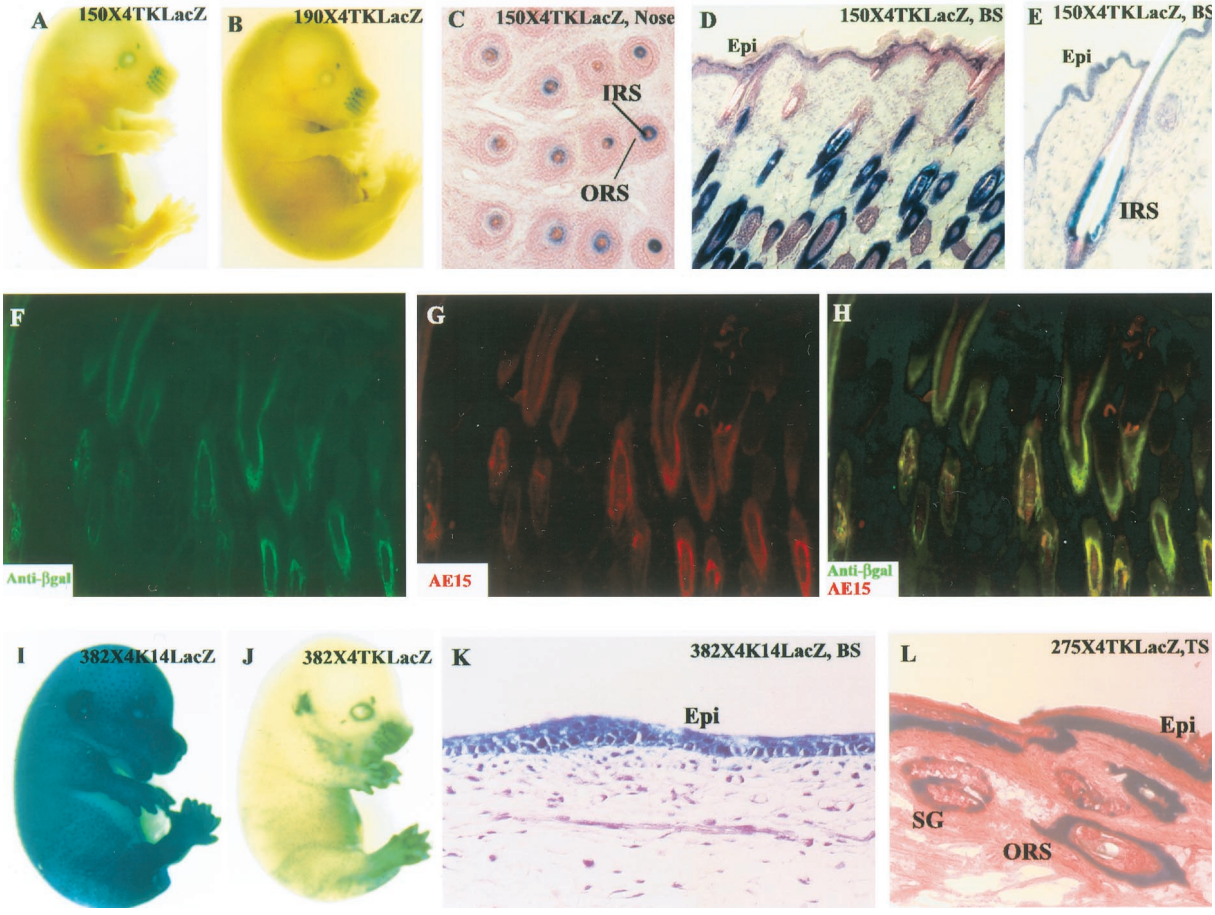

Fig. 2. Expression of various transgenes in embryos and adult tissues (see Table 1 for details). Embryos and skin sections were processed for β-gal activity by using 5-bromo-4-chloro-3-indolyl-β-D-galactoside cleavage as an assay. Transgene expressed is indicated in the upper right of each frame. (*A* and *B*) β-gal activity in 15.5-day embryos showing intense blue staining in whisker follicles. (*C*) Section of nose skin from animal shown in *A*. (*D* and *E*) Staining of the hair follicle. (*F–H*) Immunostaining with antibodies against tricohyalin (AE15) and β-gal. (*I* and *J*) β-gal activity in 15.5-day embryos. (*K*) Staining of the epidermis from the animal in *I*. (*L*) LacZ staining of the tail skin of a 1-month-old animal. Epi, epidermis; SG, sebaceous gland; BS, back skin; TS, tail skin.

