Delineation of human papillomavirus type 18 enhancer binding proteins: the intracellular distribution of a novel octamer binding protein p92 is cell cycle regulated

Hans-Dieter Royer, Markus P. Freyaldenhoven, Inge Napierski, Dimitri D. Spitkovsky*, Tobias Bauknecht and Nina Dathan
Angewandte Tumorvirologie, Deutsches Krebsforschungszentrum, Im Neuenheimer Feld 506, D 6900 Heidelberg and 1European Molecular Biology Laboratory, Meyerhofstrasse 1, 6900 Heidelberg, FRG

Received January 25, 1991; Revised and Accepted April 9, 1991

ABSTRACT

The enhancer of human papillomavirus type 18 consists of two functionally redundant domains, one is partially conserved between HPV18 and HPV16, both mediate strong transcriptional enhancement. In contrast, short fragments of the enhancer mediate low transcriptional enhancement, suggesting that there is functional cooperation between HPV enhancer binding factors. Previously interactions of the enhancer with NF-1, AP1 and steroid receptors were shown by EMSA. Here we show by binding site blotting, that four novel sequence specific proteins p110, p92, p42 and p40 bind to the enhancer. Nuclear proteins p110 and p92 bind at repeated sites in the enhancer, proteins p42 and p40 only at one site. Recognition sequences for p110 and p92 were identified in a TTGCTTGCATAA sequence motif and consist of an overlapping p110 and p92 recognition site. The specific interaction of p110 with G residues of this 12 nucleotide long sequence was demonstrated by a mutant recognition site. Single recognition sites for p42 and p40 were localized in the enhancer by the use of overlapping oligonucleotides. In addition, electrophoretic mobility shift analysis identified Oct-1 and AP2 interactions with the enhancer. The AP2 binding site was mapped to a AGGCACATATT motif. The p92 protein binds to enhancer oligonucleotides, containing at least one copy of Oct-1 like recognition sequences, these oligonucleotides also bind synthetic Oct-1 protein. During serum starvation or at high saturation density, p92 moves from the nucleus into the cytoplasm. Immunoblots of cytoplasmic extracts with anti-Oct-1 antisera showed that p92 is a novel octamer binding factor, which is not immunologically related to the Oct-1 protein. The intracellular p92 distribution is regulated at the G0/G1 boundary of the cell cycle, by nuclear-cytoplasmic translocation.

INTRODUCTION

Cervical cancer represents the second most frequent cancer in females on a worldwide scale. DNA of human papillomaviruses with high oncogenic potential is found in over 90% of cervical cancer biopsies (77). Two proteins with transforming activity, E6 (49, 70) and E7 (27, 45, 72), are encoded by viral early genes and their continuous expression is required for maintenance of the proliferative and transformed phenotype (69, 10). The transforming activities of E6 and E7 proteins are at least in part explained by the fact that they interact specifically with products of tumor suppressor genes p105* (12, 14, 20, 39, 73) and p53 (17). HPV 18 early gene expression is controlled by the upstream regulatory region (URR), which may be considered to comprise three functional units. The most 5' region, adjacent to the L1 gene, is E6 responsive and the most 3' region, which contains the early gene promoter, is E2 responsive (22, 65). The E2 protein is a papillomavirus regulatory protein (48). Between these two regions lies the enhancer. The HPV18 enhancer is located on a 230 nucleotide long Rsal-Rsal fragment, and can be subdivided into two functionally redundant domains of similar size, whose activity depends on cellular factors (65, 66).

In benign genital warts the HPV genome is normally maintained as an episome (13, 38); in tumors and tumor derived cell lines, however, at least some of the viral DNA is integrated into the host genome (6, 13, 67). Integration appears to occur early in neoplastic development (38, 55). In almost all cases, integration interrupts the E1 and E2 open reading frames (ORFs) and leads to a loss of functional E2 protein. Consequently, the early gene promoter is no longer E2 dependent and it is assumed that the promoter comes under the control of host cell factors (reviewed in 71). In tumors and tumor derived cell lines the E6 and E7 ORFs are consistently retained and expressed (3, 56). The elucidation of mechanisms which regulate the expression of these transforming genes is of critical importance for the understanding of cervical carcinogenesis.

