

# A CACCC Box-like *cis*-Regulatory Element of the Epstein-Barr Virus ED-L2 Promoter Interacts with a Novel Transcriptional Factor in Tissue-specific Squamous Epithelia\*

(Received for publication, December 4, 1996, and in revised form, April 8, 1997)

Hiroshi Nakagawa‡, Takuya Inomoto‡, and Anil K. Rustgi†§¶

From the ‡Gastrointestinal Unit and §Hematology-Oncology Unit, Massachusetts General Hospital, Harvard Medical School, Boston, Massachusetts 02114

The Epstein-Barr (EBV) virus induces a lytic state after infecting epithelial cells. Subsequently, there is infection of B lymphocytes with two types of cycles, latent and lytic. Apart from linkage of the EBV latent membrane protein-1 (LMP-1) with benign and malignant conditions of squamous epithelial cells, little is known about other EBV gene products that may be important in these processes as well as cellular transcriptional factors that regulate EBV gene expression in these epithelial cells. The EBV ED-L2 promoter, an early lytic cycle promoter, is located upstream of a transcription start site for a short open reading frame designated BNL2 and just downstream of the BNL1 (LMP-1) open reading frame. We have previously used the EBV ED-L2 promoter to target oncogenes in transgenic mice, resulting in tissue-specific expression in the tongue, esophagus, forestomach, and skin, all sharing stratifying squamous epithelia, alternatively called keratinocytes. In the present study, we have functionally dissected the ED-L2 promoter by making deletion constructs fused to the luciferase reporter gene with transient transfections into squamous and nonsquamous epithelial cell lines as well as B lymphocytes. A CACCC box-like *cis*-regulatory element has been identified that is located between -218 and -187 base pairs of the ED-L2 promoter that confers significant promoter activity only in squamous epithelial cells. This *cis*-regulatory element is active in a heterologous minimal herpes simplex virus thymidine kinase promoter reporter gene construct when transfected into squamous epithelial cells but not in nonsquamous epithelial cells. DNA gel mobility shift assays have led to the identification of DNA-protein complexes that bind the CACCC box-like element. One of these proteins is a novel transcriptional factor that is uniquely active in stratified squamous epithelial cells, designated as keratinocyte specific factor (KSF). KSF may be related to Sp1 but appears to be distinct from Sp1. In addition, KSF may interact with related or identical *cis*-regulatory elements found in human papilloma-virus-11 E6 and cytokeratin K3 promoters that are ac-

tive in keratinocytes. In aggregate, KSF may be important in the transcriptional regulation of viral and eukaryotic genes in keratinocytes.

Epstein-Barr virus (EBV)<sup>1</sup> is a double-stranded, enveloped DNA herpes virus with host specificity restricted to humans and nonhuman primates. The EBV genome is over 170 kilobases existing in a circular or a linear duplex form. Among the encoded proteins are nuclear antigens (EBNA) and latent infection membrane proteins (LMP). EBV may infect oropharyngeal epithelial cells by virtue of association between the external viral glycoprotein 350/220 and a CR-2 like receptor on host cells (1, 2). After entry into the epithelial cells, EBV initiates a lytic infection that in turn leads to infection of circulating B lymphocytes that traffic through the nasopharynx. EBV undergoes two types of cycles in B lymphocytes. The first is a lytic infection where a high copy number of EBV genomes is initiated, followed by production of viral particles that can be released from the host cell. In contrast, the second consists of a latent infection in which the circular EBV genome is episomal.

While much investigation has sought to elucidate EBV's role in the immortalization and transformation of B lymphocytes, thereby leading to lymphoproliferative diseases (1) such as Burkitt's lymphoma, Hodgkin's disease, T-cell lymphoma, among others, very little is known about the molecular mechanisms underlying EBV's lytic infection of oropharyngeal epithelial cells. It is known that latent membrane protein-1 (LMP-1) can induce transformation of rodent fibroblasts resulting in growth in reduced serum conditions, loss of contact inhibition, anchorage independence, and tumor formation in nude mice (3). LMP-1 can additionally block differentiation in epithelial cells that can be induced to differentiate terminally in a manner similar to normal keratinocytes (4, 5). Furthermore, LMP-1 transcripts are associated with nasopharyngeal carcinoma (6). In this context, although LMP-1 does play a significant role, the repertoire of genes of EBV that are involved in producing the lytic cycle in epithelial cells remains to be elucidated, and the cellular transcriptional factors in the epithelial cells that govern the transcriptional regulation of these genes require further investigation.

The EBV ED-L2 promoter, an early lytic cycle promoter, was originally defined as one of the TATA boxes found in *EcoRI* D/*Bam*HI N fragments of the EBV genome and is located 30 base pairs upstream of a transcription start site for a short open reading frame designated BNL2 whose function remains

\* This work was supported by the ADHF/AGA Fiterman and Funderberg Awards (to A. R.), an American Cancer Society Faculty Research Award (to A. R.), National Institutes of Health Grant DK40561 (to A. R.), and Department of Energy Award DE-FG-2-91-ER61228 (to H. N. and A. R.). Oligonucleotides were synthesized through the CSIBD Molecular Biology Core National Institutes of Health Grant DK43351. The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked "advertisement" in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

¶ To whom correspondence should be addressed: GI Unit, Jackson 904, Massachusetts General Hospital, 50 Blossom St., Boston, MA 02114. Tel.: 617-724-3740; Fax: 617-726-3673; E-mail: rustgi@helix.mgh.harvard.edu.

<sup>1</sup> The abbreviations used are: EBV, Epstein-Barr virus; LMP, latent membrane protein; bp, base pair(s); PCR, polymerase chain reaction; KSF, keratinocyte-specific factor; WT, wild-type; MT, mutant; Tricine, N-[2-hydroxy-1,1-bis(hydroxymethyl)ethyl]glycine; GMSA, gel mobility shift assays.

TABLE I  
Oligonucleotide primer sequences for plasmid constructions

*Bam*HI restriction enzyme site is underlined.

Primer	Sequence	Construct
<i>Xho</i> I-L2 AS	5'-GAAGATCTCGAGAGTGAGGCACAGCTG-3'	pL2-782
<i>Bam</i> HI-L2 S4	5'-CGGGATCCATGGCGCGGTGATCCACA-3'	pL2-782
<i>Bam</i> HI-L2 S8	5'-CGGGATCCTGGGAATGCCTTAGCTAATCA-3'	pL2-610
<i>Bam</i> HI-L2 S7	5'-CGGGATCCTGTCTCCACCCAGTAACT-3'	pL2-287
<i>Bam</i> HI-L2 S6	5'-CGGGATCCAAGCCACACCTAATCAT-3'	pL2-218
<i>Bam</i> HI-L2 S9	5'-CGGGATCCAGGAACACCTGTTGTTGA-3'	pL2-187
<i>Bam</i> HI-L2 S10	5'-CGGGATCCTTCTTTGCGCATAAGCAC-3'	pL2-164
<i>Bam</i> HI-L2 S11	5'-CGGGATCCTAATCCCTCTCTCACA-3'	pL2-144
<i>Bam</i> HI-L2 S5	5'-CGGGATCCGCTAGCCAAAACCTCCA-3'	pL2-114
Luciferase AS	5'-CTTTATGTTTTTGGCGTCTTCCA-3'	

to be elucidated (7). As depicted in Fig. 1A, a noncoding sequence containing the TATA element, hence designated the ED-L2 promoter region, is located just downstream of another open reading frame, BNL1, that encodes the latent membrane protein 1 (LMP1).

When a *Mln*I-*Bam*HI fragment of the EBV genome containing coding sequences for BNL1 and -2 as well as flanking sequences was expressed under the BNL1/BNLF2 promoters and a polyoma virus enhancer in transgenic mice, the BNL2 gene transcript was expressed uniquely in tissues sharing a stratified squamous epithelium such as the tongue, esophagus, and to a lesser extent the skin, whereas the BNL1 transcript was expressed ubiquitously (8).

As a means of understanding oncogenesis in the oral cavity and upper gastrointestinal tract (esophagus) with stratified squamous epithelia, we have engineered transgenic mice in which 782 bp of the ED-L2 promoter was fused to the cyclin D1 complementary DNA without any additional enhancers, yielding the expression of cyclin D1 in the tongue, esophagus, forestomach, and to a lesser extent skin (9). In addition, the transgene expression was localized immunohistochemically to the basal and suprabasal layers of the stratified squamous epithelia but was not detectable in other layers of the epithelium or layers underneath the epithelium. These findings led to the hypothesis that the ED-L2 promoter is active in the oral cavity and upper gastrointestinal (esophageal) stratified squamous epithelial cells, or keratinocytes, by virtue of cell type-specific interactions between *cis*-acting regulatory elements in this promoter and cellular *trans*-acting nuclear factors possibly without the need for transactivation by other EBV gene products. We have identified in the present study a CACCC box-like *cis*-regulatory element in the EBV-ED L2 promoter that interacts with several nuclear transcriptional factors, one of which appears to be keratinocyte-specific. This in itself is a novel finding for EBV's role in epithelial cells, but placed in a broader context, the same nuclear transcriptional factors interact with similar or identical *cis*-regulatory elements of the human papillomavirus-11 E6 promoter and some of the cytokeratin promoters that are active in keratinocytes.

#### MATERIALS AND METHODS

**Plasmid DNA Constructions**—A 782-base pair (bp) sequence extending between 168,268 and 167,487 nucleotides of the B95-8 Epstein-Barr virus genome, according to the numbering method of Baer *et al.* (10), contains the ED-L2 promoter (Fig. 1A). This promoter has been fused to the human cyclin D1 complementary DNA, and this transgene (designated pL2HD1) has been expressed in mice (9). The 782-bp fragment was isolated by the polymerase chain reaction (PCR) using pL2HD1 as a DNA template with a sense primer from 168,268 and an antisense primer from 167,487 (see Table I for primer sequences), each of which was synthesized with a *Bam*HI site in the 5'-end of the sense primer and an *Xho*I site in the 5'-end of the antisense primer. The PCR reaction was done in 1 × Native *Pfu* buffer (Stratagene) consisting of 20 mM Tris-HCl, pH 8.2, 10 mM KCl, 6 mM (NH<sub>4</sub>)<sub>2</sub>SO<sub>4</sub>, 2 mM MgCl<sub>2</sub>, 0.1% Triton X-100, and 10 μg/ml nuclease-free bovine serum albumin sup-

plemented with 200 μM each dNTP, 0.6 μM sense and antisense primers, 2.5 units of *Pfu* DNA polymerase (Stratagene), and 10 ng of plasmid DNA template. The PCR amplification conditions consisted of denaturation at 94 °C for 1 min, annealing at 55 °C for 1 min, and extension at 72 °C for 1 min for 20 cycles. After PCR amplification, the reaction product was digested with *Bam*HI and *Xho*I, agarose gel-purified, and ligated into the luciferase promoterless vector, pXP2 (11), to generate the pL2-782 plasmid. A subsequent series of deletion constructs (pL2-610, pL2-287, pL2-218, pL2-187, pL2-164, pL2-144, and pL2-114) were made in a similar fashion using pL2-782 as a template for PCR with sense primers designed at the different positions of the promoter and an antisense primer from +24 of the coding region of the luciferase reporter gene (Table I). The pL2-435 was generated by digestion of pL2-782 with *Bam*HI and *Sma*I, deleting 347 bp of the promoter sequence and followed by religation of the plasmid. Plasmids were purified by a modified alkaline lysis method (Qiagen plasmid kit and QIAprep spin plasmid kit).