was very pronounced in the IRS of back skin and whisker hair follicles, but staining was absent in the dermis or interfollicular epidermis; in some regions, weak staining was observed in the granular layers (Fig. 2 *D* and *E*). Double immunofluorescence with antibodies against β-gal and trichohyalin, a well characterized component of the IRS cells, confirmed the identity of these cells as IRS (Fig. 2 *F–H*). A similar expression pattern was observed when the K14 minimal promoter was used instead of the TK heterologous promoter (150 × 4K14LacZ; see data compiled in Table 1). Thus, the sequences of the HS III region of the K14 enhancer had information that was sufficient on its own to target gene expression to skin epithelium. However, rather than directing expression to mitotically active epidermal keratinocytes and ORS, the element directed expression to IRS, a differentiation stage that normally excludes K14 enhancer activity.

Previously we had shown that 700 bp of the K14 sequences (−2,000 to −1,300) targeted reporter gene expression to epidermis (10). To assess whether the activity sufficient for epidermis-specific gene expression was contained within HSs II and III, we designed two constructs, 382 × 4TKLacZ and 382 × 4K14LacZ, containing sequences from −1,707 to −1,325 with the heterologous TK and K14 minimal promoters, respectively. To strengthen the expression of the reporter, four copies of each segment were used. Transgenic embryos displayed remarkable tissue-specific gene expression confined to the epidermis (Fig. 2 *I* and *J*). Staining was seen in multiple layers of embryonic day-15.5 to -16.5 embryo (Fig. 2*K*), a feature also characteristic of K5 and K14 expression at these early developmental stages (20). The K14 minimal promoter seemed more effective in conferring robust expression than the TK promoter.

To delineate the relevant sequences further, another construct was created, 275 × 4TKLacZ, containing just the sequences from the HS II (−1,450 to −1,325) and HS III (−1,707 to −1,557) along with the TK minimal promoter. Transgenic mice harboring this construct displayed expression in the basal epidermal layer and in the ORS and sebaceous glands of the hair follicle (Fig. 2*L*). Overall, expression was highly tissue-specific and was not seen in surrounding mesenchyme, cartilage, or other nonepithelial cell types. Thus, although HSs II and III each exhibited either an altered expression profile or ectopic expression on their own, together they contained all of the information for highly tissue- and differentiation-specific gene expression in a pattern that markedly resembled that of the endogenous *K14* gene.

Sequence Conservation of the HS III and DNA-Binding Factors. Previously we showed that the human K14 HSs II (−1,325 to −1,450) region displayed a higher level of sequence identity with the corresponding mouse HSs than was seen in the surrounding sequences (10). This high degree of sequence conservation also extended to the HS III region (−1,752 to −1,557), which shared

Table 1. Summary of β-galactosidase expression in the transgenic embryos and adult mouse harboring various LacZ constructs that were used for regulatory analysis

Construct	Min Promoter	No. of Exp Tgs /Total Tgs	LacZ Expression
195X4TKLacZ -1,752 -1,557	TK	3/5	IRS
150X4TKLacZ -1,707 -1,557	TK	5/9	IRS
150X4K14LacZ -1,707 -1,557	K14	4/5	IRS
275X4TKLacZ -1,707 -1,557 -1,450 -1,325	TK	3/7	Epidermis*
382X4TKLacZ -1,707 -1,325	TK	2/6	Epidermis
382X4K14LacZ -1,707 -1,325	K14	8/10	Epidermis

Transgenic embryos were scored as positive or negative after overnight incubation with 5-bromo-4-chloro-3-indolyl β-D-galactoside and visual inspection for blue staining. For constructs displaying transgene expression in the skin, internal tissues were scanned for ectopic expression of the reporter gene. Shown are the number of embryos and/or mice that expressed LacZ in the epidermis and the total number of transgenic mice.