The enhancers of HPV16 and HPV18 contain recognition sites for nuclear factor I (NF1), activator protein AP1, glucocorticoid

* On leave from the All Union Cancer Institute, Moscow, USSR
receptor and an uncharacterized factor, papillomavirus enhancer associated factor (7, 8, 9, 21, 24, 59). The activities of individual cis acting elements contribute to full enhancer activity. Published data suggest that HPV enhancer function depends on the cooperative interaction of multiple factors (7), because short segments of the enhancer have only weak transactivating function, whereas the complete enhancer is a strong transactivator. Frequently recognition sites bind multiple proteins, and individual factors can interact with different recognition sequences (15, 28, 32, 34, 54). To gain more insight into the complexity of DNA-protein interactions with the human papillomavirus type 18 enhancer, we have used a binding site blotting approach to delineate HPV18 enhancer binding proteins (19, 43). With this approach we have identified several new enhancer binding factors. It is our long term goal to identify regulatory factors which play a key role in the activation of HPV early gene expression in cervical carcinomas. The delineation of HPV18 enhancer binding proteins is a first step in this analysis which will ultimately lead to the purification of these novel factors, to the isolation of corresponding cDNA clones and their functional characterization.

MATERIALS AND METHODS
Cell Culture and Cell Lines
Human cervical cancer cell lines were maintained in Dulbecco’s Modified Eagle’s Medium (DMEM) supplemented with 10% fetal calf serum. Cervical cancer derived cell lines were provided by colleagues of the Institute of Angewandte Tumorvirologie at the Cancer Center.

Extraction of Nuclear and Cytoplasmic Proteins
Nuclear extracts of cervical cancer derived cell lines and human fibroblasts were prepared according to a published procedure (58). After detergent lysis with 0.65% NP40 nuclei are prepared by low speed centrifugation at 4°C and proteins eluted with 520mM NaCl with slight agitation. For storage the eluates were dialyzed against a buffer containing 50% glycerol, 50mM NaCl, 520mM NaCl with slight agitation. For storage the eluates were dialyzed against a buffer containing 50% glycerol, 50mM NaCl, 10mM Heps (pH7.9), 0.5mM PMSF and 0.5mM DTT. Protein concentrations were determined by a colorimetric assay (Biorad) using serum albumin as a standard. Nuclear proteins from Hela cells and a B lymphoblastoid line (Laz 509) were enriched by heparin Sepharose chromatography as described (76). Cytoplasmic proteins were prepared after cell lysis and removal of nuclei by low speed centrifugation. Cytoplasmic proteins were dialyzed as described above for nuclear proteins.

Synthetic Oligonucleotides
Single stranded oligonucleotides and the complementary strands were synthesized on Applied Biosystems DNA synthesizers and purified by preparative denaturing acrylamide gel electrophoresis (40). Full length bands were identified by UV shadowing, excised, eluted by diffusion in 500mM ammoniumacetate and ethanol precipitated. To generate double stranded oligonucleotides complementary strands were annealed at temperatures minus 3°C below the specific melting points (Tm). Radiolabeled double stranded oligonucleotides were prepared with polynucleotide kinase and 32P-gamma ATP. The sequence for the recognition site of AP2 was from the human metallothionein IIA (hMT-IIA) gene control region (31). The human papillomavirus enhancer sequence (66) was used for synthesizing enhancer oligonucleotides RP2, RP3, RP3/4, RP4, RP5, S1tet and S2tet. The sequences are listed in table 1.

Binding Site Blotting
Nuclear proteins were size fractionated on 8% SDS polyacrylamide gels (35) and transferred to nitrocellulose at 150mA overnight at room temperature with a horizontal blotting chamber (IBI) in Laemml running buffer without SDS. After blotting bound proteins were denatured in situ by 6M guanidinium hydrochloride and renatured by sequential dilution of guanidinium as described (19, 68). Nitrocellulose membranes were blocked with 5% nonfat dry milk (Carnation) for 30 min at room temperature and subsequently incubated with end labeled double stranded oligonucleotides at 5 x 10^5 cpm/ml in the presence of 5ug/ml poly(dIdC)(dIdC) as nonspecific competitor. The buffers for DNA binding and conditions for washing binding site blots were as described (19).

Immunoblotting with anti-Oct-1 antiserum
Immunoblots were prepared as for binding site blotting, with a guanidinium denature renature cycle. For immunodetection of Oct-1 the blocked membranes were incubated with a 100 fold dilution of Oct-1 antiserum (49) in 0.5% dry milk. For the detection of bound antibodies, iodinated protein A (Amersham) was used at 100000 cpm/ml. After washing in 0.5% dry milk 20min with three changes blots were dried and exposed to X ray film.

