Multiple interactions between cellular factors and the non-coding region of human papillomavirus type 16

Gary J. Sibbet and M. Saveria Campo*

Beatson Institute for Cancer Research, Garscube Estate, Switchback Road, Bearsden, Glasgow G61 1BD, UK.

The interaction between cellular factors and the non-coding region (NCR) of human papillomavirus type 16 (HPV-16) has been studied using different approaches: DNase I hypersensitive sites (HSS) analysis of HPV-16 chromatin in SiHa cells, footprinting (FP) of the NCR in vitro with nuclear extracts from SiHa cells and band shift analysis of synthetic oligonucleotides corresponding to the HSS and footprint sites with nuclear extracts from a variety of cells. These analyses have shown that cellular DNA-binding factors (CDBFs) bind to at least 13 sites, FP-A to M, over approximately 500 bp in the 3' half of the NCR. FP-E to H overlap with the keratinocyte-dependent enhancer and most sites contain the consensus sequence GCCAAA, which is similar to both the cytokeratin promoter element AARCCAAA and the ubiquitous CCAAT motif. FP-E appears to be the target for C/EBP and AP1 binding, but FP-F, G and H appear to bind novel specific cellular factors. The CDBF binding to FP-G has been tentatively identified as the same factor binding to the cytokeratin promoter element.

Introduction

The aetiology of cervical cancer strongly suggests the involvement of sexually transmissible agents. Recently, several specific types of human papillomavirus (HPV), particularly HPV-16, have been found to be associated with the great majority of genital cancers (zur Hausen, 1989), pointing to the involvement of the virus in these malignancies. Infection of the mucosa of the uterine cervix by HPV leads to flat lesions which can progress to cancer via cervical intraepithelial neoplasia and carcinoma in situ (Pfister, 1987). This process is often but not always associated with a configurational change of the HPV genome. The viral DNA, maintained as a high copy number episome in most benign lesions, becomes integrated into the host genome in cancers (Dürst et al., 1985); but see Schneider-Maunoury et al. (1987), for integration of viral DNA in early pre-malignant lesions, and Matsukura et al. (1989), for episomal forms of viral DNA in frank cancers. Integration disrupts the viral E2 open reading frame (ORF) which encodes transcriptional regulatory factors (Phelps & Howley, 1987), thereby allowing unregulated expression of the transforming ORFs E6 and E7 (Matlashewski et al., 1987).

The cell line SiHa, derived from a cervical squamous cell carcinoma, contains a single copy of integrated HPV-16 DNA deleted in the E2 ORF (El Awady et al., 1987; Baker et al., 1987), and is therefore incapable of transcriptional autoregulation; the viral DNA still maintains the non-coding region (NCR) which directs transcription of the intact E6 and E7 ORFs from the P97 promoter (Smotkin & Wettstein, 1986). The cell line thus provides the possibility of studying the interaction between cellular factors and the HPV-16 NCR without interference from viral autoregulatory factors.

In order to analyse the transcriptional control imposed by the cell on the viral genome, we have mapped the sites of interaction between cellular DNA-binding factors (CDBFs), which may be regulatory proteins, and their target binding sites in the NCR. Such studies were performed both in situ, by DNase I digestion of viral chromatin in SiHa nuclei, and in vitro, by footprinting and gel retardation experiments.

Other laboratories have previously identified a constitutive enhancer within the HPV-16 NCR, which was responsive to dexamethasone (Gloss et al., 1987) and greatly stimulated expression from a heterologous promoter in HeLa and SiHa cells and primary foreskin keratinocytes (Cripe et al., 1987; Gloss et al., 1987). The enhancer may therefore be dependent upon cell-type specific factors for its activity. More recently, Gloss et al. (1989) reported the detection by DNase I footprinting of numerous in vitro binding sites for various transcription factors in the NCR of HPV-16. These included seven binding sites for NF1, three for AP1 and one glucocorticoid receptor element (GRE), which had been shown to bind partially purified glucocorticoid receptor in vitro (Gloss et al., 1987).
In this study, we have confirmed many of the footprints described by Gloss et al. (1989) and have more importantly shown that the minimal cell-type specific enhancer of HPV-16 is bound by CDBFs in the nuclei of SiHa cells. Furthermore, gel retardation analysis of oligonucleotides across this region suggests that it may be bound by AP1, C/EBP and three novel CDBFs, but not by NF1.