**Minimal Promoter DNA Constructions**—Minimal promoter DNA constructs containing the wild-type (WT) or mutant (MT) nucleotides spanning -218 to -184 of the pL2-782 were generated by ligation of kinased double-stranded synthetic oligonucleotides into the *Bam*HI site of the heterologous thymidine kinase promoter vector, pT81, with the luciferase reporter gene (11).

**Oligonucleotides**—All oligonucleotides were synthesized by the phosphoramidite procedure (Applied Biosystems) and purified by gel electrophoresis.

**DNA Sequencing**—All plasmid and minimal promoter DNA constructions were verified by DNA sequencing with the dideoxy-mediated chain termination method using the Sequenase version 2.0 DNA sequencing kit (U. S. Biochemical Corp.).

**Tissue Culture Cell Lines**—Human esophageal squamous carcinoma cell lines TE-11, TE-12 (gift of Dr. T. Nishihira), T.T (JCRB 0262, Japanese Cancer Research Resources Bank), and HCE4 (gift of Dr. C. C. Harris) were grown as described previously (12). Panc-1 (ATCC CRL 1739), HepG2 (ATCC HB 8065), HeLa (ATCC CCL 2), and AGS (ATCC HTB 22), obtained from the American Type Culture Collection (ATCC Rockville, MD), were cultured under standard conditions, namely Dulbecco's modified Eagle's medium (Sigma) supplemented with 10% fetal calf serum (BioWhittaker), 100 units/ml penicillin, and 100 μg/ml streptomycin (Sigma). A human B cell line, BJAB (gift of Dr. B. Cherayil), was grown in RPMI 1640 medium (Sigma) with 10% fetal calf serum (BioWhittaker), 100 units/ml penicillin, and 100 μg/ml streptomycin (Sigma). Human skin (SCC-13) and tongue (SCC-25) squamous cell carcinoma cell lines (gifts of Dr. J. Rheinwald) were grown in a 1:1 mixture of Ham's F12 medium and Dulbecco's modified Eagle's medium (Sigma), supplemented with 0.4 μg/ml hydrocortisone (Sigma) and 10% fetal calf serum, and 100 units/ml penicillin and 100 μg/ml streptomycin (Sigma) (13).

**Luciferase and Human Growth Hormone Assays**—Transient transfection of the plasmid and minimal DNA constructions in cultured cells was done by the calcium phosphate method (14) using the CaHPO<sub>4</sub> transfection kit (5' → 3', Inc., Boulder, CO), except for BJAB that was transfected by electroporation using the Gene Pulser system (Bio-Rad). For the calcium phosphate transfections, cells were plated at a density of 1 × 10<sup>6</sup> cells/35-mm well and transfected 24 h later with 2 μg of the luciferase reporter plasmid and 2 μg of pXG5, a plasmid containing the mouse metallothionein-I promoter fused to the human growth hormone gene (Nichols Institute), in a 250 μl of solution consisting of 125 mM CaCl<sub>2</sub>, 25 mM Hepes, pH 7.05, 0.75 mM Na<sub>2</sub>HPO<sub>4</sub>, 5 mM KCl, 140 mM NaCl, 6 mM glucose. After an 8-h incubation, SCC-13, SCC-25, Panc-1, HepG2, and HCE4 cells were subjected to a glycerol (15%) shock for 3

TABLE II  
Sense sequences of double-stranded oligonucleotides of the EBV ED-L2 promoter used for GMSA

Nucleotide positions in the promoter are indicated in parentheses after each of the wild-type sequences. Altered nucleotides are underlined. Note the CACACCT motif in DWT, FWT, and EWT probes. Wild-type (DWT) and mutated (DMT 1–5) oligonucleotides were also fused to the herpes simplex virus thymidine kinase promoter in the pT81 vector.

CWT	5'-CTCATGCCAGCAGAGGCAGG-3'	(-203 to -184)
DWT	5'-CCAAGCCACACCTAACTCATGCCAGCAGAGGCAGG-3'	(-218 to -184)
DMT 1	5'-CCAAGCCACACCT <u>T</u> CAGATTGCCAGCAGAGGCAGG-3'	
DMT 2	5'-CCAAGAA <u>T</u> GTTAGAACTCATGCCAGCAGAGGCAGG-3'	
DMT 3	5'-CGTCC <u>C</u> CACACCTAACTCATGCCAGCAGAGGCAGG-3'	
DMT 4	5'-CCAAGCA <u>A</u> CACCTAACTCATGCCAGCAGAGGCAGG-3'	
DMT 5	5'-CCAAGCCACAC <u>A</u> TAACTCATGCCAGCAGAGGCAGG-3'	
EWT	5'-CCAAGCCACACCTAACTCATG-3'	(-218 to -198)
FWT	5'-AGCCACACCTAA-3'	(-215 to -204)
FMT 1	5'- <u>C</u> CCACACCTAA-3'	
FMT 2	5'-AGT <u>C</u> ACACCTAA-3'	
FMT 3	5'-AGTAA <u>C</u> ACCTAA-3'	
FMT 4	5'-AGC <u>A</u> CACCTAA-3'	
FMT 5	5'-AGCC <u>G</u> CACCTAA-3'	
FMT 6	5'-AGCCAT <u>A</u> CCTAA-3'	
FMT 7	5'-AGCCACG <u>C</u> CTAA-3'	
FMT 8	5'-AGCCACA <u>T</u> CTAA-3'	
FMT 9	5'-AGCCACA <u>T</u> ATAA-3'	
FMT 10	5'-AGCCACAC <u>A</u> TAA-3'	
FMT 11	5'-AGCCACAC <u>C</u> CAA-3'	
FMT 12	5'-AGCCACACCT <u>C</u> C-3'	

min, washed three times with phosphate-buffered saline (BioWhittaker), and fresh medium was exchanged. The other cell lines (TE-11, TE-12, T.T, AGS, and HeLa) were transfected in an identical fashion but without glycerol shock as this was not found to be optimal for the cells. For electroporation of BJAB,  $6 \times 10^6$  cells were incubated with 15  $\mu$ g each of the plasmids in 600  $\mu$ l of complete medium on ice for 10 min and then electroporated at 450 V and 125 microfarads. All cells were harvested for the luciferase assay at 48 h post-transfection. The luciferase assay was performed in the following manner. Cells were washed twice with phosphate-buffered saline and lysed in 200  $\mu$ l of  $1 \times$  cell culture lysis reagent (Promega), and 40  $\mu$ l of the lysate was mixed with 100  $\mu$ l of luciferase assay reagent consisting of 20 mM Tricine, 1.07 mM  $MgCO_3$ , 2.67 mM  $MgSO_4$ , 0.1 mM EDTA, 33.3 mM dithiothreitol, 270  $\mu$ M coenzyme A, 530  $\mu$ M ATP, and 470  $\mu$ M luciferin. The luciferase activities were measured with a monolight luminometer (Analytical Luminescence Laboratory), and transfection efficiency was normalized by quantitatively measuring the human growth hormone secreted by transfected cells according to the manufacturer's directions as provided in the HGH-TGES 100T Kit (Nichols Institute). Transfections were carried out at least three times and variation between experiments was not greater than 15%.

**Nuclear Extracts and Gel Mobility Shift Assays (GMSA)**—Nuclear extracts from cultured cells were prepared essentially as described by Schreiber *et al.* (15) except the buffers were supplemented with a mixture of 0.5  $\mu$ g/ml protease inhibitors (aprotinin, chymostatin, pepstatin) (Boehringer Mannheim). The protein concentration was determined by the Bradford assay (16). To make the  $^{32}P$ -labeled oligonucleotide DNA probes, 5 pmol of a double-stranded oligonucleotide was radiolabeled by the Klenow fill-in reaction in a buffer consisting of 10 mM Tris-HCl, pH 7.5, 5 mM  $MgCl_2$ , 7.5 mM dithiothreitol, 33  $\mu$ M dATP, 33  $\mu$ M dGTP, 33  $\mu$ M dTTP, 0.33  $\mu$ M [ $\alpha$ - $^{32}P$ ]dCTP (NEN Life Science Products), 1 unit of DNA polymerase I Klenow fragment (Amersham Corp.) and then polyacrylamide gel-purified. GMSAs were carried out by incubating 5  $\mu$ g of nuclear extract with 5 fmol of the  $^{32}P$ -labeled oligonucleotide DNA probe (20,000 cpm) in a 20- $\mu$ l binding reaction containing 10 mM Tris-HCl, 50 mM NaCl, 1 mM dithiothreitol, 1 mM EDTA, 10% glycerol, and 1.0  $\mu$ g of poly(dA-dT)/poly(dA-dT) (Pharmacia Biotech Inc.). After incubation at room temperature for 15 min, the samples were loaded onto a 6% polyacrylamide, 0.25  $\times$  Tris borate gel and electrophoresed at 10 V/cm for 2 h.

For competition experiments, the nuclear extract was preincubated with 100-fold excess of unlabeled oligonucleotides prior to the addition of the  $^{32}P$ -labeled oligonucleotide DNA probe (see Table II for sequences of competitor oligonucleotides). Immune supershift assays were performed using a polyclonal anti-Sp1 antibody (Santa Cruz), a monoclonal anti-AP2 antibody (gift of Dr. T. Williams), a monoclonal anti-Sp3 antibody (gift of Dr. G. Suske), and a polyclonal anti- $\alpha$  basonuclin antibody (gift of Dr. H. Green) and preincubated with the nuclear extract at room temperature for 10 min prior to the addition of the  $^{32}P$ -labeled oligonucleotide DNA probe, as described previously (17).