Electrophoretic mobility shift analysis (EMSA)
EMSA was carried out as described (60). For EMSA of enhancer domain I, functional domain I of the HPV18 enhancer was isolated from pURR 18, a clone containing the upstream regulatory region of HPV18, which is located on a 1050 bp BamHI fragment (6). In this fragment the enhancer is present on a Rsal-Rsaf fragment (NT 582—810). This fragment was isolated by polycarylamide gel electrophoresis and electrophoresis after Rsal and Fnu4HI I digestion. The digestion with Fnu4HI I (position 910) removes a comigrating promoter fragment. For EMSA of one functional domain, which is located on a BstNSI-Rsaf fragment (nucleotides 697—810), the Rsaf(582)-Rsaf(810) fragment was digested with BstNI and radiolabeled with polynucleotide kinase. The fragment containing enhancer domain I was isolated by polycarylamide gel electrophoresis and electrophoresis. The incubations of enhancer domain I with nuclear protein contained 5 μg nucleic extract and 2.5 μg poly(dIdC)(dIdC) as nonspecific competitor. For EMSA DNA’s were incubated in binding buffer (60), for 25 min at room temperature and loaded on a low salt polycarylamide gel (4% total monomer,30:1 acrylamide/N,N’-methylenebisacrylamide ratio). Electrophoresis was at 11V/cm for 90 min at room temperature. Gels were dried and exposed to X ray film overnight.

Synthesis of Oct-1 protein in vitro
Recombinant Oct-1 protein was synthesized from Oct-1 mRNA in a rabbit reticulocyte lysate (Promega), with S35 methionine as radioactive amino acid. The Oct-1 mRNA was transcribed in vitro from HindIII cleaved pBS Oct-I (64) with T7 polymerase (Statragene).
RESULTS

Delineation of HPV18 enhancer binding proteins

In this study we have used a binding site blotting approach for the visualization of HPV18 enhancer binding proteins, using double stranded enhancer oligonucleotides. We focused on one of the two functionally redundant domains of the enhancer, domain 1. The domain is represented by oligonucleotides RP2, RP3, RP3/4, RP4, whose sequences are given in table 1. The relative positions of the oligonucleotides within the enhancer are shown in Figure 1E. In nuclear extracts of human fibroblasts, Hela cells and Hela-fibroblast hybrid cells (61) DNA binding activities were detected, which bind to oligonucleotides RP2, RP3, RP3/4 and RP4 (Fig 1 A,B,C,D). Based on the apparent molecular weights, these proteins were termed p110, p92, p42 and p40. In addition, minor DNA binding activities were present in nuclear extracts, which were not analyzed further. Protein p85, present wherever p92 was detected, is a derivative of p92, generated after cell lysis and nuclear extraction (unpublished observation). The recognition sequences for p110 and p92 apparently are reiterated amongst the oligonucleotides. p110 binds to RP2, RP3, RP3/4, and RP4, whilst p92 binds to RP3, RP3/4 and RP4 (Fig 1). p110 was not detected in normal human fibroblasts (four independent preparations). The recognition sites for p40 and p42 are present once in enhancer domain I within oligonucleotide RP4. The failure to detect p40 and p42 with the partially overlapping oligonucleotide RP3/4 confines the recognition sequences to the twelve nucleotides unique to RP4: AGTTTGTGTTTTTA. This sequence has a distant similarity to the AP2 recognition site (24). Oligonucleotides RP3 and RP5 (lanes 2 and 5) did not inhibit complex formation whereas oligonucleotides RP3/4 and RP4 efficiently competed with AP2 binding. This result maps the AP2 recognition site to the overlap between the RP3/4 and RP4 sequences. The sequence of this region, TAATTAAACGCGTTTTAGGCACATATTTTT, contains a motif present in the URMs of many human papillomaviruses: AGGCACATAT. By mutational analysis it was shown that a 92 nucleotide long HPV16 enhancer fragment, which consisted of two AP1 sites, two NF1 sites and the AGGCACATAT motif, lost 45% of its activity in the reporter plasmid pBLCAT2, when the AGGCACATAT site was mutated (7). The mutation changes the AGGCACATAT motif to AGgtatctc (7). The AP2 oligonucleotide (31) used here to identify the AP2 DNA interaction with the enhancer contained two AP2 binding sites from the human metallothionein gene IIa (hMTIIa) gene control region. The sequence of one AP2 site, 5'-GGTGTTTGCACCC-TGGAG, reads on the antisense strand: 5'-TTCCAGGGAGA-AACC. Comparison between the RP3/4 sequence, the conserved AGGCACATAT motif and the AP2 recognition site (lower strand) suggests 5'-AGGC as a common motif. The mutation reducing activity of the 92 nucleotide HPV16 enhancer fragment, destroys this 5'-AGGC motif. In addition, the AGGC motif in the AP2 recognition site was shown to bind purified AP2 with only little reduced affinity (42). From these published data we conclude that the AP2 interaction identified here by EMSA, occurs most likely by AP2 binding to the AGGC sequence of the larger AGGCACATAT motif. The functional significance of this interaction with respect to early gene promoter regulation will be part of future work. Further experiments will determine whether AP2 is the only factor binding to the AGGCACATAT motif. The 52K AP2 protein was not detected by binding site blotting, presumably because AP2 DNA binding activity cannot be renatured after SDS-PAGE (42).