Methods

Cells. SiHa and CaSki cervical carcinoma cell lines (American Type Culture Collection, Maryland, U.S.A.) were grown in roller bottles (Falcon) in Special Liquid Medium supplemented with 10% foetal calf serum and 37.5 μg/ml penicillin (Gibco).

Nuclei and nuclear extract. Nuclei and their extracts were prepared essentially as described by Dignam et al. (1983), with a few modifications. All solutions were supplemented with a mixture of protease inhibitors: 0.5 mM-PMSE, 1 μg/ml pepstatin A, 1 μg/ml leupeptin, 1 μg/ml aprotinin and 0.5 mM-benzamidine (Sigma). Nuclear proteins were extracted with 0.35 M-NaCl buffer and the cleared extract was precipitated by 20 mM-EDTA, 1% SDS and 0.25 mg/ml tRNA followed by phenol/chloroform extraction and ethanol precipitation. The pellets were dissolved, denatured and run on 6% polyacrylamide sequencing gels.

DNase I hypersensitive site (HSS) mapping. DNase I HSS were mapped essentially as described by Wu (1989). SiHa nuclei were resuspended in 0.25 M-sucrose, 10 mM-Tris pH 7.5, 60 mM-KCl and 5 mM-2-mercaptoethanol plus protease inhibitors. Nuclei, in lots equivalent to 30 μg of DNA, were digested with up to 10 μg DNase I (Cooper Biomedical) for 3 min at 24 °C. However, conditions were determined empirically for each batch of nuclei. DNase I digestion was stopped by adding 0.5% SDS and 25 mM-EDTA. DNA was purified from the digested nuclei by proteinase K digestion, phenol extraction stopped by adding 0.5% SDS and 25 mM-EDTA. DNA was finally dissolved, denatured and run on 6% polyacrylamide sequencing gels.

DNase I footprinting. Footprinting analyses were performed by a modification of the method described by Jones et al. (1985). The 5'- or 3'-labelled DNA probes (Fig. 2) were pre-mixed with 5% polyvinyl alcohol (Sigma) and 1 μg poly(dIdC) (Pharmacia) before addition to nuclear extract, typically 10 to 60 μg protein. This was then incubated on ice for 15 min, followed by 2 min at room temperature, before adding 2.5 mM-CaCl2 and DNase I (Cooper Biomedical), typically 0.4 ng for digestion without nuclear extract or 50 to 450 ng for digestion with nuclear extract. The extent of DNase I digestion of DNA probes is critically dependent on the presence of Ca2+ and the concentration of nuclear extract proteins and was therefore determined empirically. Digestion was stopped by adding an equal volume of 200 mM-LiCl, 20 mM-EDTA, 1% SDS and 0.25 mg/ml tRNA followed by phenol/chloroform extraction and ethanol precipitation. The pellets were dissolved, denatured and run on 6% polyacrylamide sequencing gels.

Gel retardation. Band shift analyses were performed after the methods of Dorn et al. (1987) and Goding et al. (1987). Typically, 0.5 ng

Results

Mapping of DNase I hypersensitive sites in HPV-16 chromatin

SiHa nuclei were treated with increasing concentrations of DNase I before purifying the DNA and digesting it with PvuII. The restricted DNA was electrophoresed on a high resolution agarose gel, blotted and hybridized with a HaeIII-Sau3A probe which abuts the proximal PvuII site at nt 552 (Fig. 1). Increasing DNase I digestion revealed a large and complex HSS extending from approximately 500 bp to approximately 1000 bp upstream of the PvuII site (Fig. 1a). The HSS contained seven sub-bands, of which sub-bands 2, 4, 5 and 6 were the strongest (Fig. 1a and 5). The clearest feature displayed by the densitometric analysis of the autoradiograph (Fig. 1b) is a major gap or footprint between sub-bands 2 and 4 of the HSS, which maps to the minimal cell-type specific enhancer first described by the laboratories of H.-U. Bernard (Gloss et al., 1987) and L. Turek (Cripe et al., 1987).