**Ultraviolet Cross-linking**—To make an oligonucleotide DNA probe

for UV cross-linking with nuclear proteins, 2.5 pmol each of a sense single-stranded oligonucleotide extending between -218 and -184 (probe D), 5'-gatCAAGCCACACCTAACTCATGCCAGCAGAGGCAGG-3', and an antisense oligonucleotide, 5'-CCTCTGCTGGCATGAG 3', extending between -184 and -199 were annealed and subjected to a Klenow fill-in reaction in a fashion similar to the preparation of the oligonucleotide DNA probe for the GMSA, except [ $\alpha$ - $^{32}P$ ]dGTP (NEN Life Science Products) and 5-bromo-2'-deoxyuridine 5'-triphosphate (Sigma) were incorporated instead of dGTP and dTTP, respectively. 75  $\mu$ g of crude nuclear extract from TE-11 cells was incubated with the oligonucleotide DNA probe (100,000 cpm) in a 150- $\mu$ l binding reaction, which was increased in volume but otherwise equivalent to that for GMSA reaction, at room temperature for 30 min, and then UV-irradiated at 310 nm at room temperature for 15 min. UV cross-linking experiments were performed in an identical fashion with probes for Keratin 3 and HPV-11 E6 (sequences are in Fig. 9A). The UV-irradiated DNA-protein complexes of interest were electrophoretically fractionated in a 6% polyacrylamide gel, isolated by electroelution at 4  $^{\circ}C$ , and analyzed on a 10% SDS-polyacrylamide gel.

## RESULTS

**The EBV ED-L2 Promoter Is Active in Keratinocytes Derived from Esophageal and Skin Stratified Squamous Epithelia**—Since 782 bp of the EBV ED-L2 promoter proved to be sufficient in achieving gene expression in a tissue-specific fashion, this region was subjected to functional dissection by transient transfection of various cell lines with a series of plasmid constructs in which the full-length or serially deleted EBV ED-L2 promoter sequences were ligated to the luciferase reporter gene (Fig. 1A).

A prominent transcriptional activity of the EBV ED-L2 promoter was observed in an esophageal squamous cell carcinoma cell line, TE-11, by transfection of the full-length EBV ED-L2 promoter-luciferase reporter gene construct, designated pL2-782 (Fig. 1B), nearly 2-fold more active than the potent promoter, pRSV-luc. The EBV ED-L2 promoter was also active, although to a lesser extent, in another esophageal squamous carcinoma cell line, T.T (Table III). Importantly, activity was also observed in the SCC-13 skin squamous carcinoma cell line (Table III). However, the EBV ED-L2 promoter was weakly active in the esophageal HCE-4 and TE-12 cell lines and tongue SCC-25 cell line (Table III). The ED-L2 promoter's activity may reflect the degree of differentiation in these cell lines, as depicted by the expression of cytokeratins. In this context, TE-11, T.T, TE-12, and SCC-13 express K5 and K14, typically expressed in proliferating basal cells, whereas HCE-4 and

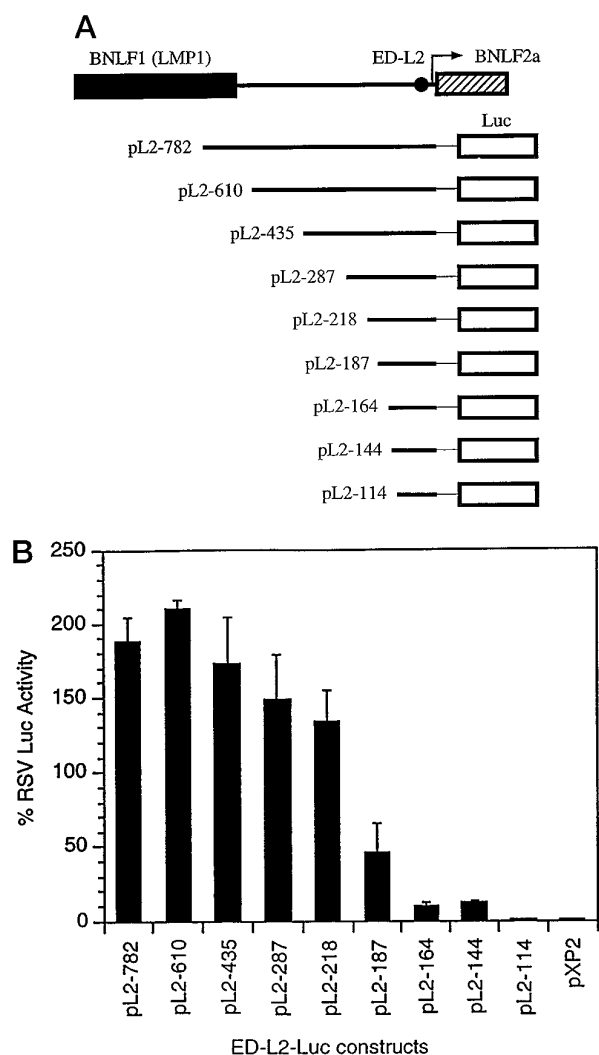


FIG. 1. A, structure of the EBV ED-L2 promoter region with a series of deleted promoter-luciferase reporter gene constructs. The ED-L2 promoter region is flanked by the 3'-end of a coding sequence for BNLF1 (LMP1) and 180 base pairs of an open reading frame for BNLF2a. A transcription start site (vertical bar with a horizontal arrow) and a TATA box (TATAAAA) (closed circle) are located at 9 base pairs and 39 base pairs upstream of the initiator AUG codon, respectively. A sequence of 782 base pairs of the ED-L2 promoter region spanning nucleotide positions 168,268 and 167,487 of the EBV genome by the numbering system of Baer *et al.* (10) was subcloned to 5'-luciferase reporter gene in a basic promoterless vector, pXP2, as described under "Materials and Methods." Boxes indicate coding sequences. B, deletion analysis of the ED-L2 promoter in an esophageal squamous carcinoma cell line, TE-11. The ED-L2 luciferase reporter gene constructs were transfected into TE-11 cells. L2 luciferase activity was assayed after 48 h and expressed as a percentage of Rous sarcoma virus luciferase activity. Activities in all transfection experiments represent the mean  $\pm$  S.D. of a minimum of four transfections done in parallel. Each transfection experiment was independently repeated at least four times.

SCC-25 do not have the same keratin profile (18).<sup>2</sup> SCC-25 was the only tongue cell line tested, and we cannot exclude the possibility of the EBV-ED L2 promoter being more active in other cell lines derived from the tongue which would be commensurate with the pattern in transgenic mice.

The EBV ED-L2 promoter was inactive in cells of nonsquamous epithelial origin (Table III) including the stomach (AGS), liver (HepG2), and pancreas (Panc-1), as well as B lymphocytes (BJAB). Minimal activity was evident in HeLa, a simple epi-

thelial derived cervical cancer cell line (Table III). Cytokeratins K5 and K14 are not expressed in HepG2 and HeLa cells (19), consistent with the notion that cell type and its differentiation status may affect EBV ED-L2 promoter activity.

The TE-11 cell line was selected for testing deletions of the EBV ED-L2 promoter. Deletional analysis of the EBV ED-L2 promoter in TE-11 cells revealed maximum activity with the pL2-610 construct (Fig. 1B). Promoter activity continuously declined from -610 to -218 suggesting the presence of multiple positive *cis*-regulatory elements. However, the most obvious loss of activity was observed between -218 and -114, indicating that the intervening sequences contain *cis*-regulatory elements that account for nearly 70% total activity. This tendency was also observed in T.T, TE-12, and SCC-13 cells (Table III). pL2-114 yielded basal activity comparable with that of the promoterless pXP2. Further functional mapping of *cis*-regulatory elements was carried out with TE-11 cells, indicating the presence of *cis*-regulatory element(s) in the sequence between -218 and -187 and also between sequences -187 and -164 (Fig. 1B). We concentrated our studies on the sequences between -218 and -187 since the greatest reduction in promoter activity occurred in this region.

*Cell Type-specific EBV ED-L2 Promoter Activity Is Associated with a CACCC Box-like cis-Regulatory Element Residing between Nucleotide Positions -218 and -187*—The region between nucleotide positions -218 and -187, designated arbitrarily as region D, contains two sequences that resemble known *cis*-acting regulatory elements, namely one similar to the CACCC box and another that has 70% homology to KER1 (Fig. 2 and Tables IV and V). The CACCC box was identified originally in promoters of the rabbit, human, and mouse  $\beta$ -globin genes. The CCACACCC *cis*-regulatory element plays an essential role in the transactivation of the  $\beta$ -globin promoter through its interaction with the erythroid cell-specific erythroid Krüppel-like factor as well as a ubiquitous factor Sp1 (20, 21). The CACCC core motif also has been found in viral promoters or enhancers, for example human papillomavirus-11 E6 promoter (22) and the SV40 enhancer (23, 24). Whereas the CACCC motif is more prevalent in promoters, another motif for consideration is KER1, a palindromic sequence, 5'-GCCTG-CAGGC-3', that was identified originally in the 5'-regulatory region of the cytokeratin K14 gene (19).

The functional consequences of the isolated region D from the EBV ED-L2 promoter were tested in a heterologous promoter system. Wild-type and mutated sequences of region D, spanning nucleotide positions -218 and -184 (Table II), were ligated 5' to an enhancerless, minimal thymidine kinase promoter luciferase reporter gene construct, designated pT81, and these constructs were transfected into TE-11 cells. As shown in Fig. 2A, the transcriptional activity of region D in the pT81 vector was enhanced in TE-11 cells from 10- to 40-fold, depending upon the number of ligated fragments. These data indicate that region D is active in a heterologous promoter system in TE-11 cells. In contrast, no enhanced activity was observed in HepG2 cells (Fig. 2B), lending further evidence to the notion that the ED-L2 promoter is active in a cell type-specific fashion.

The *cis*-regulatory element within region D was functionally mapped to the CACCC-like motif. As shown in Fig. 2B, a construct pT81-DMT2, containing a block mutation of the CACCC like motif, was entirely inactive compared with the pT81-DWT (wild type) construct. Furthermore, point mutations created within the CACCC-like motif (pT81-DMT4 and pT81-DMT5) appeared sufficient in the inactivation of region D, suggesting that those altered nucleotides are critical for interaction between the *cis*-regulatory element and *trans*-acting factor(s). However, a block mutation in the 5'-flanking

<sup>2</sup> H. Nakagawa and A. K. Rustgi, unpublished data.

TABLE III  
The EBV ED-L2 promoter activity in various cell lines

The full-length (pL2-782) and two of the deletion constructs (pL2-218 and pL2-114) were transfected in various cell lines to determine the EBV ED-L2 promoter activity by luciferase assay. The data are represented as % of RSV luciferase activity. RSV luciferase activities in each cell line were at least 200-fold over background levels observed in pXP2. Averages  $\pm$  S.D. of 3–6 independent transfections are represented. ND, not done.