A sequence motif: AGGCACATAT which is present in many human papilloma virus URMs interacts with activator protein AP2

In a separate line of experiments, DNA protein interactions with enhancer domain I of the HPV 18 URR were examined by electrophoretic mobility shift analysis (EMSA). A fragment representing domain I (65) forms multiple retarded protein DNA complexes with nuclear proteins from tumorigenic Cgl3 cells (Figure 2). It is possible to identify known transcription factors by EMSA, by the inclusion of synthetic recognition sites in the binding reaction as unlabeled competitors in molar excess, in order to eliminate specific complex formation. In this way, the interaction of activator protein 2 (AP2) with enhancer domain I was identified (Figure 2, lane 6). AP2 is a 52Kd transcription factor which binds to specific recognition sequences and mediates induction by two different signal transduction pathways: protein kinase C and cAMP (31, 42, 74). The arrow in Figure 2 points to a retarded protein:DNA complex which is not formed in the presence of excess AP2 recognition sequences (GGGTGTTT-CGCCCTGGA)2 and may thus be assumed to contain AP2. Subsequently unlabeled enhancer oligonucleotides RP3, RP3/4, RP4 and RP5 were used as competitors to map the AP2 recognition sequence within the enhancer. RP5 (table 1) is part of the enhancer and contains a well characterized NF1 recognition site (24). Oligonucleotides RP3 and RP5 (lanes 2 and 5) did not inhibit complex formation whereas oligonucleotides RP3/4 and RP4 efficiently competed with AP2 binding. This result maps the AP2 recognition site to the overlap between the RP3/4 and RP4 sequences. The sequence of this region, TAATTAAACGCGTTTTAGGCACATATTTTT, contains a motif present in the URMs of many human papillomaviruses: AGGCACATAT. By mutational analysis it was shown that a 92 nucleotide long HPV16 enhancer fragment, which consisted of two AP1 sites, two NF1 sites and the AGGCACATAT motif, lost 45% of its activity in the reporter plasmid pBLCAT2, when the AGGCACATAT site was mutated (7). The mutation changes the AGGCACATAT motif to AGgtatctc (7). The AP2 oligonucleotide (31) used here to identify the AP2 DNA interaction with the enhancer contained two AP2 binding sites from the human metallothionein gene IIa (hMTIIa) gene control region. The sequence of one AP2 site, 5'-GGTGTTTTGCCCTGGA, reads on the antisense strand: 5'-TTCCAGGGAGAAAAACC. Comparison between the RP3/4 sequence, the conserved AGGCACATAT motif and the AP2 recognition site (lower strand) suggests 5'-AGGC as a common motif. The mutation reducing activity of the 92 nucleotide HPV16 enhancer fragment, destroys this 5'-AGGC motif. In addition, the AGGC motif in the AP2 recognition site was shown to bind purified AP2 with only little reduced affinity (42). From these published data we conclude that the AP2 interaction identified here by EMSA, occurs most likely by AP2 binding to the AGGC sequence of the larger AGGCACATAT motif. The functional significance of this interaction with respect to early gene promoter regulation will be part of future work. Further experiments will determine whether AP2 is the only factor binding to the AGGCACATAT motif. The 52K AP2 protein was not detected by binding site blotting, presumably because AP2 DNA binding activity cannot be renatured after SDS-PAGE (42).

Table 1. HPV18 enhancer oligonucleotides used in this study.

<table>
<thead>
<tr>
<th>Oligo</th>
<th>Map position</th>
<th>Sequence</th>
</tr>
</thead>
<tbody>
<tr>
<td>RP2</td>
<td>664-706</td>
<td>AAGGGGCACCTGTATTACCTCATTNTTCTTGTCAGTGGCGCTA</td>
</tr>
<tr>
<td>RP3</td>
<td>705-742</td>
<td>TACAACAATTTGTTTCGATAACTATAACTCCTCCTATG</td>
</tr>
<tr>
<td>RP3/4</td>
<td>728-771</td>
<td>TGACACTCTCCTATGTAATAAAACTGCNTTAGGCACATATTTT</td>
</tr>
<tr>
<td>RP4</td>
<td>741-783</td>
<td>TACTTAAGGCGACCTGGTTTGCACCTGGTAATAAAACTGCCTGTTAGGCACATATT</td>
</tr>
<tr>
<td>RP5</td>
<td>783-816</td>
<td>GCTGCTTTTAGGCACATATTTTACTATATCCCTCATCCTCCTATG</td>
</tr>
<tr>
<td>S1tet</td>
<td>713-724</td>
<td>TTTTTCTGAATTGTCATTTGTGGCAGTGAACAATG</td>
</tr>
<tr>
<td>S2tet</td>
<td>787-800</td>
<td>AAGGTCATATTGCAATA</td>
</tr>
<tr>
<td>AP2</td>
<td>hMT-IIA</td>
<td>GGTTGTTTGGCTGGA</td>
</tr>
</tbody>
</table>

