The EMBO Journal vol.7 no.2 pp.533-539, 1988

The carboxy-terminal domain shared by the bovine papillomavirus E2 transactivator and repressor proteins contains a specific DNA binding activity

Alison A.McBride, Richard Schlegel and Peter M.Howley

Laboratory of Tumor Virus Biology, National Cancer Institute, National Institutes of Health, Bethesda, MD 20892, USA

Communicated by B.E.Griffin

The E2 open reading frame of bovine papilloma virus 1 (BPV-1) has been shown to encode both positive and negative acting transcriptional regulatory factors. The DNA binding properties of these factors were analysed to investigate the mechanism by which they might regulate viral gene expression. Polypeptides corresponding to the full-length E2 product and a shorter protein thought to represent the repressor function were synthesized *in vitro* by translation of T7 polymerase generated transcripts. Using rabbit antisera generated against synthetic peptides from the E2 open reading frame, it was possible to immunoprecipitate each of these products and show that each was capable of binding the same specific sequence located at several sites in the BPV-1 genome. This DNA binding property was mapped to a conserved carboxy-terminal domain of 101 amino acids by analysis of truncated polypeptides synthesized from the E2 open reading frame.

Key words: DNA-protein interaction/regulatory elements/*in vitro* transcription-translation

Introduction

Bovine papilloma virus 1 (BPV-1) has served as the prototype for the molecular and genetic analysis of the papillomaviruses, a group of small DNA tumour viruses which induce proliferative squamous epithelial and fibroepithelial lesions. BPV-1 causes cutaneous fibropapillomas in cattle and can readily transform certain rodent cells in tissue culture (Dvoretzky *et al.*, 1980). In the latter case the viral genome can be maintained extrachromasomally and only 'early' viral genes are expressed (Law *et al.*, 1981; Heilman *et al.*, 1982). This latent, non-productive infection of rodent cells is thought to be analogous to that of the basal epithelial cells and dermal fibroblasts of a fibropapilloma. Viral gene expression is tightly regulated in transformed rodent cells, in which the level of transcription is low and restricted to the early transforming region (Heilman *et al.*, 1982), and in a fibropapilloma in which late gene expression occurs only in the terminally differentiating keratinocytes (Amtmann *et al.*, 1982; Engel *et al.*, 1983; Baker and Howley, 1987). The viral and cellular mechanisms involved in the control of these transcriptional regulatory pathways are not yet well understood.

The E2 open reading frame (ORF) of BPV-1 encodes at least two regulatory factors. The full-length E2 ORF encodes a transactivator which increases transcription from the 7940 and 89 promoters by acting at specific sites within the long control region (LCR) (Spalholz *et al.*, 1985, 1987; Haugen *et al.*, 1987). More recently, a transcriptional repressor function, E2-TR, has been identified which inhibits E2 transactivation and appears to act at the same sites. This function has been mapped to the 3' portion of the E2 ORF and can be expressed from the 3080 promoter (Lambert *et al.*, 1987a). Genetic studies indicate that an E2 function is also important for cellular transformation and plasmid maintenance, however, this may be indirect and due to the regulatory functions encoded by this ORF. There is a class of E2 mutants which map to, or downstream from, the splice acceptor at nucleotide 3225 and which cannot be complemented by the full-length E2 ORF (Rabson *et al.*, 1986; Hermonat and Howley, 1987), suggesting that additional functions may be encoded by the E2 ORF.

Recent studies have shown that bacterially generated E2 fusion proteins can bind dsDNA with specificity for a consensus sequence repeated several times in the BPV-1 genome (Androphy *et al.*, 1987; Moskaluk and Bastia, 1987). Several copies of this motif, $ACCN_6GGT$, are present in the LCR of the BPV-1 genome and there are two additional copies