Cell lines	Constructs			
	pL2-782	pL2-218	pL2-114	pXP2
T.T.	32.0 $\pm$ 6.7	16.8 $\pm$ 7.0	0.3 $\pm$ 0.0	0.1 $\pm$ 0.0
TE-12	3.4 $\pm$ 0.6	1.9 $\pm$ 0.1	0.0 $\pm$ 0.0	0.0 $\pm$ 0.0
HCE-4	1.0 $\pm$ 0.3	ND	ND	0.0 $\pm$ 0.0
SCC-13	20.6 $\pm$ 1.9	26.3 $\pm$ 6.5	0.3 $\pm$ 0.0	0.5 $\pm$ 0.0
SCC-25	3.1 $\pm$ 1.3	ND	ND	0.2 $\pm$ 0.0
HeLa	3.8 $\pm$ 0.5	1.9 $\pm$ 0.2	0.6 $\pm$ 0.0	0.4 $\pm$ 0.0
AGS	0.3 $\pm$ 0.0	ND	ND	0.0 $\pm$ 0.0
HepG2	0.0 $\pm$ 0.0	0.0 $\pm$ 0.0	0.0 $\pm$ 0.0	0.0 $\pm$ 0.0
Panc-1	0.1 $\pm$ 0.0	0.0 $\pm$ 0.0	0.0 $\pm$ 0.0	0.0 $\pm$ 0.0
BJAB	0.9 $\pm$ 0.0	1.1 $\pm$ 0.1	0.6 $\pm$ 0.1	0.5 $\pm$ 0.1

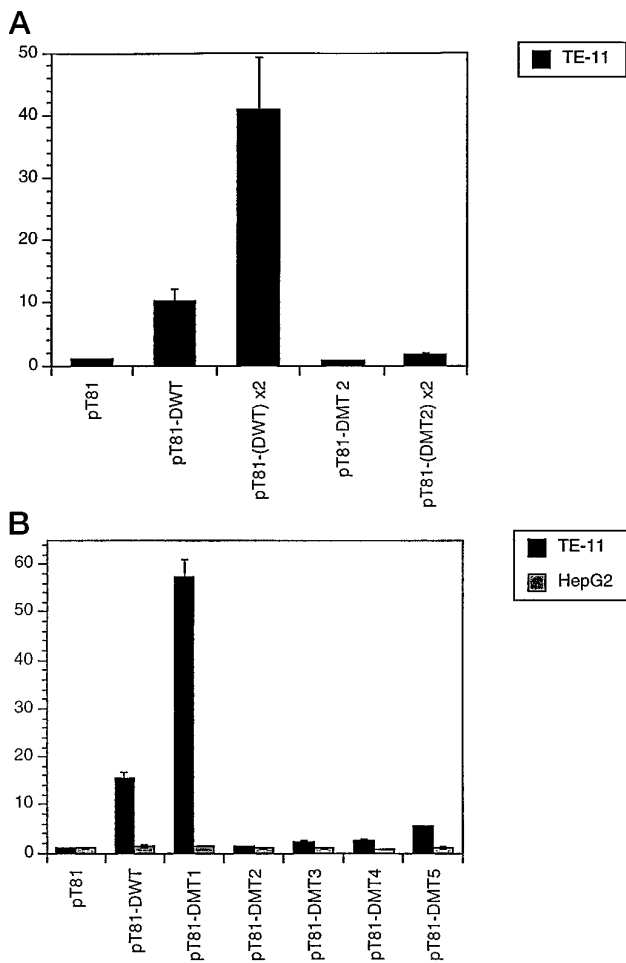


FIG. 2. Functional analysis of the CACCC-like motif in region D of the ED-L2 promoter. The wild-type DWT and mutated DMT sense oligonucleotides (see Table II for sequences) were subcloned in the enhancerless herpes simplex thymidine kinase vector, pT81, and transfected into TE-11 cells. Luciferase activity was assayed after 48 h. Activity is expressed relative to basal pT81-luciferase activity and represents the mean  $\pm$  S.D. of a minimum of four transfections done in parallel. Each transfection experiment was independently repeated at least three times. *A*, transcriptional activation by monomeric and dimeric concatamers of the wild-type and mutated sequences containing the CACCC-like element was determined in TE-11 cells. *B*, the constructs were transfected into TE-11 cells and HepG2 cells to determine cell type-specific activity of the CACCC-like element. *Solid box*, TE-11; *shaded box*, HepG2.

sequence of the CACCC-like motif (pT81-DMT3) also abolished transcriptional activity, suggesting that the intact 5'-flanking sequence is necessary but not sufficient for this activity. Inter-

estingly, a construct pT81-DMT1 in which a block mutation of 3'-flanking sequence of the CACCC-like motif was created had greater activity comparing with the wild-type construct, perhaps indicating the existence of an adjacent negative *cis*-regulatory element. The CACCC-like motif and its 5'-flanking sequence of region D is nevertheless likely to be of central importance as a cell-specific positive *cis*-regulatory element, especially since the mutant pT81-DMT1 construct did not have stronger transcriptional activity than the wild-type pT81-DWT construct in HepG2 cells.

Although mutations were not created in the KER-1 like motif, it seemed likely not to play a critical role since mutations in the sequence 5' to the KER1-like motif were sufficient to inactivate the transcriptional activity of region D.

*cis-Acting Regulatory Elements between Nucleotide Positions -218 and -187 Interact with Multiple Nuclear trans-Acting Factors*—To explore DNA-protein interactions between nuclear trans-acting factors and *cis*-regulatory elements that reside between nucleotide positions -218 and -187, GMSAs were carried out with a  $^{32}$ P-radioactively labeled double-stranded oligonucleotide, designated DWT. When crude nuclear extracts from TE-11 cells were examined, three distinct complexes, designated I, II, and III (Fig. 3, lane 1), were detected. The sequence specificity of these binding activities was further determined by competition experiments with 100-fold molar excess of unlabeled oligonucleotides. As depicted in Fig. 3, complexes I, II, and III were competed away by unlabeled DWT (lane 2) but not by other unlabeled oligonucleotides containing unrelated consensus motifs such as EGR (lane 10), AP1, AP3, GRE, CREB, NF $\kappa$ B, and Oct-1 (data not shown), strongly suggesting that complexes I, II, and III are sequence-specific complexes.

Interestingly enough, despite the absence of an Sp1 consensus motif in region D, complexes I and II but not complex III were competed away by the Sp1 oligonucleotide (Fig. 3, lane 9, and Fig. 6), suggesting that complexes I and II might comprise Sp1 or Sp1-related transcriptional factors and an atypical Sp1 site in region D. Whereas region D contained a KER1-like sequence as described previously, a competitor oligonucleotide containing an AP2 consensus sequence did not compete away complexes I, II, or III (Fig. 3, lane 11, and Fig. 6). Indeed, the KER1 motif is known to interact with the AP2 transcriptional factor in the keratin 14 promoter (18), and this was not observed within region D of the EBV ED-L2 promoter.

Although the intensity of the signals was very faint even after prolonged x-ray film exposure, there were three other specific complexes in TE-11 nuclear extracts, two of which migrated above complex III and one migrated between complex III and free probe (Fig. 3). The weak intensity of the signals did not lead to further investigation of these particular complexes. Of note, there was a nonspecific complex that migrated be-

TABLE IV  
The CACCC box, GT-I motif, and homologous elements and interacting transcription factors

Sequences identified as essential for DNA-protein interactions were underlined. Note these *cis*-acting elements often play a role in cell type-specific gene expression.

<i>cis</i> -Acting element	Sequence	Gene	<i>trans</i> -Acting factor	Ref.
CACCC-box	5'- <u>GCCACACCCCT</u> -3'	$\beta$ -Globin	EKLF, Sp1	20, 21, 36
GT-I motif	5'-ACTTTCACACCCCT-3'	SV40 enhancer	TEF-II	23, 24
GT box	5'-CCTTGCCACACCCCT-3'	Uteroglobin	Sp3, Sp4	33, 39
GT box	5'-CTCCACCCACC-3'	T-cell receptor	Sp2, Sp3	34
Sp1	5'-CGGGCGGGGC-3'	Ubiquitous/SV40 enhancer	Sp1	26

TABLE V  
*cis*-acting regulatory elements and transcriptional factors implicated in regulation of gene expression in keratinocytes

<i>cis</i> -Acting element	Sequence	Gene	<i>trans</i> -Acting factor	Ref.
AP2	5'-GCCNNNGGC-3'	Ubiquitous/SV40 enhancer	AP2	40
KER1	5'-GCCTGCAGGC-3'	Cytokeratins K5, K14, K3	AP2	18, 19
CK 8-mer	5'-AANCCAAA-3'	Cytokeratins, involucrin	Not identified	38

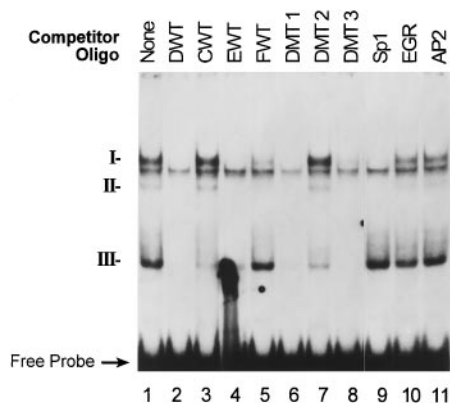


FIG. 3. Gel mobility shift assay of nuclear extracts from TE-11 cells and region D sequence of the ED-L2 promoter. Reactions containing 5  $\mu$ g of a crude nuclear extract and 5 fmol of a  $^{32}$ P-labeled wild-type oligonucleotide probe DWT were performed with or without 100-fold molar excess of unlabeled competitor oligonucleotides, including truncated or mutated sequences of DWT (see Table III for sequences) and those corresponding to Sp1, 5'-ATTTCGATCGGGCGGGGCGC-AGC-3'; EGR, 5'-CGCCTCGCCCCGCGCCGG-3'; and AP2, 5'-GATCGAAGTACCGCCCCGCGCCCGT-3'.

tween complexes I and II (Fig. 3).

**Complexes I and II, but Not Complex III, Interact with a CACCC-like cis-Regulatory Element**—To further delineate the nucleotide sequences comprising complexes I and II, competition experiments were done employing truncated and mutated sequences of region D as competitor oligonucleotides (Table II). As depicted in Fig. 3, a competitor oligonucleotide, designated CWT (lane 3), did not compete complexes I and II. However, oligonucleotide EWT competed away complexes I and II (lane 4), thereby indicating the sequence within oligonucleotide E could bind complexes I and II. A shorter competitor oligonucleotide FWT competed away both complexes I and II but with less efficiency than EWT (lane 5). However, under similar electrophoretic mobility shift assay conditions, the  $^{32}$ P-radioactively labeled FWT detected two distinct complexes which upon competition experiments showed a pattern similar to that of complexes I and II with probe D (data not shown). This observation suggests that a sequence within FWT is sufficient for the formation of complexes I and II, but flanking nucleotide sequences are likely necessary.