*Coordinates of oligonucleotides are nucleotide positions in the upstream regulatory region (URR) according to the published sequence (66)
Proteins p110 and p92 bind to a TTGCTTGCATAA motif
In order to investigate the sequence specific binding of p110 and p92, we sought to identify single recognition sites for p110 and p92, and subsequently establish the presence of such sites in all oligonucleotides which bind p92 and p110. The sequence of oligonucleotide RP3 is partially protected from DNaseI digestion at low protein concentration in DNaseI footprint experiments (unpublished data, and 21). The protected sequence is TTGCTTGCATAA. This sequence was synthesized as tetramer (S1tet) for binding site blotting. S1tet was found to bind p110 and p92 in nuclear extracts of the cervical cancer derived cell line SiHa (Figure 3, lane 1) and in nuclear extracts from Hela cells and Hela-fibroblast hybrid cells (not shown). The nature of two proteins with lower molecular weight (Figure 3) was not examined further. This result suggests that recognition sites for both p110 and p92 are present in this twelve nucleotide long sequence.

Nuclear protein p110 is a sequence specific DNA binding protein
The S1 sequence TTGCTTGCATAA, which is a subsequence of RP3, contains two GC pairs separated by two T residues. The HPV18 enhancer contains a very similar sequence, AAGCGTAATTGCATA, which is present in oligonucleotide RP5. A tetramer of this sequence (S2tet) does not bind p110 in binding site blots (Figure 4A). In S2tet the TTGCTCATAA motif of the S1 sequence TTGCTTGCATAA is created at the head to tail fusions of the S2 monomers: 5′AAGCTTAATTGCATA-AAGCTTAATTGCATA. The 5′GC pair of the S2 monomer is separated by 5 residues TAATT from the 3′GC pair of the S2 monomer sequence. Therefore, the S2tet sequence is a S1 mutant, where the TTGCTCATAA motif is present in both oligonucleotides, and the spacing between the two GC pairs is increased to five nucleotides. As S1 mutant, the sequence of S2 tet can be written: 5′aagCTtaTTGCATA-aagCTtaTTGCATA-b. In nuclear extracts of cervical cancer derived cell line SiHa p110 and p92 are present, (Figure 4A, lane 1), as revealed by RP3 binding, and also bind to S1tet (lane 2). p110, however, does not bind to S2tet (lane 3). This result suggests that p110 and p92 binding sites are non-identical but may overlap, and that the 5′GC pair is important for p110 binding. It is also possible that the replacement in S2 of the two most 5′ S1 T residues by two A residues affects p110...
Octamer related sequences are present and overlap in enhancer oligonucleotide RP3 (see table 2). The octamer related sequence 1: AATTGCTT has a 6/8 match and sequence 2: TTGCATAA a 5/8 match with the octamer consensus sequence. The underlined residues, ATTTGCAT, required for Oct-1 binding are also present in octamer related sequence: AATTGCTT (32). S1tet and S2tet bind p92, and octamer related motifs are present in both sequences. In S1tet the octamer related motif 1 AATTGCTT is present at the S1 monomer head to tail fusions 5'TTGC-TTGCATAA-TTGCATTGCATAA in addition to the TTGC-TTGCATAA motif of S1. In oligonucleotide S2 a similar Oct-1 motif is present with a 7/8 match: AAGCTAATTGCATA. In S2tet AAGCTAATTGCATA-AAGCTAATTGCATA the S1 TTGCATAA motif is restored at the head to tail junctions (see table 2). The molecular weight of Oct-1 is 92K. To investigate whether Oct-1 binds to RP3, we used Oct-1 protein in EMSA of RP3. First Oct-1 mRNA was synthesized from a cloned cDNA with T7 polymerase and the RNA was then translated in reticulocyte lysate (64). EMSA of RP3 with the in vitro synthesized Oct-1 protein demonstrates that retarded Oct-1 RP3 complexes are formed (Figure 5, lane 2), and we conclude that the octamer related sequences in RP3 bind Oct-1 protein. This result supports the possibility that p92 is in fact Oct-1. Experiments which are described below show that p92 is distinct from Oct-1.