*Ectopic expression in the bone and cartilage was observed only for one line of the 275X4TKLacZ.

≈81% sequence identity with the corresponding HS III from the 5′ upstream region of the mouse *K14* gene (Fig. 3). The conservation of sequence and DNase HSs was intriguing, especially in light of the fact that the human HSs II and III functioned faithfully in the mouse.

Analysis of the Transcription Factors Involved in HS III. Computer analysis of the HS III region suggested putative binding sites for many transcription factors (Fig. 3). To explore whether these sites actually bind keratinocyte nuclear proteins, we conducted EMSAs by using radiolabeled oligonucleotides and mouse and human keratinocyte nuclear extracts. Initially, we generated six oligonucleotides covering the sequences from −1,710 to −1,557 as shown in Fig. 4*A*. Of the six oligonucleotides, four of them showed specific DNA protein complexes and were chosen for further characterization.

One group of DNA–protein complexes bound to oligonucleotide 1. EMSAs with this labeled oligonucleotide showed three complexes that were competed by excess of wild-type oligonucleotides as well as by oligonucleotides mutation (MT) 1, containing mutations in the GATA element (Fig. 4*B*, lanes 2 and 3 and 6 and 7, respectively). In contrast, a 20- or 100-fold excess of mutant oligonucleotide MT2, containing mutations in the 3′ region of the oligonucleotide, failed to compete the two complexes (Fig. 4*B*, lanes 4 and 5). The lower band was competed with random oligonucleotides (data not shown), suggesting that it was probably a nonsequence-specific DNA–protein complex. The site occupied by the two specific complexes did not match to the binding sites of any known transcription factors.

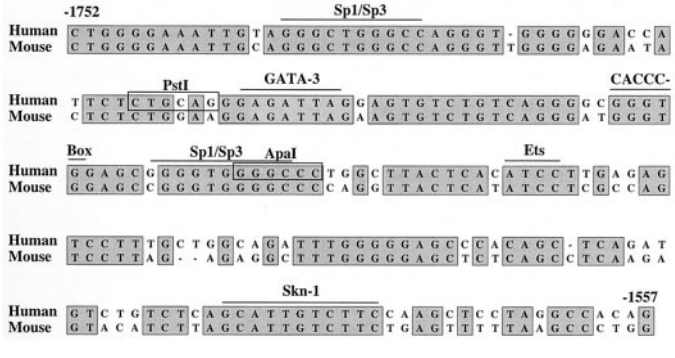

Fig. 3. Sequence identity. Sequence alignment of the human HS III region and the corresponding region of the mouse. Sequence alignment was performed by using the CLUSTALW program of MacVector. The overall sequence similarity is 81% over this stretch of DNA. The putative transcription factor binding sites are indicated. The *Pst*I and *Apa*I restriction enzyme sites are boxed.