Furthermore, complexes I and II were competed by the oligonucleotide DMT1 as efficiently as DWT, and almost by DMT3, but not by DMT2 (lanes 6–8), suggesting that the altered sequence in DMT2, namely the CACCC-like element, is most essential for formation of complexes I and II and that the

altered 5'-flanking sequence in DMT3 is also necessary for binding. These results obtained with mutated sequences of region D are compatible with those of competitor oligonucleotides EWT and FWT and those with  $^{32}$ P-radioactively labeled WTF.

Since the competitor oligonucleotide FWT did not interfere with complex III while both competitor oligonucleotides EWT and CWT almost abolished complex III (Fig. 3, lanes 3–5), we deduced that complex III binding is with a sequence overlapping between CWT and EWT. However, the disappearance of complex III with competitor oligonucleotide DMT1 suggested that complex III may potentially require less stringent sequence specificity for its formation. Although the binding site for complex III is not as clear as that for complexes I and II, it is unlikely that complex III shares the binding motif with complexes I and II since the FWT and Sp1 (Fig. 3, lane 9) competitor oligonucleotides competed away both complexes I and II but not complex III. The weak binding activities (different from complexes I, II, and III) to region D were not clearly mapped. However, they showed a pattern similar to complex III with the same competitor oligonucleotides (Fig. 3 and data not shown).

To further delineate the critical nucleotides necessary for binding of complexes I and II, competition experiments were done with the FWT competitor oligonucleotide and its serially mutated sequences. As shown in Fig. 4, mutated competitor oligonucleotides with single nucleotide substitutions at the four cytosine bases in the CACCC-like motif (designated FMT 4, 6, 8, and 10) did not interfere with formation of complexes I and II (lanes 4, 6, 10, and 12). Other competitors without substitutions at the cytosine bases also partially competed away complexes I and II compared with FWT (lane 1). In particular, complexes I and II were entirely abolished by a competitor oligonucleotide FMT 11 whose single nucleotide substitution had previously been reported to affect binding of nuclear proteins to the CACCC-like motif in the  $\beta$ -globin promoter (20). This is consistent with the notion that the CACCC-like element in the EBV ED-L2 promoter region D may bind the same factors that interact with the true CACCC box in other promoters. Overall, these competition experiments indicate that each of the four cytosine bases within the CACCC-like element plays an essential role in the formation of the complexes I and II, but simultaneously 5'-flanking nucleotides are contributory. We conclude that the essential sequence in region D is 5'-CCAAGCCACACCTAA-3' (–218 to –204 bp). Finally, no differences in complexes I and II patterns were observed with the same panel of competitor oligonucleotides, suggesting that both complexes bind the same DNA motif.

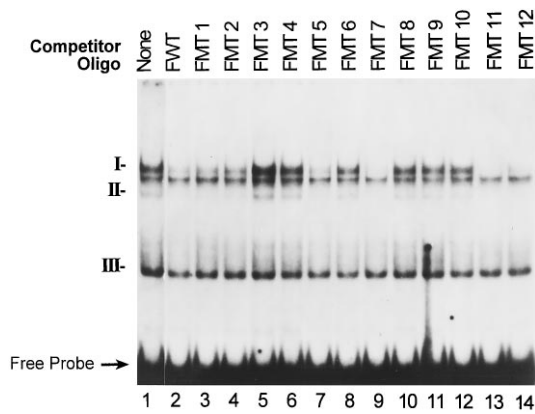


FIG. 4. Gel mobility shift assay with mutational analysis of nucleotides essential for interaction between the CACCC-like element and transcriptional factors in TE-11 nuclear extracts by competition assays. Wild-type oligonucleotide F (FWT) and its serially mutated sequences (FMT 1–12) (see Table III for sequences) were used as unlabeled competitor oligonucleotides. A nuclear extract was incubated with  $^{32}$ P-labeled probe DWT in the presence (lanes 2–14) or absence (lane 1) of 100-fold molar excess of unlabeled competitor oligonucleotides. Note that competitor oligonucleotides altered at any of the cytosine bases *underlined* in oligonucleotide F, 5'-AGCCACACCTAA-3' (FMT 3, 4, 8, 9, and 10), failed to compete away complexes I and II.

#### Complex I Comprises a Keratinocyte-specific Factor (KSF)—

Since the presence of a positive *cis*-regulatory element in region D and potentially cell type-specific interacting *trans*-acting factors in TE-11 cells were determined by transfection and GMSA experiments, further GMSA experiments were performed to elucidate whether there was indeed cell type specificity in the binding activities with probe D.

Fig. 5 shows that the complex I detected in TE-11 nuclear extracts (Fig. 5, lane 1) also exists in nuclear extracts from TE-12, T.T, and SCC-13 cells (Fig. 5, lanes 2–4), although the intensity of the signal is variable. In addition, complex I was barely detectable in SCC-25 nuclear extracts (Fig. 5, lane 5). However, complex I was undetectable in nuclear extracts from cell lines of selected nonsquamous epithelial origin, namely HepG2 (liver) and Panc-1 (pancreas) or a B cell line, BJAB (lanes 8–10). It was also undetectable in HeLa nuclear extracts (Fig. 6, lane 6). Instead, a prominent binding activity, designated complex IA (Fig. 5, lanes 6 and 8–10), was evident in HepG2, Panc-1, BJAB, and HeLa nuclear extracts and was undetectable in the esophageal, tongue, and skin squamous cell carcinoma cell nuclear extracts. Another complex, designated IB, was detectable in HepG2 and Panc-1 nuclear extracts (Fig. 5, lanes 8 and 9) and had a distinct mobility from the complex I or IA (Fig. 5, lane 7) as further confirmed by prolonged electrophoretic separation (data not shown). These data suggest that complex I comprises nuclear transcriptional factor(s) specifically expressed or activated in cells of stratified squamous epithelial origin, alternatively called keratinocytes. As a result, we have designated complex I as keratinocyte-specific factor (KSF). In contrast, while the intensity of the signals corresponding to complexes II and III varied in the electrophoretic mobility shift assays, they were not keratinocyte-specific based upon their detection in BJAB nuclear extracts. In summary, these electrophoretic mobility shift assays indicated the 5'-CCAAGCCACACCTAA-3' *cis*-regulatory element(s) in region D of the EBV ED-L2 promoter interacts with a cell type-specific factor, KSF, and also with ubiquitous factors.

**Complex IA Comprises Sp1 in Nuclear Extracts from Nonkeratinocytes**—To further characterize nuclear *trans*-acting factors interacting with the CACCC-like element to form complexes I and II, immune supershift assays were performed with

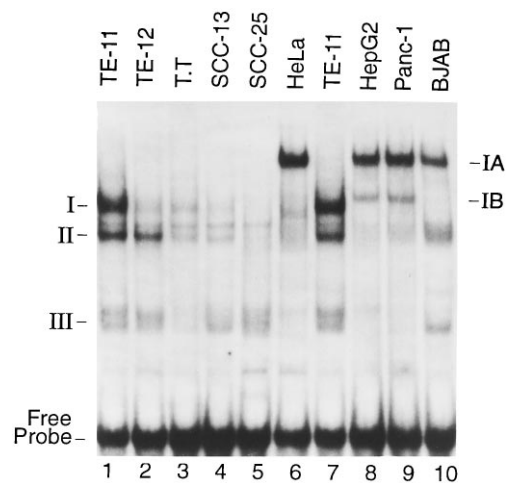


FIG. 5. Gel mobility shift assay showing cell type-specific interaction between region D of the ED-L2 promoter and nuclear extracts from various cell lines. The  $^{32}$ P-radioactively labeled oligonucleotide probe DWT was incubated with 5  $\mu$ g each of nuclear extracts prepared from cells of squamous epithelial origin or keratinocytes, TE-11, TE-12, and T.T (esophagus), SCC-13 (skin), SCC-25 (tongue); other epithelial cells, HeLa (cervix), HepG2 (liver), and Panc-1 (pancreas); and B lymphocytes, BJAB. A TE-11 nuclear extract used in lane 7 was an identical preparation to the one in lane 1.

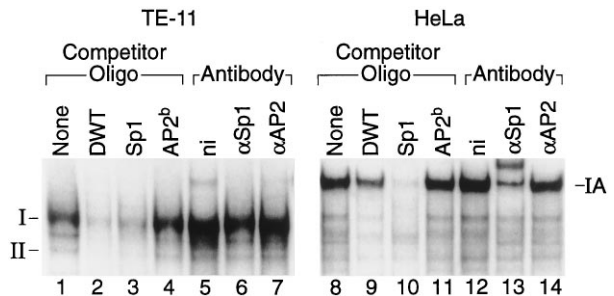


FIG. 6. Gel mobility shift assay with competition and immune supershift reactions were performed to characterize complexes between the  $^{32}$ P-radioactively labeled oligonucleotide probe DWT and nuclear extracts to determine immunological cross-reactivity between the complexes and known transcriptional factors. 5  $\mu$ g of nuclear extract from TE-11 or HeLa cells was incubated with 1.0  $\mu$ l of a monoclonal AP2 antibody, 1.0  $\mu$ l of a rabbit polyclonal Sp1 antisera, or 1.0  $\mu$ l of control rabbit nonimmune sera (*ni*) at room temperature for 10 min prior to addition of the  $^{32}$ P-radioactively labeled oligonucleotide DWT probe.

available antibodies against factors known to interact with the CACCC element. Initially, an immune supershift assay was done with an anti-Sp1 antibody. Fig. 6 shows that complexes I and II detected in TE-11 cells were competed with an oligonucleotide containing the Sp1 consensus motif (Fig. 6, lane 2) but were not interfered with by anti-Sp1 antibody (Fig. 6, lane 6). This would indicate that the factors composing complexes I and II are immunologically distinct from Sp1. In contrast, complex IA detected in a HeLa nuclear extract was also clearly competed away by the Sp1 consensus oligonucleotide (Fig. 6, lane 8) as well as an unlabeled wild-type competitor oligonucleotide DWT (not shown) but not with the AP2 oligonucleotide (Fig. 6, lane 9). Furthermore, this complex was supershifted by anti-Sp1 antibody (Fig. 6, lane 11) indicating that complex IA comprises Sp1. Although the data are not shown, anti-Sp3 antibody did not interfere with complexes I or II on the premise that these complexes may be related to Sp1 (antibodies to Sp2 and Sp4 were unavailable for testing). These data clearly indicate that complex IA, observed in cells of nonsquamous epithelial origin or B lymphocytes (Fig. 5), is Sp1 and most likely binds the CACCC-like motif in region D of the EBV ED-L2 promoter,