The intracellular distribution of p92 is cell cycle dependent

The visualization of individual enhancer binding factors enabled us to ask whether they are regulated e.g. during the cell cycle. To study this question, cells were grown in vitro to high saturation density, which leads to cell cycle arrest at the G0/G1 boundary of the cell cycle (46). Tumorigenic Cgl3 cells were grown to confluence, and kept for 48 hours in culture, and were then lysed and nuclear extracts prepared. We found that in nuclear extracts of high density cells p92 DNA binding activity was drastically reduced (Figure 6A, lane 2). This could mean that under these conditions the DNA binding activity may have been lost due to modification, or proteolytic processing. As anti-p92 sera are not yet available, we could not test this possibility. A third possibility to explain loss of DNA binding activity was to assume a change in the intracellular p92 localization, e.g. transport into the cytoplasm. In order to test this possibility cytoplasmic extracts of the same Cgl3 cells, which were used to prepare nuclear extracts shown in Fig 6A, were examined by binding site blotting. In exponentially growing cells p92 is not detectable in the cytoplasm (Fig. 6B, lane 1), whereas at confluence traces of p92 and the 85kD form of p92 appear in the cytoplasm (lane 2), and in high density cells most p92 is found in the cytoplasm (lane 3). At this time p92 is no longer detectable in the nucleus (Fig.6A, lane 2). We have analyzed four additional cervical cancer derived cell lines and human fibroblasts and in all cases, p92 was found in the cytoplasm at high density (Figure 7B). High density culture conditions lead to a similar block in cell cycle as serum starvation (46). Cgl3 cells were grown to half confluence and serum starved in 0.5% fetal calf serum for 24 or 48 hours and cytoplasmic extracts from these time points were analyzed by binding site blotting. P92 is cytoplasmic in Cgl3 cells after serum starvation (Figure 6, lanes 4 and 5). These results show that the intracellular distribution of p92 is cell cycle regulated, and that under growth conditions which lead to cell cycle arrest at the G0/G1 boundary p92 is translocated into the cytoplasm. Most likely, active binding. In order to reproduce this result, we prepared additional nuclear extracts from cervical cancer derived cell lines C33, Caski and ME180. Nuclear extracts of all cells contain p110 which was detected by RP3 in a binding site blot (Figure 4B), with the highest concentration in C33 (Fig. 4B, lane 1). p92 is not detected owing to the conditions under which the cells were grown. Again, p110 was shown to bind S1tet but not the S2tet sequence (Fig. 4C). In conclusion, this result shows that p110 is a sequence specific DNA binding protein, which is present in all cervical cancer derived cell lines we have examined so far.

p92 binds to sequences which are similar to POU factor binding sites

Oligonucleotide RP3 contains two potential POU factor binding sites. Members of the POU homeodomain family of transcriptional activators (30) include Oct-1, the ubiquitous octamer binding protein, which has a molecular weight of 92KD (18, 63). The Oct-1 protein binds to the consensus sequence ATGCAAAT (16, 50). Mutational analysis of an Oct-1 binding site from the immunoglobulin heavy chain promoter revealed three residues which are required for Oct-1 binding ATGCAAAT (or ATTTGCAT in the antisense orientation) (32, 44). Two
Table 2. Octamer related sequences and homeodomain recognition sites in enhancer oligonucleotides which bind nuclear and cytoplasmic p92 protein.