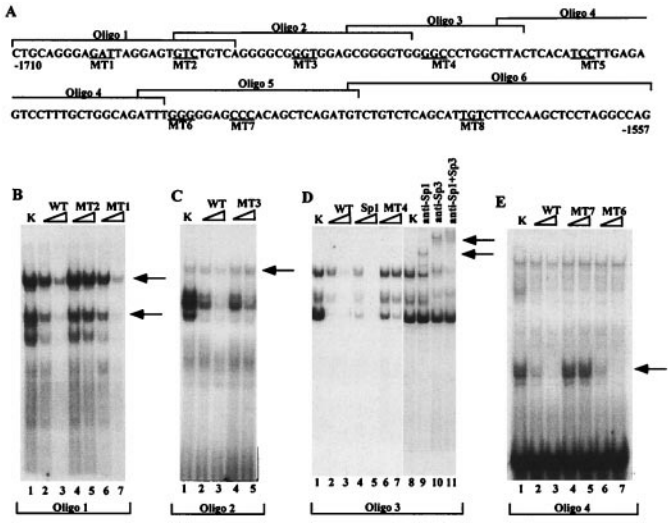

Fig. 4. EMSAs of nuclear extracts showing binding of transcription factors to the HS III region. (*A*) The upper-strand sequences of the oligonucleotides (1–6) are shown. The mutated sequences are underlined. MT1 (GAT > CGC), MT2 (GTC > AAA), MT3 (GGT > AAA), MT4 (GGC > AAA), MT5 (TCC > AAA), MT6 (GGG > AAA), MT7 (CCC > AAA), and MT8 (TGT > AAA). MT5 and -8 were used in experiments described in Fig. 5. (*B*) EMSAs were performed with radiolabeled oligonucleotide 1 and nuclear extracts from keratinocytes (K). Competition experiments with different oligonucleotides were performed by using radiolabeled oligonucleotide 1 (WT) and a 20- or 100-fold excess of unlabeled double-stranded oligonucleotides (WT, MT1, and MT2) as indicated over each lane. (*C*) EMSAs were performed as in *B*. Arrow indicates the specific complex. (*D*) EMSAs were performed by incubating oligonucleotide 3 with keratinocyte nuclear extracts. Competitor used in each EMSA is indicated over each lane. Arrows indicate supershifts, reflective of complexes between Sp1 and/or Sp3, DNA, and antibody. Consensus Sp1 oligonucleotides were obtained from Promega. (*E*) EMSAs were performed with oligonucleotide 5. Specific complex is shown with an arrow.

Oligonucleotide 2 also complexed specifically with nuclear proteins (Fig. 4*C*). This oligonucleotide contained a conserved CACCC-box sequence (GGGTG in the upper strand). Although wild-type oligonucleotides competed all of the complexes, mutations of the CACCC box (MT3) failed to compete the uppermost complex (shown with an arrow), while effectively competing the lower complexes (Fig. 4*C*, lanes 2–5). In addition, EMSAs conducted with mutant oligonucleotide showed loss of the upper complex (data not shown). Because the lower complexes also were competed with nonspecific oligonucleotides, they probably represent nonspecific DNA–protein complexes (data not shown). Antibody against Klf4, an epidermally expressed CACCC-box binding factor did not affect the upper complex. Thus, either Klf4 is not the CACCC-box binding protein that interacts with this site or the antibodies are not effective at recognizing Klf4 when complexed to DNA. At present, the identity of this CACCC-box protein remains unknown.

Sequences corresponding to oligonucleotide 3 contained a GC-rich binding site for Sp1 family members, and this site was found to interact with these proteins (Fig. 4*D*). Two major

complexes were detected, both of which were competed by wild-type and consensus Sp1 oligonucleotides but not with the oligonucleotides with the Sp1-binding site mutated (Fig. 4*D*, lanes 2–7). To confirm the identity of these complexes we incubated the reactions with antibodies against two of the Sp1-family members, Sp1 and Sp3, which are highly expressed in keratinocytes. Both these antibodies were capable of producing a supershift on the DNA–protein complexes formed with the wild-type oligonucleotides (Fig. 4*D*, lanes 9–11, arrow). Collectively these data suggest that there is a bona fide Sp1/Sp3 binding site within this segment of the enhancer.

A DNA–protein complex also was detected by oligonucleotide 5. To test the specificity of this DNA–protein complex, two mutant oligonucleotides were generated, MT6 and MT7, which had three base-pair mutations each. Competition experiments revealed that although wild-type oligonucleotides and MT6 competed, the MT7 oligonucleotide did not compete this complex (Fig. 4*E*). This result suggested that the complex was specific and interacted with the GC-rich region of this segment. However, a transcription-factor binding-site database search did not reveal any similarities with any known binding sites. Further studies will be needed to identify the exact transcription factor(s) that binds to this element.

Finally, computer analysis also revealed a binding site (GGA, core) for an Ets factor in oligonucleotide 4. Despite the fact that functional Ets sites have been implicated in many epidermal-specific gene regulatory elements, including one in the HS II region (4, 10, 21), oligonucleotide 4 did not produce an EMSA under the conditions tested. Oligonucleotide 6 corresponding to the 3′ end of the sequence also failed to detect any specific complex in an EMSA, although it contained consensus sequences for Skn-1-like factors (22). It is still possible that these regions may bind to transcription factors that are not detectable in our assays. Although not exhaustive, these EMSA experiments were sufficient to reveal that HS III is complex and contains multiple sequences that can bind multiple keratinocyte nuclear factors.