Footprinting analysis of HPV-16 NCR

The NCR was digested with HaeIII and Sau3A and the two fragments, of 334 bp and 665 bp respectively, were subjected to footprinting analysis with nuclear protein extracts from SiHa cells (Fig. 2). Regions of viral DNA bound to cellular proteins are less accessible to DNase I

Oligonucleotides. The oligonucleotides used in these studies were synthesized and purified by Oswei DNA Service, Edinburgh University, U.K. or by C. Lang with an Applied Biosystems 381A DNA synthesizer. Complementary oligonucleotides were annealed as described by Dorn et al. (1987) and purified by polyacrylamide gel electrophoresis. Table 2 displays the sequences of the oligonucleotides used in this study, aligned to maximize their similarity. Oligonucleotides E,F,G and H contain the corresponding footprints in the NCR of HPV-16. Oligonucleotides CP1, CP2 and NF1 are described by Chodosh et al. (1988) and are derived respectively from the promoters of the heat-shock protein 70 gene, the gamma-fibrinogen gene and the replication origin of adenovirus. Oligonucleotide NFY is described by Dorn et al. (1987) and contains the major histocompatibility complex E, Y-box. Oligonucleotide HK6 derives from the human cytokeratin-6 gene promoter (Blessing et al., 1987). TK and TKm oligonucleotides derive from the herpes simplex virus thymidine kinase gene promoter (Graves et al., 1986; Johnson et al., 1987). Oligonucleotides API and ATF derive from elements of the adenovirus E3 gene promoter (Morgan, 1989).
HPV-16 NCR interaction with cellular factors

Fig. 1. DNase I HSSs in HPV-16 chromatin. (a) Southern blot of SiHa chromatin digested for 3 min with 1, 2, 3, 4, 5, 6, 7, 8 and 10 μg DNase I (lanes 2 to 10 respectively). Control lanes (lanes 1 and 11) represent chromatin samples incubated without DNase I. The position of the HSS is related to the viral NCR, depicted to the right with the Sau3A–HaeIII fragment used as a hybridization probe. (b) Densitometric analysis of autoradiographs of HSS bands; low (--), medium (- -) and high ( ) represent 3.0, 5.0 and 8.0 μg DNase I respectively. The densitometric profiles are cumulative.

digestion than naked DNA and appear as relatively clear regions or footprints on a sequencing type gel. In total, 13 footprints were identified, FP-A to M, from both DNA strands (Fig. 2). FP-A and C to H contain a sequence showing different degrees of identity to the K box or cytokeratin promoter element AARCCAAA (Blessing et al., 1987) and to the CCAAT box motif (Dorn et al., 1987) (Table 1). In FP-A, E, G and H the sequence is found in the lower strand (Table 1). FP-A and FP-E, in addition, contain a sequence homologous to the GRE (Gloss et al., 1987). FP-E also contains two sequences similar to the AP1 binding site (Angel et al., 1987; Lee et al., 1987), one of which overlaps the GRE element and the CCAAT-like box (Table 1). FP-I contains a third sequence homologous to the AP1 site (Table 1), FP-L contains a TATA box and FP-M contains one of the E2-responsive elements (Cripe et al., 1987) (Fig. 2b). FP-B, J and K do not show any easily recognizable sequence elements (Fig. 2b).

Gel retardation analysis

The footprinting analysis described above defines only areas of the NCR which bind CDBFs but does not allow identification of the CDBFs themselves. To this end we have conducted gel retardation analyses with synthetic oligonucleotides covering footprints E to H and nuclear extracts from SiHa, CaSk and HeLa cells (epithelial), MRC5 cells (fibroblast) and K562 cells (myeloid). Oligonucleotide-protein complexes migrate on gels more slowly than free oligonucleotide and the shift in migration is dependent on the particular proteins binding.

FP-E to H were chosen because, as discussed below, they overlap with the in situ footprint in the HSS and the cell-type specific or keratinocyte-dependent enhancer (Gloss et al., 1987; Cripe et al., 1987; Fig. 5) and are therefore the most likely targets for epithelium-specific CDBFs. Oligonucleotides E, F, G and H were compared with each other and with a panel of oligonucleotides known to contain high affinity binding sites for specific CCAAT box factors (Table 2). Initial experiments showed that all four oligonucleotides readily bound an apparently ubiquitous CDBF in all nuclear extracts used. However, the prolonged preincubation of nuclear extracts with poly(dIdC) greatly reduced the binding of the factor to oligonucleotides F, G and H and revealed other DNA–protein complexes, F1, G1 and H1 (Fig. 3). Complex formation with oligonucleotide E was little affected.