<table>
<thead>
<tr>
<th>Oligonucleotide</th>
<th>Octamer Related Sequence</th>
<th>Homeodomain Recognition Site</th>
</tr>
</thead>
<tbody>
<tr>
<td>RP3</td>
<td>TACAACGGCTTGCATAACATATATCTTGCATAG</td>
<td>octamer related sequence 1: 6/8 match</td>
</tr>
<tr>
<td></td>
<td>TACAACGGCTTGCATAACATATATCTTGCATAG</td>
<td>octamer related sequence 2: 5/8 match</td>
</tr>
<tr>
<td>S1tet</td>
<td>TTGCTTGCATA-ATGCTTGCATA-ATGCTTGCATA-ATGCTTGCATA</td>
<td>Octamer related sequence 1: 6/8 match</td>
</tr>
<tr>
<td></td>
<td>TTGCTTGCATA-ATGCTTGCATA-ATGCTTGCATA-ATGCTTGCATA</td>
<td>octamer related sequence 2: 5/8 match</td>
</tr>
<tr>
<td>RP3/4</td>
<td>TATCCACTCCCATGTAATTTAGGCAATATT</td>
<td>highT affinity homeodomain recognition sequence: 8/9 match*</td>
</tr>
<tr>
<td></td>
<td>TGTAATTTTGAGGCAACATATTTGGTTGGTTTTTA</td>
<td>contains part of the high affinity homeodomain site of RP3/4 and a low affinity homeodomain recognition sequence*</td>
</tr>
<tr>
<td>RP5</td>
<td>ACTTAAGCTAATTGCTA-ATTGCTAATTGCTA-ATTGCTAATTGCTA</td>
<td>octamer related sequence 3: 7/8 match</td>
</tr>
<tr>
<td></td>
<td>ACTTAAGCTAATTGCTA-ATTGCTAATTGCTA-ATTGCTAATTGCTA</td>
<td>octamer related sequence 4: 6/8 match</td>
</tr>
<tr>
<td>S2tet</td>
<td>AAGCTAATTGCTA-ATGCTAATTGCTA-ATGCTAATTGCTA-ATGCTAATTGCTA</td>
<td>octamer related sequence 2: 5/8 match</td>
</tr>
<tr>
<td></td>
<td>AAGCTAATTGCTA-ATGCTAATTGCTA-ATGCTAATTGCTA-ATGCTAATTGCTA</td>
<td>octamer related sequence 3: 7/8 match</td>
</tr>
<tr>
<td></td>
<td>AAGCTAATTGCTA-ATGCTAATTGCTA-ATGCTAATTGCTA-ATGCTAATTGCTA</td>
<td>octamer related sequence 4: 6/8 match</td>
</tr>
</tbody>
</table>

* the high and low affinity homeodomain recognition sites were defined by EMSA and used to form homeodomain-DNA complexes, which were used for crystallography (33)

Transport is required for the observed nucleo-cytoplasmic translocation because nuclear pores are not passable for molecules with a molecular weight larger than 60kD (37). In order to test if this kind of regulation is also operating in non-tumorigenic Hela-fibroblast hybrids, cytoplasmic extracts of 444 cells at high saturation density were examined. In non-tumorigenic 444 cells p92 DNA binding activity appears in the cytoplasm at high saturation density as well (Figure 6C).

**p92 is a novel octamer binding protein**

We suspected that p92 is in fact the ubiquitous Oct-1 protein, which is also involved in mediating S phase dependent gene expression of the human histone H2B gene (36). Regulation of Oct-1 intracellular distribution has, to our knowledge, not been reported and we were interested to demonstrate whether the cytoplasmic p92 is indeed the Oct-1 protein. For this experiment, nuclear extracts of cervical cancer derived cell lines C4I, ME180, SW756, human fibroblasts and C4II cells at high saturation density were analyzed by binding site blotting. In all cell lines p92 is found in the cytoplasm (Fig 7B), whereas very little p92 was detected in the nucleus (not shown). We then used a polyclonal antiserum which was raised against purified Oct-1 protein (49), to examine whether cytoplasmic p92 would be recognized by the antiserum in an immunoblot. Cytoplasmic p92 of C4I, ME180, SW756, Fibroblasts and C4II cells, however, does not react with this polyclonal antiserum (Figure 7A). This result suggests that the cytoplasmic p92 protein is immunologically distinct from the Oct-1 protein. Therefore, we conclude that p92 is a novel factor, which binds to octamer like recognition sequences. To exclude the possibility that the particular batch of antiserum was inactive, and to determine the fate of the Oct-1 protein at the G0/G1 boundary of the cell cycle, we examined nuclear extracts of high density cells by immunoblotting with Oct-1 antiserum. In nuclear extracts of cervical cancer derived cell lines C33, ME180, SW756, C4I and C4II similar amounts of the 92K Oct-1 protein are present in high density nuclear extracts (Fig 8). This result clearly demonstrates that the antiserum was active, that the Oct-1 protein remains in the nucleus under conditions of high saturation density,
and that the cytoplasmic p92 is most likely a novel octamer binding protein whose intracellular distribution is regulated by extracellular growth conditions.