Mutational Analysis Reveals Functional Importance of the CACCC Box, SP1 Sites, and a Repressor Element Within the HS III Region. Having elucidated several potential binding sites for CACCC-box factors, Sp1/Sp3, and other previously uncharacterized factors, we used mutagenesis to test the functional role of these binding sites on the transcriptional activity of the K14 enhancer element in transient-transfection experiments. Previously we found that the HS III region does not activate reporter luciferase activity at high levels unless coupled to the HS II (10). Therefore, we engineered mutations in the 195-bp HS III region (−1,752 to −1,557) and linked this segment to the 125 bp (−1,450 to −1,325) of HS II and to the TK minimal promoter driving luciferase (320TKLuc, Fig. 5). We tested the effects of mutations that abolished the binding of specific DNA–protein as well as those that abolished conserved potential transcription-factor binding sites, which had not revealed DNA–protein complexes. Of these, the MT3 mutation of the CACCC box decreased transcriptional activity of the 320TKLuc transgene by approximately 65% in transiently transfected keratinocytes. The MT4 mutation of the Sp1/Sp3 site also exhibited a deleterious effect, dropping activity by 50%. Conversely, the mutation that abolished the complex binding to oligonucleotide 5 resulted in a 2-fold increase in reporter activity, suggesting that this site may interact with a repressor. MT1 had virtually no effect on the activity, whereas MT2 had a modest 25% reduction. MT5, which abolished a potential Ets site, and MT8, which abolished a Skn-1-binding site, did not have any effect on the enhancer activity, in agreement with the lack of any DNA-binding in our EMSA experiments. Together, our findings compiled in Fig. 5 suggest that HS III contains at least three key regulatory elements: a CACCC box, an Sp1/Sp3-binding site, and a sequence motif for the binding of a previously uncharacterized repressor.

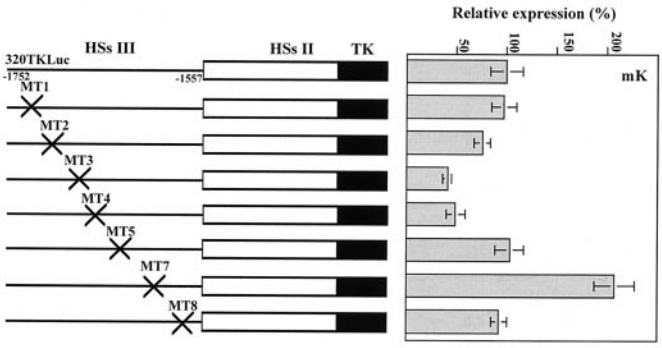

Fig. 5. Delineating the active elements within the HS III K14 enhancer element. 320TKLuc was constructed to contain sequences −1,752 to −1,557 and −1,450 to −1,325 from the K14 upstream region. These sequences were introduced into the pLTK vector (10). Transfections were conducted with mouse keratinocytes (mK). Cell extracts were assayed subsequently for luciferase and β-gal values as described in *Materials and Methods*. Mutant plasmids contained the mutations as described in Fig. 4. After harvesting the cells, luciferase activities were determined and normalized against the β-gal values. Each experiment was repeated at least three times. Results shown are the averages of three experiments with standard deviations. The activity of the 320TKLuc was set at 100%.

Discussion

Identification of DNase I HSs has led often to the subsequent discovery of key cis-regulatory elements and/or locus control regions for many genes such as those of the β-globin clusters (23). In this and our previous study, we used chromatin structure mapping in primary keratinocytes in identifying potentially important regulatory segments of the 5′ upstream region of epidermally expressed genes. We now have analyzed both the HS II and HS III regions of the human *K14* gene and examined how these elements act in concert to coordinate skin epithelial-specific gene expression.