With oligonucleotide E, two main complexes, E1 and E2, were observed (Fig. 3) with both epithelial and non-epithelial nuclear extracts. The E2 doublet was competed by oligonucleotide E and to a lesser extent by oligonucleotides TK, TKm and HK6 (Fig. 4). Moreover it was heat-stable and comigrated with the complexes formed with oligonucleotides TK or TKm (G. J. Sibbet, unpublished results; Dorn et al., 1987; Raymondjean et al., 1988), thus leading to the tentative identification of this factor as the CCAAT box or enhancer-binding protein C/EBP (Graves et al., 1986; Johnson et al., 1987). The competition by oligonucleotide HK6 may be a reflection of the binding site promiscuity of C/EBP (Table 2) for, as well as binding avidly to CCAAT box and enhancer core homologies, C/EBP also binds tightly to promoter elements of the albumin, alpha1-antitrypsin and transthyretin genes, binding sites which share minimal, if any, DNA sequence similarity (Costa et al., 1988).

The slower migrating complex E1 (Fig. 3) was fully competed by oligonucleotides E and AP1 and to a lesser
Fig. 2. Footprinting analysis of the viral NCR. (a) The two *Sau3A–HaeIII* fragments (7432 to 7766 nt (lanes 1 to 4) and 7766 to 525 nt (lanes 5 to 8)) were digested with DNase I with (lanes 2, 4, 6 and 8) or without (lanes 1, 3, 5 and 7) prior incubation with SiHa nuclear extracts. The footprints on the upper (lanes 1, 2, 5 and 6) and lower (lanes 3, 4, 7 and 8) strand are shown for both fragments (b) The *Sau3A* HaeIII fragments used in footprinting experiments. (c) Nucleotide sequence of the footprinted region of the NCR. The footprints A to M are indicated above the sequence for the upper strand and below it for the lower strand. The nucleotide numbering is in relation to P97, the main HPV-16 promoter.
Table 1. Common motifs within footprints

<table>
<thead>
<tr>
<th>Identity to Footprint</th>
<th>Sequence</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cytokeratin promoter element</td>
<td>AARCCAAA TGTGCCAAA A7472</td>
</tr>
<tr>
<td></td>
<td>C 755GCGCCCAAA 7559</td>
</tr>
<tr>
<td></td>
<td>D 7684TTGCCAAC 7592</td>
</tr>
<tr>
<td></td>
<td>E 7654TGAACAAAT 7646</td>
</tr>
<tr>
<td></td>
<td>F 7672TGCACCAAT 7680</td>
</tr>
<tr>
<td></td>
<td>G 7718AAGGCCAA 7710</td>
</tr>
<tr>
<td></td>
<td>H 7749TATGCGCAAT 7741</td>
</tr>
<tr>
<td>C/EBP consensus</td>
<td>CCAAT box consensus</td>
</tr>
<tr>
<td>TGACTCA</td>
<td>GCAAT</td>
</tr>
<tr>
<td>API consensus</td>
<td>AP1 consensus</td>
</tr>
<tr>
<td></td>
<td>A761TTAGCTA 7617</td>
</tr>
<tr>
<td>GRE consensus</td>
<td>GTTACANNNNTTGTTC</td>
</tr>
<tr>
<td>A 7475GCCACAAAAAATGCT 7489</td>
<td></td>
</tr>
<tr>
<td>E 7641TGTACATTGTCAT 7655</td>
<td></td>
</tr>
</tbody>
</table>