**DISCUSSION**

The search for consensus sequences in the URR’s of HPV18 and HPV16, in combination with the use of purified factors in in vitro analyses and specific antisera, has established interactions with AP1, NF1, SP1 and the glucocorticoid receptor (7, 9, 23, 24). Here we have identified four novel sequence specific enhancer binding proteins by binding site blotting. The recognition sites for p42 and p40 occur once in domain I: in contrast multiple recognition sites for p10 and p92 were identified in this enhancer domain. Multiple sites for the same factors frequently occur in viral enhancers (11, 29), and the redundancy of information in enhancers seems to be a general principle of mammalian transcriptional control. We are currently engaged in efforts to clone p92, p42 and p40 cDNA’s for functional expression in eukaryotic cells. With this approach we intend to determine the role of these proteins with respect to HPV18 early gene regulation. A p110 recognition sequence was mapped to twelve nucleotides, TTGCTTGCATAA, with oligonucleotide S1tet. For p110 binding two GC pairs separated by two T residues are required. A search of similar sequences in oligonucleotides RP2, RP3/4 and RP4, which also bind p110, did not reveal strongly conserved consensus sequences, but in all cases GC rich stretches exist which are likely candidate sequences for p110 binding. Enhancer binding protein p110 may have relaxed DNA binding activities, like Oct-1, AP2 and SP1 (4, 25, 42). We have begun to purify p110, preliminary experiments show that p110 elutes from heparin Sepharose at 500mM KCl as a sharp peak. We have examined a large variety of human tumor samples and found high levels of p110 in all tumors including a retinoblastoma (unpublished data). In binding site blots neither AP1 nor AP2 were detected with oligonucleotides containing AP1 and AP2 recognition sites. This is explained by the fact that AP1 binding depends on heterodimeric interactions with the FOS protein (51), and on the observation that the 52kD AP2 protein cannot be renatured after PAGE (42).

HPV enhancer binding protein p92 may be a novel member of the POU homeodomain family of transcription factors

The homeodomain is a DNA binding motif that plays a central role in eukaryotic gene regulation (57). The amino acid sequences of homeodomains of many transcription factors from different species are highly conserved. Recently the crystal structure of a homeodomain-DNA complex was established (33). Two sites for a homeodomain-DNA interaction were described, with a 100 fold difference in affinity for homeodomain binding (33). The high affinity site ATGTAATT is also present in HPV enhancer oligonucleotide RP3/4 with a 8/9 match ATGTAATTA (table 2).
The central motif for homeodomain binding is TAAT. Oct-1 protein also interacts specifically with this motif (33). We have shown that p92 is immunologically distinct from the Oct-1 protein. p92 binds to octamer related sequences and to the high affinity homeodomain binding site (unpublished observation). Whether p92 protein contains a homeodomain will become evident after sequencing of p92 cDNA clones. Enhancer binding protein p92 did not react with anti-Oct1 antiserum, and one might expect that a putative homeodomain of p92 would have been recognized by the anti-Oct-1 antiserum. The homeodomain of Oct-2, however, does not react in immunoblots with the same antiserum (49), although the two homeodomains of Oct-1 and Oct-2 are over 80% identical (63, 64), indicating that the homeodomain is not recognized by the antiserum. Homeodomain containing transcription factors form a large family of regulatory proteins, some of which are involved in developmental and tissue specific gene expression (15, 28, 44). The ubiquitous Oct-1 protein is also involved in cell cycle regulation of the histone H2B gene (36).

Cell cycle regulation of p92

Regulation of the intracellular distribution of transcription factors is beginning to emerge as a general regulatory principle. For example NFκB is held in the cytoplasm by a cytoplasmic anchoring protein (IkB) and is released to move into the nucleus after ligand binding. The process of nuclear entry of NFκB would be negatively regulated by p92. The presence of at least four recognition sites (TAAT) Oct-1 protein also interacts specifically with this motif (33). We have shown that p92 is immunologically distinct from the Oct-1 protein. p92 binds to octamer related sequences and to the high affinity homeodomain binding site (unpublished observation). Whether p92 protein contains a homeodomain will become evident after sequencing of p92 cDNA clones. Enhancer binding protein p92 did not react with anti-Oct1 antiserum, and one might expect that a putative homeodomain of p92 would have been recognized by the anti-Oct-1 antiserum. The homeodomain of Oct-2, however, does not react in immunoblots with the same antiserum (49), although the two homeodomains of Oct-1 and Oct-2 are over 80% identical (63, 64), indicating that the homeodomain is not recognized by the antiserum. Homeodomain containing transcription factors form a large family of regulatory proteins, some of which are involved in developmental and tissue specific gene expression (15, 28, 44). The ubiquitous Oct-1 protein is also involved in cell cycle regulation of the histone H2B gene (36).

ACKNOWLEDGMENTS

We are grateful to H. zur Hausen for continuous support. We thank T. Rupp for oligonucleotide synthesis, R. Webler for darkroom work. The pBS-Oct1 plasmid was a gift of W. Herr. We thank B. Royer-Pokora, I. Mattaj, R. Di Lauro and K. Kelsey for valuable comments on the manuscript.

REFERENCES