During embryonic development, pluripotent ectoderm gives rise to both epidermis and hair follicles. Within the hair follicle are three concentric regions: the ORS, the IRS, and the shaft (24). The ORS is contiguous with the basal layer of the epidermis and shares similar biochemistry. The IRS and the hair shaft differentiate upward from dividing matrix cells in the hair bulb (24). Presumably, a combination of intrinsic and environmental factors controls the production of transcription factors that enable skin epithelial cells to differentiate along separate pathways.

We were surprised to find that HS III on its own contains intrinsic information to direct expression of target genes specifically to the IRS of skin epithelium. Because the *K14* gene is expressed naturally in the basal layer of the epidermis and the ORS, rather than the IRS, a hitherto unrecognized complexity of differentiation-specific regulation is unveiled. Moreover, because transgene expression in other tissues was not detected, the information necessary for keratinocyte-specific gene expression seemed to be contained within this element. The identification of a small 150-bp DNA segment that can target expression to the IRS specifically now enables us not only to probe the mechanism underlying differentiation-specific gene expression in the IRS further, but also to target expression of foreign genes to this set of cells *in vivo*.

The 125-bp HS II on its own exhibited consistent and prominent expression in bone/cartilage and dermis in addition to the epidermis (10). Given this individual behavior and that of HS III, it was particularly interesting to find that both elements in

combination resulted in epidermal-specific and ORS-specific gene expression. The developmental pattern generated with this combined element was similar to that of the endogenous *K14* gene. Thus, sequences encompassing HS II and HS III together can achieve remarkable tissue specificity, strengthening the notion that epidermal-specific gene expression is governed by multiple sequences and factors.

Conservation of enhancer sequences across species has been valuable in identifying putative regulatory sequences and transcription-factor binding sites. Both the HS II and HS III regions show a high degree of sequence conservation between the human and mouse *K14* genes, and the HS III region contains conserved sequence motifs for the binding of Sp1 and CACCC-box binding factor families. Involvement of Sp1/Sp3 has been implicated in the regulation of several epidermal and/or ORS genes including keratins *K5*, *K16*, *K17*, and *K19* (4, 7, 21). Sp1 binds to GC-rich sequences, and its binding sites have been identified in the proximal promoter region of the *K5* gene (7). Given the ubiquitous expression of Sp1 proteins, it seems most likely that Sp1 acts cooperatively with other transcription factors in specifying keratinocyte gene expression. CACCC-box binding factor families include some that are relatively tissue- and differentiation-restricted such as erythroid cell-enriched Kruppel-like factor 1 (Klf1), lung enriched (Klf2), and in the epidermis, Klf4 (16, 25, 26). Interestingly, it has been shown recently that Klf4 and Sp1 together modulate K19 promoter activity and contribute to its keratinocyte-specific expression (21). Thus, the identification of functional Sp1 and CACCC-box binding sequences in the HS III elements underscore the importance of these families of proteins in skin epithelial gene expression.

When taken together with our prior studies on the K14 promoter/enhancer, we now know that binding sites AP2, Sp1, AP-1, Ets, CACCC-box binding factors, and a few as yet unidentified factors all contribute to proper temporal and spatial patterns of transcriptional activity in skin epithelium. To assess the specific role of each factor will necessitate testing in transgenic mice the effects of individual mutations in each binding motif. Although preliminary evaluation of mutants in cultured keratinocytes can be helpful, our experience has told us that it is not necessarily reliable in predicting the key regulatory elements involved. Given that the skin possesses >20 different cell types, the impact and complexity of environmental cues on keratinocyte-specific gene expression perhaps is not surprising.

In conclusion, the identification of a bona fide epidermal enhancer element is an important step in understanding the molecular mechanisms underlying keratinocyte-specific gene expression. In addition, our partial dissection of this element has begun to unravel some of the mysteries underlying differentiation specificity and begins to explain why so many of the same elements are found in epidermally expressed genes irrespective of differentiation stage. Our studies now set the stage for further research exploring the role of these groups of transcription factors and the role of chromatin in mediating keratinocyte-specific and differentiation-specific gene expression.