Table 2. Oligonucleotides in competition experiments*

<table>
<thead>
<tr>
<th>Oligonucleotide</th>
<th>Sequence†</th>
</tr>
</thead>
<tbody>
<tr>
<td>Oligo E†</td>
<td>gatcTTATATGACAAATGCTGATAGTACGATGAGTGGC</td>
</tr>
<tr>
<td>Oligo F†</td>
<td>gatcCAGATCGGCCCAAGGCCCATCAGA</td>
</tr>
<tr>
<td>Oligo G†</td>
<td>gatcTTACAGTTCCAGCCCAAACATTTTA</td>
</tr>
<tr>
<td>Oligo H†</td>
<td>gatcTTATTCATGGCAATGGGAGA</td>
</tr>
<tr>
<td>CPI†</td>
<td>gatcTTCTCTCTCTTTGACCAATCGCA</td>
</tr>
<tr>
<td>CP2†</td>
<td>gatcTTGACCGGTTCAAGCCACCTTTTA</td>
</tr>
<tr>
<td>NF1†</td>
<td>gatcTTTGGCGATTGACGAATGCAATAG</td>
</tr>
<tr>
<td>NFY†</td>
<td>gatcTTTTAACACATGAAAAATG</td>
</tr>
<tr>
<td>HK6†</td>
<td>gatcCAAGGAAGCGACAAATGCAATCA</td>
</tr>
<tr>
<td>TK†</td>
<td>gatcTGAAATTGCCCAATGACAAAGCG</td>
</tr>
<tr>
<td>TKm†</td>
<td>gatcTGAAATTGCCCAATGACAAAGCG</td>
</tr>
<tr>
<td>API</td>
<td>AGCTTCGAAATGACAGATGTCATACCTGAGG</td>
</tr>
<tr>
<td>ATF</td>
<td>GACCGACACCTGTGACGAAGGCCCGCCAGG</td>
</tr>
</tbody>
</table>

* One strand of the sequence only is displayed.
† E to TKm are aligned to indicate CCAAT box-like homology; API and ATF aligned according to their respective motifs. 
‡ E to TKm are terminated with BamH1/BglII linkers.

extent by oligonucleotide ATF (Fig. 4), suggesting that FP-E contains a functional API or API-like binding site; oligonucleotides CPI, TK and TKm were also strong but less effective competitors. However, the CDBF in the E1 complex is unlikely to be a CPI-like CCAAT-box binding factor because oligonucleotide NFy, which contains a strong binding site for the CPI-like factor (Dorn et al., 1987), did not compete with oligonucleotide E (G. J. Sibbet, unpublished results). In addition, both oligonucleotides TK and TKm were equally effective as competitors of the E1 complex, although the CCAAT→GCAAT mutation in TKm

Fig. 3. Band shift analysis of oligonucleotides E to H. Lane 1, oligo E; lane 2, oligo F; lane 3, oligo G; lane 4, oligo H. The labelled oligonucleotides were incubated with SiHa nuclear extracts. Free unreacted oligonucleotide is indicated; the complexes due to the binding of oligonucleotides to CDBFs are indicated as E1 to H1; the doublet band due to interaction of oligonucleotides with C/EBP is indicated only for oligonucleotide E as E2.

Fig. 4. Competition of protein–oligonucleotide complex formation. The competed oligonucleotides are indicated at the left and the competing oligonucleotides at the top of the figure; the unlabelled competing oligonucleotides were added at a 100-fold molar excess over the labelled oligonucleotides. The two complexes formed by oligonucleotide E are indicated 1 and 2 in brackets.
prevented the oligonucleotide binding the CP1-like factor (G. J. Sibbet, unpublished results; Goding et al., 1987). A more likely explanation for the competition of E1 by oligonucleotides CP1, TK and TKm is that they all contain motifs similar to the AP1-binding site (oligonucleotide CP1, TGAGCCA; oligonucleotide TK/Tkm, TGACAAG) (Table 2). Furthermore the E1 complex and that formed with labelled oligonucleotide AP1 were indistinguishable (G. J. Sibbet, unpublished results). These results strongly suggest that E1 is a complex between AP1 and oligonucleotide E.

With oligonucleotide F, a complex F1 was detected (Fig. 3) which, although comigrated with the H1 complex, was competed only by an excess of oligonucleotide F (Fig. 4). Despite the presence of a CCAAT box-like sequence in oligonucleotide F, none of the other oligonucleotides had any effect on complex formation, suggesting that the CDBF binding to the F sequence in these and in the footprinting experiments is a novel factor. No F1 complex was observed with MRC5 fibroblast extracts (G. J. Sibbet, unpublished results).

Oligonucleotide G formed a complex G1 which migrated faster than either E1, F1 or H1, but slower than a residual non-specific complex comigrating with E2 (Fig. 3). G1 was competed by oligonucleotides G, E and H (Fig. 4), reflecting the high level of homology between these oligonucleotides (Tables 1 and 2), and by oligonucleotide HK6, suggesting that G may be a target for the same CDBF(s) that binds the conserved cytokeratin promoter element. Indeed, oligonucleotide HK6 forms a complex with SiHa nuclear extract which comigrates with G1 and is specifically competed by oligonucleotide G (G. J. Sibbet, unpublished results).