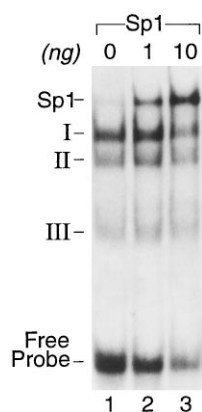


FIG. 7. Gel mobility shift assay showing the interaction of Sp1 and region D of the ED-L2 promoter. The indicated amount of purified human recombinant Sp1 protein was added to the reaction containing 5  $\mu$ g of a TE-11 crude nuclear extract and  $^{32}$ P-radioactively labeled oligonucleotide probe DWT.

whereas complex I (KSF) recognizes the same motif but is distinct from Sp1 and Sp3. Further confirmation of the interaction of Sp1 and region D was done by incubating radiolabeled oligonucleotide probe D with increasing concentrations of purified human recombinant Sp1 protein and maintaining a constant concentration of TE-11 nuclear extract. As shown in Fig. 7, a complex migrating slower than complex I (KSF) appeared, depending upon the amount of the Sp1 protein added in the mixing experiment. Furthermore, purified human recombinant Sp1 protein competed away complex I (Fig. 7, lane 3), suggesting that KSF and Sp1 may interact with the same CACCC-like cis-regulatory element but in a competitive fashion.

Although the data are not shown, the Sp1 binding site in region D was mapped by competition experiments with the same panel of oligonucleotides as ones used to determine the binding site for complexes I and II. As expected, competitor oligonucleotides DWT and EWT, but not CWT, competed complex IA in a HeLa nuclear extract, indicating that Sp1 binds to the sequence contained in the 5'-half of region D. Furthermore, complex IA was competed away by a competitor oligonucleotide DMT1 carrying a block mutation at the 3'-flanking sequence of the CACCC-like motif. Consistent with the identification of the CACCC box as a minimal binding sequence in the  $\beta$ -globin promoter (20, 25), we demonstrated that complex IA failed to be competed by oligonucleotide DMT2 (data not shown). The other block mutation in the 5'-flanking sequence of the CACCC-like element in region D (DMT3) did not compete away complex IA (data not shown) which is different from the results of competition experiments with complexes I and II. Furthermore, complex IA was not interfered with by competitor oligonucleotide WTF nor by any one of its mutated series. Although 5'-AGC-CACACCTAA-3' was sufficient for binding of complexes I and II, these data suggest that Sp1 also requires a sequence extending over the minimal CACCC-like element into 5'-flanking nucleotides.

**KSF Consists of a Protein with Molecular Mass of Approximately 65 kDa**—As an approach to further characterizing the transcriptional factors comprising complexes I, II, and III, an UV cross-linking experiment was done to determine their molecular masses. As shown in Fig. 8, the isolated complex I was resolved on SDS-polyacrylamide gel electrophoresis into a component with apparent molecular masses of approximately 65–70 kDa (Fig. 8, lane I). Furthermore, complex II appears to consist of a 40–45 kDa protein (Fig. 8, lane II), although this band has a faint signal. Finally, consistent with the suggestion on DNA gel shift that complex III comprises several transcriptional factors, the UV cross-linking data confirm this observa-

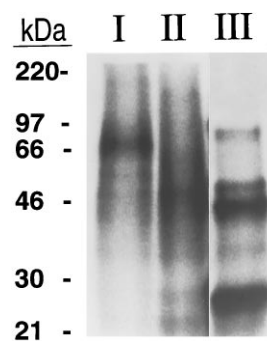


FIG. 8. Denaturing SDS-polyacrylamide gel analysis of UV cross-linked complexes formed with TE-11 nuclear extract proteins and the  $^{32}$ P-radioactively labeled oligonucleotide probe DWT. The components corresponding to complexes I, II, and III are shown in lanes I, II, and III, respectively. The positions of the molecular size markers are indicated.

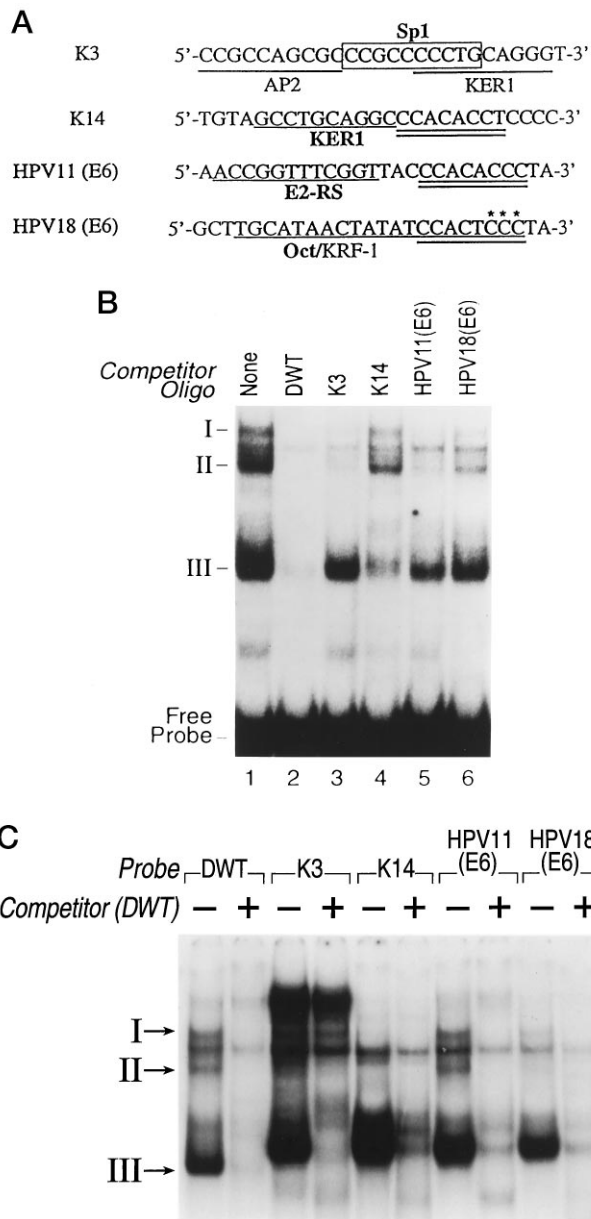
tion (Fig. 8, lane III).

**KSF May Interact with Other Viral and Eukaryotic cis-Regulatory Elements Located in Promoters Regulating Keratinocyte-specific Gene Expression**—The transcriptional regulation of gene expression in keratinocytes is marked by the identification of some cis-regulatory elements in eukaryotic and viral promoters (19, 22, 26–29). Interestingly, some of these previously identified cis-regulatory elements consist of or are flanked by a GC-rich sequence identified as an Sp1 site as well as the CACCC core motif of GT-I in the SV40 enhancer and its homologous sequences (19, 22, 26–29) (Table IV and Fig. 9A). This prompted us to hypothesize that KSF may participate in the transcriptional regulation of other viral and eukaryotic promoters active in keratinocyte-specific gene expression.

To test this idea, a competition experiment was performed with competitor oligonucleotides containing the sequences of cis-regulatory elements in the cytokerin K3 and K14 promoters and the human papillomavirus-11 and -18 E6 promoter (Fig. 9B and Table V). Fig. 9B shows that all of these competitors interfered to varying degrees with the formation of complex I (Fig. 9B, lanes 3–6), thereby indicating that KSF might interact with the CACCC-like element in other promoters. Of note, the K14 oligonucleotide abolished the binding of complex III (Fig. 9B, lane 4).

To further test the notion that complex I might bind related CACCC elements in other promoters, a cross-competition experiment was performed in which TE-11 nuclear extracts were incubated with  $^{32}$ P-radioactively labeled sequences in the K3, K14, HPV-11 E6, and HPV-18 E6 promoters (Fig. 9C). This reveals that the putative complex I appears to bind radioactively labeled K3 and HPV-11 E6 sequences and possibly HPV-18 E6 (very faint). It should be emphasized that the inference of DNA binding is based upon migration patterns only. Apart from this consideration for complex I, it is possible that the putative complex III pattern with these promoters is due to migration differences, or alternatively, there is no binding to any complex III. Competition with unlabeled wild-type D oligonucleotide reveals that it competes away the binding of the putative complex I with HPV-11 E6 and only partially with K3 (Fig. 9C).

Having established the UV cross-linking pattern of complex I with probe D as an approximately 65–70-kDa protein, we performed UV cross-linking experiments with the putative complex I binding to K3 and HPV-11 E6 sequences. This revealed that the molecular mass of each putative complex I is in the same range of 65–70 kDa (Fig. 10), suggesting that this complex may indeed be KSF. However, the DNA gel shifts and the UV cross-linking data remain suggestive, and actual clon-

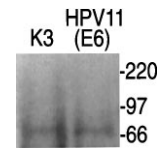


**FIG. 9. Gel mobility shift assay showing the interaction of complex I (keratinocyte-specific factor) and cis-regulatory elements identified in other eukaryotic and viral promoters active in keratinocytes.** A, sequences containing cis-acting regulatory elements identified in cytokeratin K3 and K14 promoters and in human papilloma virus 11 and 18 E6 promoters were tested as unlabeled competitor oligonucleotides. The identified DNA motifs are *underlined* or *boxed*. CACCC (GT-I) or its homologous motif is *underlined twice*. DNA motifs known to interact with identified transcriptional factors are in *bold*. Important nucleotides within DNA motifs as determined by methylation interference assays are marked with *asterisks*. B, the competition assay was performed by incubating 5  $\mu$ g of TE-11 nuclear extract and the  $^{32}$ P-radioactively labeled oligonucleotide probe DWT in the presence or absence of 100-fold molar excess of unlabeled competitor oligonucleotides. C, oligonucleotides from DWT, K3, K14, HPV-11 (E6), and HPV-18 (E6) were  $^{32}$ P-radioactively labeled and incubated with TE-11 nuclear extracts. Competition was performed in the presence or absence of 100-fold molar excess of unlabeled DWT oligonucleotide.

ing of the gene for KSF would be definitive proof. Furthermore, KSF may have different affinities for related cis-regulatory elements in promoters that are active in keratinocytes as reflected in the DNA gel shift experiments.