We thank Linda Degenstein and Jie Fan for generation of the transgenic mice and Cedith Copenhaver for technical support. We thank Grazina Traska for excellent cell-culture help. S.S. is currently the recipient of a postdoctoral fellowship from the National Institutes of Health. E.F. is an Investigator of the Howard Hughes Medical Institute. This work was supported by funds from the Howard Hughes Medical Institute and from the National Institutes of Health, RO1-AR31737.

1. Fuchs, E. (1993) *J. Cell Sci.* **17,** 197–208.
2. Fuchs, E. (1995) *Annu. Rev. Cell Dev. Biol.* **11,** 123–153.
3. Stark, H. J., Breitkreutz, D., Limat, A., Ryle, C. M., Roop, D., Leigh, I. & Fusenig, N. (1990) *Eur. J. Cell Biol.* **52,** 359–372.
4. Byrne, C. (1997) *BioEssays* **19,** 691–698.
5. Eckert, R. L., Crish, J. F., Banks, E. B. & Welter, J. F. (1997) *J. Invest. Dermatol.* **109,** 501–509.
6. Fuchs, E. & Byrne, C. (1994) *Curr. Opin. Genet. Dev.* **4,** 725–736.
7. Byrne, C. & Fuchs, E. (1993) *Mol. Cell. Biol.* **13,** 3176–3190.
8. Leask, A., Byrne, C. & Fuchs, E. (1991) *Proc. Natl. Acad. Sci. USA* **88,** 7948–7952.
9. Leask, A., Rosenberg, M., Vassar, R. & Fuchs, E. (1990) *Genes Dev.* **4,** 1985–1998.
10. Sinha, S., Degenstein, L., Copenhaver, C. & Fuchs, E. (2000) *Mol. Cell. Biol.* **20,** 2543–2555.
11. Maatta, A., Ruhrberg, C. & Watt, F. M. (2000) *J. Biol. Chem.* **275,** 19857–19865.
12. Lee, J. H., Jang, S. I., Yang, J. M., Markova, N. G. & Steinert, P. M. (1996) *J. Biol. Chem.* **271,** 4561–4568.
13. Jang, S. I., Steinert, P. M. & Markova, N. G. (1996) *J. Biol. Chem.* **271,** 24105–24114.
14. Andersen, B., Weinberg, W. C., Rennekampff, O., McEvilly, R. J., Bermingham, J. R., Jr., Hooshmand, F., Vasilyev, V., Hansbrough, J. F., Pittelkow, M. R., Yuspa, S. H. & Rosenfeld, M. G. (1997) *Genes Dev.* **11,** 1873–1884.
15. Morasso, M. I., Markova, N. G. & Sargent, T. D. (1996) *J. Cell. Biol.* **135,** 1879–1887.
16. Segre, J. A., Bauer, C. & Fuchs, E. (1999) *Nat. Genet.* **22,** 356–360.
17. Seitz, C. S., Lin, Q., Deng, H. & Khavari, P. A. (1998) *Proc. Natl. Acad. Sci. USA* **95,** 2307–2312.
18. Vassar, R., Rosenberg, M., Ross, S., Tyner, A. & Fuchs, E. (1989) *Proc. Natl. Acad. Sci. USA* **86,** 1563–1567.
19. DasGupta, R. & Fuchs, E. (2000) *Development (Cambridge, U.K.)* **126,** 4557–4568.
20. Byrne, C., Tainsky, M. & Fuchs, E. (1994) *Development (Cambridge, U.K.)* **120,** 2369–2383.
21. Brembeck, F. H. & Rustgi, A. K. (2000) *J. Biol. Chem.* **275,** 28230–28239.
22. Pal, S., Lo, M. C., Schmidt, D., Pelczer, I., Thurber, S. & Walker, S. (1997) *Proc. Natl. Acad. Sci. USA* **94,** 5556–5561.
23. Fraser, P. and Grosveld, F. (1998) *Curr. Opin. Cell Biol.* **10,** 361–365.
24. Hardy, M. H. (1992) *Trends Genet.* **8,** 55–61.
25. Turner, J. & Crossley, M. (1999) *Int. J. Biochem. Cell Biol.* **31,** 1169–1174.
26. Wani, M. A., Wert, S. E. & Lingrel, J. B. (1999) *J. Biol. Chem.* **274,** 21180–21185.