Oligonucleotide H formed a complex H1 (Fig. 3) which comigrated with E1 and F1, but not with G1. Although H1 was competed by oligonucleotides E and G as well as oligonucleotide H, it was not competed by oligonucleotide F, or by oligonucleotides AP1 and HK6, which do compete for CDBF binding to oligonucleotides E and G respectively (Fig. 4). These data show that the CDBF binding to oligonucleotide H is distinct from those binding to oligonucleotides E, F and G. Complexes G1 and H1 were observed both with epithelial and fibroblast extracts (G. J. Sibbet, unpublished results).

The minor bands seen with oligonucleotides F and H (Fig. 3) were competed by the panel of oligonucleotides to the same extent as the major bands F1 and H1.

Discussion

The interaction of the NCR of HPV-16 with host cell factors has been investigated in situ by DNase I digestion of viral chromatin in SiHa nuclei, and in vitro by footprinting and gel retardation analysis of cloned and synthetic fragments of NCR DNA with nuclear extracts. There is good agreement between the results obtained from the different experiments: the footprints overlap and interdigitate with the sub-bands within the DNase I HSS (Fig. 5) and oligonucleotides corresponding to selected footprints do associate with specific cellular elements.
proteins. All the results point to the same conclusions, namely that there is extensive interaction between cellular DNA-binding factors and the viral regulatory region, evidence of complex transcriptional regulation.

Most of the interactions with CDBFs, if not all, take place in the 3' half of the NCR; we have not observed any HSSs in the 5' half, not did we detect any footprints. This agrees with previously published HSS mapping data for HPV-18 in HeLa cells (Rossl et al., 1989), but is at variance with the footprint data of Gloss et al. (1989) who found footprints extending further upstream when high concentrations of nuclear extract were used. This may reflect the difference between the actual binding of CDBFs in HSS mapping in situ and their potential to bind in footprinting analysis in vitro, which is dependent upon the binding conditions, the concentration of added nuclear extract and the binding affinities of individual CDBFs.

The size of the HSS is consistent with the displacement of two nucleosomes from the 3' half of the NCR (Fig. 1 and 5), and the complexity of the sub-band pattern within the HSS is thought to reflect the binding of CDBFs in place of the nucleosomes (Wu, 1989). The large gap within the HSS between the stronger sub-bands 2 and 4 (Fig. 1 and 5) probably represents an extended footprint bound in situ by CDBFs. This region is of particular interest because it maps to the minimal cell-type specific HPV-16 enhancer (Gloss et al., 1987; Cripe et al., 1987) and encompasses the area of sequence homology in the NCRs of HPV-16, 18 and 33 (Fig. 5; Table 3).

To try to identify the CDBFs interacting with the NCR, we performed a DNase I footprinting analysis across the NCR with SiHa nuclear extract and a band shift analysis with oligonucleotides containing selected footprints; 13 footprints were detected. Footprints were mapped to the previously recognized AP1 and GRE sites in FP-E (Gloss et al., 1989) and to the viral E2 site in FP-M (Gloss et al., 1987; Cripe et al., 1987). In addition, other footprints were identified containing a common motif 5'GCCAA3' related to both the cytokeratin promoter-like element AARCCAAA (Blessing et al., 1987) and the CCAAT box (Dorn et al., 1987; Raymondjean et al., 1988). However, despite their sequence homology with the CCAAT box motif, none of the footprinted sites bound the factors known to associate with the motif, namely CP1, CP2, or NF1, as shown by the lack of competition by the oligonucleotides containing their respective high affinity binding sites. Gloss et al. (1987) have suggested that NF1 binds as a monomer to a number of sites, including FP-F, G and H (respectively fp5e, fp6e and fp7e in their nomenclature). We have no evidence to support this. In our hands, these sites bind three distinct CDBFs, none of the complexes band shifts with NF1 (G. J. Sibbet, unpublished results) nor are they successfully competed by an oligonucleotide with a binding site for NF1, although complex G1 is partially competed by a 100-fold excess.