#### DISCUSSION

Insights into transcriptional regulation of gene expression in squamous epithelia, or keratinocytes, have been gained



**FIG. 10. Denaturing SDS-polyacrylamide gel analysis of UV cross-linked complexes formed with TE-11 nuclear extract proteins and the  $^{32}$ P-radioactively labeled oligonucleotide probes K3 and HPV-11 (E6).** The positions of the molecular size markers are indicated.

through analysis of the eukaryotic cytokeratin promoters. Complementary information has been acquired through investigation of DNA viruses that can infect and replicate in squamous epithelia. Foremost among these viruses are human papillomavirus and Epstein-Barr virus.

The Epstein-Barr virus infects oropharyngeal squamous epithelial cells prior to uptake in B lymphocytes. EBV is associated with epithelial conditions such as oral leukoplakia, nasopharyngeal carcinoma, and esophageal squamous cell carcinoma (30). Some insights into EBV's role in squamous epithelial cells are starting to emerge from transgenic mice studies. The EBV ED-L2 promoter has been used to target cyclin D1 in mice with resultant tissue-specific expression in the tongue, esophagus, forestomach, and skin (9). These findings suggest that cellular transcriptional factors may interact with the EBV ED-L2 promoter in a relatively tissue-specific fashion. In a broader context, functional analysis of this promoter may provide insights if such factors might interact with other viral and eukaryotic promoters in squamous epithelial cells.

*Keratinocyte-specific EBV ED-L2 Promoter Activity in Human Squamous Epithelial Carcinoma Cell Lines Correlates with Transgene Expression in Mice*—Our transfection experiments with a series of EBV ED-L2 promoter deletion constructs suggested that the majority of promoter activity resided between nucleotide positions -218 and -187, a fragment designated as region D. This activity was evident exclusively in cell lines of squamous epithelial origin but not in nonsquamous epithelial origin or in B lymphocytes. Furthermore, the same DNA fragment extending from nucleotides -218 and -187 functioned as an enhancer in a heterologous minimal herpes simplex virus thymidine kinase promoter-luciferase reporter gene construct in a cell type-specific fashion. These transfection data strongly imply that this region contains cis-acting regulatory element(s) that interact with nuclear transcriptional factors to allow gene expression in a squamous epithelial cell or keratinocyte-specific fashion.

In our ED-L2-cyclin D1 transgenic mice, the transgene is expressed only in basal and suprabasal cells but not in terminally differentiated cells in the stratified squamous epithelium of the tongue, esophagus, forestomach, and skin (9). Since cytokeratins K14 and K5 are expressed in actively dividing basal cells in all stratified epithelia (27), and K1 and K10 cytokeratins are expressed in terminally differentiated cells, it is tempting to speculate that ED-L2 promoter activity may be dependent as well on the degree of differentiation in the squamous epithelial cells which in fact is suggested by our cell line data. The undetectable ED-L2 promoter activity in cell lines of nonsquamous epithelial origin paralleled findings in transgenic mice (8, 9). Curiously enough, the ED-L2 promoter was inactive in an EBV-negative B cell line, BJAB. This observation suggests that B lymphocytes may require EBV gene products to up-regulate the ED-L2 promoter, as is the case in other EBV promoters (31). It is also possible that certain signaling pathways may need to be induced in B cells to up-regulate the ED-L2 promoter. In fact, a 0.6-kilobase transcript correspond-

ing to BNL2 (Fig. 1A) is induced in EBV-infected B lymphocytes upon 12-*O*-tetradecanoylphorbol-13-acetate treatment which experimentally allows latently infected lymphocytes to enter the lytic cycle (7). Perhaps, in this context, the ED-L2 promoter is regulated by a protein kinase C-mediated signal transduction pathway in B cells. Of potential parallel importance, protein kinase C activation also plays a critical role in keratinocyte differentiation through inhibition of cytokeratins 1 and 10 (32).

**ED-L2 Promoter Activity Is Associated with a CACCC-like Element in a Cell Type-specific Fashion**—The transfection experiments in TE-11 cells functionally mapped nucleotides responsible for the keratinocyte-specific transcriptional activity of the ED-L2 promoter to a CACCC-like motif and its 5'-flanking nucleotides of region D, namely 5'-CAAGCCACACCTAA-3'. GMSAs with a <sup>32</sup>P-labeled probe D revealed multiple binding activities, including the keratinocyte-specific complex I, designated keratinocyte-specific factor (KSF), as well as complexes II and III. Upon competition experiments, single nucleotide substitutions in mutant competitor oligonucleotides, FMT 4 and FMT 10 (Fig. 4A), did not interfere with the formation of complexes I and II. This observation correlated with transfections of heterologous minimal promoter-luciferase reporter gene constructs in which pT81-DMT4 and pT81-DMT5, containing the corresponding point mutations (Fig. 2B) of FMT 4 and FMT 10, inhibit the transactivation of region D. This clearly implies the interaction between the CACCC-like cis-regulatory element and trans-acting factors contributes significantly to EBV ED-L2 transcriptional activity.

While the result with pT81-DMT3 (Fig. 2B) indicated the functional importance of the 5'-flanking sequence of the CACCC-like element in the ED-L2 promoter, the GMSAs further corroborated this notion. In particular, the EWT competitor oligonucleotide more efficiently competed complexes I and II than FWT or DMT3 (Fig. 3), suggesting 5'-flanking nucleotides are necessary. The importance of the flanking sequence of the CACCC-like element is further underscored by the absence of a diminution in transcriptional activity between pL2-782 and pL2-610, despite the presence of an identical 5'-CCA-CACCT-3' element at nucleotide position -766. Furthermore, there is only basal activity with pL2-114, despite an identical 5'-CCCACACCT-3' element at -99 of the EBV ED-L2 promoter.

As a separate consideration, although the transfection data with construct pT81-DMT1 suggested the presence of a negative cis-regulatory element, there was no detectable complex that was competed specifically with DWT but not with DMT1 (Fig. 3). Thus, we believe that the CACCC-like cis-regulatory element acts positively and represents the main cis-regulatory element in region D.

The role of complex III in the heterologous promoter system remains unclear. However, the abrogation of activity in the pT81-DMT2 transfection (Fig. 2B) suggests that complexes I and II are more critical for promoter function than complex III since the GMSA competition experiment showed DMT2 competed complex III but not complexes I and II (Fig. 3A, lane 7).

Although the effect of the mutated KER1-like motif in region D was not formally tested in the minimal promoter-luciferase reporter gene constructs, it is nonetheless unlikely to play an intrinsic role in contributing to promoter activity since block mutation of the CACCC-like motif was sufficient to inactivate the transcriptional activity of region D. In addition, neither competition nor immune supershift experiments in the GMSAs indicated involvement of AP2 in the formation of complexes I and II (Fig. 6, lanes 3 and 6) or III (data not shown), supporting the notion that complex I is distinct from AP2 which is an

important transcriptional factor involved in keratinocyte-specific gene expression.

**KSF Interacts with a CACCC-like Element**—The GMSAs demonstrated that one of the binding activities interacting with the CACCC box-like element, KSF, was detectable only in squamous cell carcinoma cell lines retaining features of keratinocytes. The KSF binding activity correlated closely with functional data from the transfection studies.

It is noteworthy that the CACCC box was originally identified as an evolutionally conserved cis-acting regulatory element in  $\beta$ -globin promoters (25) and that it interacts with the erythroid Krüppel-like factor which functions as a transactivator in a cell type-specific fashion (21) and also interacts with Sp1. Furthermore, the CACCC box element has been identified as part of the GT-I motif in the SV40 viral enhancer (24) and subsequently in some eukaryotic keratin promoters. Interestingly, some of them are also identified as cis-regulatory elements responsible for cell type-specific gene expression (see Tables IV and V). Thus far, three transcriptional factors have been cloned and shown to interact with the GT-1 or GT box elements. All are highly homologous to Sp1 at their zinc finger DNA binding domains and, therefore, have been designated Sp2, Sp3, and Sp4 (33, 34). In this context, KSF may be related to these factors based upon our competition assays with the Sp1 consensus oligonucleotide and recombinant Sp1 protein. However, the molecular mass of KSF, estimated to be 65–70 kDa by UV cross-linking, suggests that it is distinct from Sp1, Sp2, Sp3, and Sp4, the molecular masses of which range from approximately 80 to 100 kDa (33, 34). In addition, the tissue or cell distribution of the Sp1 protein family is not consistent with that of KSF, at least as implied by the GMSAs.

Sp1 binding activity was barely detectable in nuclear extracts from squamous cell carcinoma cell lines in contrast to nuclear extracts from cells of nonsquamous epithelial origin or B lymphocytic origin (Fig. 5). This finding is compatible with the low expression levels of Sp1 in squamous epithelia as noted in immunohistochemistry studies (35). The lack of transcriptional activity of region D in HepG2 cells (Fig. 2C), despite abundant Sp1 activity in HepG2 nuclear extracts with probe D (Fig. 5), clearly indicates differences in function between Sp1 and KSF for the CACCC-like cis-regulatory element. While both Sp1 and KSF potentially may share an identical DNA binding motif, differences in function may be attributable to variations in transactivation domains as is the case with erythroid Krüppel-like factor and Sp1.

TEF-II is another cell-specific transcriptional factor that has been shown to interact with the SV40 GT-I element (23). While the molecular mass of TEF-II is similar to that of KSF, we doubt that these two factors are identical since TEF-II was originally purified from HeLa cells, and under our experimental conditions, KSF binding activity was not present in HeLa nuclear extracts.

**Possible Interaction of KSF and the CACCC Box-like Element in the Regulation of Genes Essential in Keratinocytes**—The GMSA data indicate that CACCC box elements in other viral and eukaryotic promoters may interact with complex I (KSF). This suggests that KSF may interact with promoters aside from the EBV ED-L2 promoter, in particular those that are important in regulating gene expression in keratinocytes. It is useful to dissect such DNA motifs (Tables IV and V) in viral and eukaryotic promoters and to analyze them in relation to the motif that binds KSF.

The AP2/Sp1/KER1 oligonucleotide (also referred to as K3 in our GMSAs) contains a keratinocyte-specific cis-regulatory element identified in the rabbit cytokeratin K3 promoter (28, 29), whose Sp1 motif is flanked by AP2 and KER1 sites (Fig. 9A). It

was demonstrated to be functionally important and specifically bound by a transcriptional factor most likely to be Sp1 in rabbit corneal keratinocytes. Our data indicate that KSF interacts with this GC-rich atypical Sp1 motif (Fig. 9, *B* and *C*) as well as the consensus Sp1 motif (Figs. 3 and 6).