Our differing findings probably reflect the distinct origins of the NF1 used in the respective studies. The NF1-binding activity we observed appears to be identical to that of the NF1/CTF isolated from HeLa cells by Jones et al. (1987), whereas the activity described by Gloss and colleagues is derived from porcine liver and produces a distinct band pattern. Furthermore it is not clear whether NF1/CTF from HeLa cells can bind as a monomer because its footprint on the α-globin CCAAT box (equivalent to a 'half-NF1 site') is as large as that over a full NF1 site (Jones et al., 1987) and the affinity of NF1 for the CCAAT box is 100-fold lower than that for the full NF1 site (R. Hay, personal communication). Indeed, even increasing the spacing between the halves of the NF1 site's dyad symmetry by one nucleotide greatly reduces the affinity of NF1 for its binding site (Miksicek et al., 1987). Thus, apart from nomenclature, the NF1 from porcine liver and that from HeLa or SiHa cells may be very different.

An element in FP-E is known to confer responsiveness to glucocorticoid and to bind purified glucocorticoid receptor (Gloss et al., 1987; Chan et al., 1989). Our results indicate that it may also bind C/EBP and AP1 at two distinct sites. Interestingly the binding of these CDBFs can be separated by cleavage of oligonucleotide E in such a way that AP1-binding is retained in both 'half' oligonucleotides but C/EBP binding is completely lost (G. J. Sibbet, unpublished results). Similar AP1-binding sites have been demonstrated in the NCR of HPV-18 by Garcia-Carranca et al. (1988). Intriguingly the binding sites in HPV-16 for these three factors abut or overlap, which may imply an interaction between phorbol ester-mediated and glucocorticoid-mediated regulation of HPV-16 expression. Indeed, Chong et al. (1990) have recently shown that mutagenizing both AP1-binding

### Table 3. NCR homology block shared by HPV-16, -18 and -33*

<table>
<thead>
<tr>
<th>Virus</th>
<th>Sequence</th>
</tr>
</thead>
<tbody>
<tr>
<td>HPV-16</td>
<td>G</td>
</tr>
<tr>
<td></td>
<td>H</td>
</tr>
<tr>
<td>HPV-18</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td>HPV-33</td>
<td></td>
</tr>
</tbody>
</table>

* Footprints G and H in HPV-16 and non-homologous nucleotides in HPV-18 and -33 are indicated.
sites disrupts the constitutive enhancer function but stimulates its response to dexamethasone fivefold.

FP-F, G and H appear to bind novel specific factors which are under further investigation. The CDBF involved in complex F1 is particularly interesting on two counts: it is strongly sequence-specific and appears to be present only in epithelial cells. The location of FP-F in the keratinocyte-specific enhancer supports the speculation that F is responsible for the expression of HPV-16 being confined to epithelial cells. These data will be presented elsewhere.

The successful competition of the complex G1 by oligonucleotide HK6, the cross-competition of the HK6-CDBF complex by oligonucleotide G and the equivalent comigration of the two complexes suggest that FP-G is most likely to be bound by the same factor(s) as bind to the cytokeratin gene promoter element and thus this factor(s) may have a role in cell-type or differentiation-specific expression of HPV-16. However, the element conferring cell-type specificity to the cytokeratin gene promoters is not the AARCCAA sequence (Blessing et al., 1989) and, accordingly, complex G1 is also found in non-epithelial cells, albeit with a slightly different mobility (G. J. Sibbet, unpublished results). It will be interesting to examine any cell-type specific modifications of this factor.

The data we have presented here do not directly examine the regulation of HPV-16 expression but they do suggest a role for a number of CDBFs. Experiments to link the binding of particular CDBFs to HPV-16 gene regulation are currently being pursued by mutational analysis of the HPV-16 NCR. In addition to the role of cellular factors in determining epithelium-specific viral transcription, one might expect that the further modulation of viral expression in the different layers of the epithelium would be achieved through synergistic interactions between several CDBFs. This is also being examined by multiple mutational analysis.

We wish to thank the following colleagues in the Beatson Institute: Caron Wilson for her initial contribution to the work, Andrew and Christine Lang for providing HeLa and K562 nuclear extracts, Jane Parker for providing MRCS nuclear extracts, Iain Morgan for providing oligonucleotides AP1 and ATF and David Gillespie and Mark Plumb for critically reading the manuscript. Part of this work was presented to the 1989 UCLA Symposium on Papillomaviruses. We are grateful to the Cancer Research Campaign for continuous financial support. M.S.C. is a CRC Fellow.

References


Identification of a rat liver protein that binds to the enhancer core element of three animal viruses *Genes and Development* 1, 133–146.


(Received 11 May 1990; Accepted 13 July 1990)