The KER1 oligonucleotide (also referred to as K14 in our GMSAs) contains the KER1 element, identified as important in keratinocyte-specific transcriptional regulation of several cytokeratin genes, including cytokeratins K14 and K5 which are expressed in the basal cell layer of the squamous epithelium (18, 19). Significantly, this KER1 motif is flanked by a sequence, 5'-CCACACCT-3', identical to the CCACACCT element in region D of the ED-L2 promoter (Fig. 9A). Less efficient competition of this oligonucleotide for complex I (Fig. 9B, *lane* 4) may be explained by the difference of the 5'-flanking sequence of the ED-L2 CACCC box-like element with that of the 5' KER1 motif. Furthermore, there does not appear to be any binding of complex I (KSF) with a radioactively labeled K14 oligonucleotide probe (Fig. 9C). It is conceivable that binding of the AP2 transcriptional factor to the KER1 motif in the KER1 competitor oligonucleotide may have interfered with the interaction between KSF and the nucleotides flanking the CACCC box-like element. This may help to explain why Leask *et al.* (18, 19) did not detect a binding activity corresponding to KSF under their experimental conditions with the KER1 oligonucleotide and SCC-13 nuclear extracts. Although it is not the case with the EBV ED-L2 promoter, it is worth emphasizing that AP2 (or KER1) and Sp1 (or CACCC) motifs are often located adjacent to each other, as exemplified in cytokeratins K3 (28, 29), K14 (18, 19) (Fig. 9A), and K5 (27), thereby potentially allowing for more sophisticated transcriptional regulation of these genes in keratinocytes.

The E2-RS/GT-I oligonucleotide (also referred to as HPV-11 E6 in our GMSAs), derived from a *cis*-acting regulatory element in E6 promoter of human papilloma virus type 11, contains a perfect CACCC box and an adjacent E2 protein-responsive sequence. A cellular factor has been implicated in cervical carcinoma-derived simple epithelial-like cells, C-33A, to interact with the CACCC box that abrogates repression by the E2 protein, although it was suggested to be neither Sp1 nor related proteins (22).

The KRF-1 oligonucleotide (also referred to as HPV-18 E6) is derived from a constitutive enhancer in the E6 promoter of human papilloma virus type 18. It was shown to interact with a transcriptional factor expressed in HeLa cells and SCC-13 cells but not in BJAB cells (36). It was not shown whether this HPV sequence interacts with Sp1 in HeLa nuclear extracts (36). Although mapping of nucleotides essential for this interaction in the relatively long DNase I-protected sequence may not be enough to define a responsible motif, three cytosine bases were nonetheless implicated by a methylation interference assay (36). Since nucleotides at the 3'-end of this region include cytosine bases resembling the CACCC box, it may help to explain why complex I was competed away by the KRF-1 competitor oligonucleotide (Fig. 9B, *lane* 6), and also why there was only faint binding to the HPV-18 E6 oligonucleotide probe (Fig. 10).

Our data suggest that KSF may potentially play a role in the transcriptional regulation of genes in keratinocytes through interaction with CACCC and GT-I elements in viral and certain cytokeratin promoters, some of which remain to be identified since we only tested a limited number. Furthermore, the transactivation of these promoters by KSF may vary depending upon the affinity of DNA binding and parallel interaction(s) with other transcription factors. Although clearly there are many positive (and negative) *cis*-regulatory elements in these pro-

motors as well as other interacting transcriptional factors, the presence of a cell type-specific factor such as KSF may be very critical, for example in the switch from proliferation to early differentiation in the stratified squamous epithelium. Although we have emphasized the relative importance of the interaction between the CACCC-like element and KSF, this does not exclude the possible importance of other DNA-protein interactions such as those constituting complexes II and III, especially in the context of the EBV ED-L2 promoter. Furthermore, we found that the region between -187 and -164 also contributes to ED-L2 promoter activity, likely through the interaction of the helix-loop-helix protein USF with an E-box motif (data not shown). However, this latter interaction was not found to be cell type-specific.

The results from the UV cross-linking experiments indicate that complex I comprises a 65–70-kDa keratinocyte-specific factor. It is possible that KSF and complex II form a heterodimer in binding to the CACCC-like motif. The contribution of complex III is unclear presently. Our data indicate that KSF is not Sp1, Sp3, AP2, or  $\alpha$ -basonuclin (data not presented) (37), all of which have been implicated in keratinocyte gene expression. Efforts are ongoing to identify the gene encoding KSF.

In the present study, cellular transcriptional factors, including a cell type-specific one in keratinocytes, were implicated in contributing to the transactivation of the EBV ED-L2 promoter through the CACCC box-like element between -218 and -187 bp. This does not preclude the contribution to promoter activity by other *cis*-regulatory elements in the ED-L2 promoter. For example, the CK 8-mer motif, found in the involucrin and many cytokeratin promoters (38), occurs twice at -746 bp and at -230 bp in the antisense orientation of the ED-L2 promoter and the E-box motif between -187 and -164 bp of the ED-L2 promoter. Overall, our findings provide novel information about EBV's role in epithelial cells. In a broader context, the data presented have important implications in our understanding of how viral and eukaryotic genes are transcriptionally regulated in keratinocytes.

*Acknowledgments*—We are grateful to Daniel C. Chung, Timothy D. Jenkins, Timothy C. Wang, and Joanna Wilson for discussions.

#### REFERENCES

- Cohen, J. I. (1993) *Ann. Intern. Med.* **118**, 45–48
- Kieff, E. (1996) in *Virology* (Felds, B. N., Knipe, D. M., and Howley, P. M., eds) Chapter 74, pp. 2343–2396, Lippincott-Raven Press, Ltd., New York
- Wang, D., Liebowitz, D., and Kieff, E. (1985) *Cell* **43**, 831–840
- Dawson, C. W., Rickinson, A. B., and Young, L. S. (1990) *Nature* **344**, 777–780
- Fahraeus, R., Rymo, L., Rhim, J. S., and Klein, G. (1990) *Nature* **345**, 447–449
- Raab-Traub, N., Hood, R., Yang, C. S., Henry, B. I., and Pagano, J. S. (1983) *J. Virol.* **48**, 580–590
- Hudson, G. S., Farrell, P. J., and Barrell, B. G. (1985) *J. Virol.* **53**, 528–535
- Wilson, J. B., Weinberg, W., Johnson, R., Yupsa, S., and Levine, A. J. (1990) *Cell* **61**, 1315–1327
- Nakagawa, H., Wang, T., Zukerberg, L., Odze, R., Togawa, K., May, G., Wilson, J., and Rustgi, A. (1997) *Oncogene* **14**, 1185–1190
- Baer, R., Bankier, A. T., Biggin, M. D., Farrell, P. J., Gibson, T. J., Hatfull, G., Hudson, G. S., Satchwell, S. C., Seguin, C., Tuffnell, P. S., and Barrell, B. G. (1984) *Nature* **310**, 207–211
- Nordeen, S. K. (1988) *BioTechniques* **6**, 454–457
- Nakagawa, H., Zukerberg, L., Togawa, K., Meltzer, S. J., Nishihira, T., and Rustgi, A. K. (1995) *Cancer* **76**, 541–549
- Rheinwald, J. G., and Beckett, M. A. (1981) *Cancer Res.* **41**, 1657–1663
- Graham, F. L., and van der Eb, A. J. (1973) *Virology* **52**, 456
- Schreiber, E., Matthias, P., Müller, M. M., and Schaffner, W. (1989) *Nucleic Acids Res.* **17**, 6419
- Bradford, M. M. (1976) *Anal. Biochem.* **72**, 248–254
- Chung, D. C., Brand, S. J., and Tillotson, L. G. (1995) *J. Biol. Chem.* **270**, 8829–8836
- Leask, A., Byrne, C., and Fuchs, E. (1991) *Proc. Natl. Acad. Sci. U. S. A.* **88**, 7948–7952
- Leask, A., Rosenberg, M., Vassar, R., and Fuchs, E. (1990) *Genes Dev.* **4**, 1985–1998
- Cowie, A., and Myers, R. M. (1988) *Mol. Cell. Biol.* **8**, 3122–3128
- Miller, I. J., and Bieker, J. J. (1993) *Mol. Cell. Biol.* **13**, 2776–2786
- Dong, G., Broker, T. R., and Chow, L. T. (1994) *J. Virol.* **68**, 1115–1127
- Davidson, I., Xiao, J. H., Rosales, R., Staub, A., and Chambon, P. (1988) *Cell* **54**, 931–942
- Xiao, J.-H., Davidson, I., Macchi, M., Rosales, R., Vigneron, M., Staub, A., and

- Chambon, P. (1987) *Genes Dev.* **1**, 794–807
25. Myers, R. M., Tilly, K., and Maniatis, T. (1986) *Science* **232**, 613–618
26. Kadonaga, J. T., Carner, K. R., Masiarz, F. R., and Tjian, R. (1987) *Cell* **51**, 1079–1090
27. Byrne, C., and Fuchs, E. (1993) *Mol. Cell. Biol.* **13**, 3176–3190
28. Wu, R.-L., Chen, T.-T., and Sun, T.-T. (1994) *J. Biol. Chem.* **269**, 28450–28459
29. Wu, R.-L., Galvin, S., Wu, S.-K., Xu, C., Blumenberg, M., and Sun, T.-T. (1993) *J. Cell Sci.* **105**, 303–316
30. Jenkins, T. D., Nakagawa, H., and Rustgi, A. K. (1996) *Oncogene* **13**, 1809–1813
31. Urier, G., Buisson, M., Chambard, P., and Sergeant, A. (1989) *EMBO J.* **8**, 1447–1453
32. Dlugosz, A. A., Cheng, C., Williams, E. K., Dharia, A. G., Denning, M. F., and Yuspa, S. H. (1994) *Cancer Res.* **54**, 6413–6420
33. Hagen, G., Müller, S., Beato, M., and Suske, G. (1992) *Nucleic Acids Res.* **20**, 5519–5525
34. Kingsley, C., and Winoto, A. (1992) *Mol. Cell. Biol.* **12**, 4251–4261
35. Saffer, J. D., Jackson, S. P., and Annarella, M. B. (1991) *Mol. Cell. Biol.* **11**, 2189–2199
36. Mack, D. H., and Laimins, L. A. (1991) *Proc. Natl. Acad. Sci. U. S. A.* **88**, 9102–9106
37. Tseng, H., and Green, H. (1992) *Proc. Natl. Acad. Sci. U. S. A.* **89**, 10311–10315
38. Blessing, M., Zentgraf, H., and Jorcano, J. L. (1987) *EMBO J.* **6**, 567–575
39. Misseyanni, A., Klug, J., Suske, G., and Beato, M. (1991) *Nucleic Acids Res.* **19**, 2849–2859
40. Williams, T., and Tjian, R. (1991) *Genes Dev.* **5**, 670–682