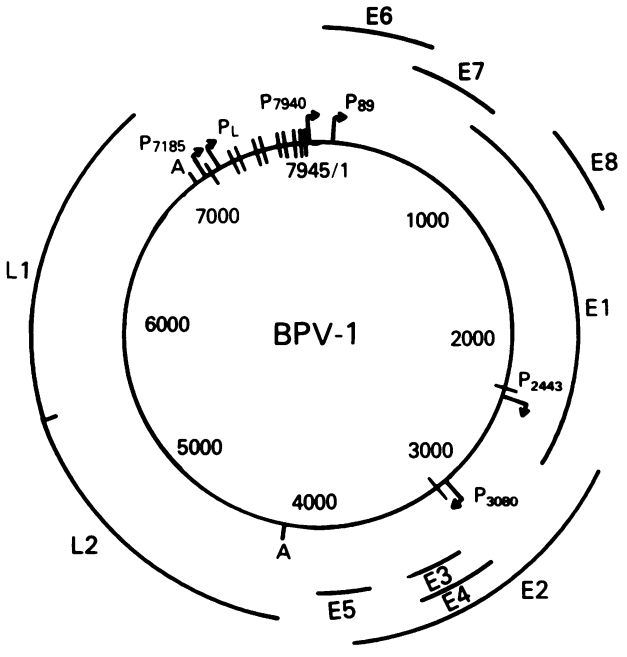

Fig. 1. Genomic organization of BPV-1 DNA. The nucleotide numbers are noted within the circular map and the major open reading frames (designated E1 to E8, L1 and L2) are shown. Promoters are indicated by P followed by the approximate position of the RNA start site except for the late promoter which is designated P_L. The symbol A represents the polyadenylation sites for early and late regions, respectively. The positions of the twelve copies of the motif, $ACCN_6GGT$, are indicated on the circle by the perpendicular lines. Ten of these are clustered in the LCR and there is one located in the vicinity of both the P_{2443} and P_{3080} promoters.

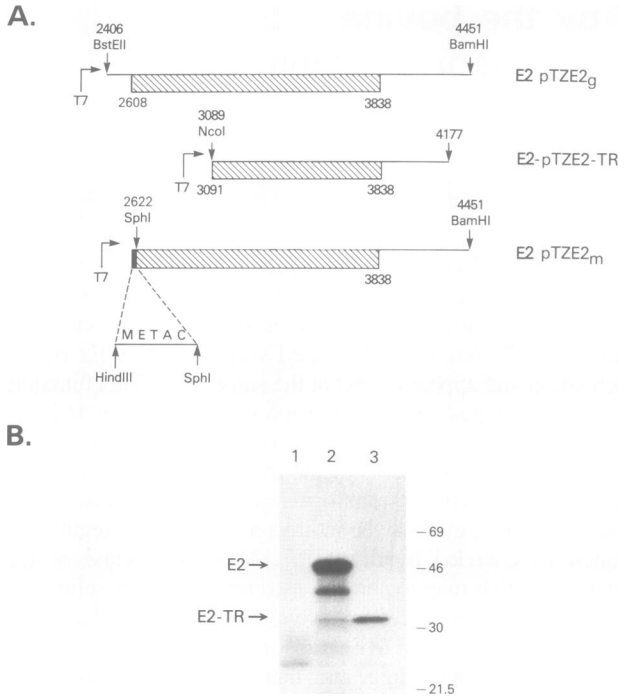

Fig. 2. Diagram of the regions of the E2 ORF cloned in the pTZ expression vectors. Both $pTZE2_g$ and $pTZE2_m$ contain the full-length E2 gene but in the latter, 215 bp of upstream sequence has been removed from the 5′ end and replaced with an oligonucleotide containing an initiation codon in an optimal translation environment. Below is an autoradiograph of *in vitro* translation products generated after *in vitro* transcription of $pTZE2_m$ (**lane 2**), pTZE2-TR (**lane 3**) or in the absence of exogenous RNA (**lane 1**). Samples were analysed on a 12% SDS polyacrylamide gel and compared to protein mol. wt standards, as indicated (units are in kd).

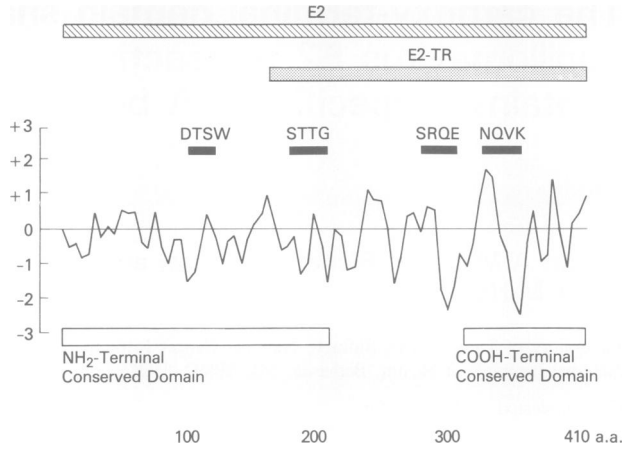

Fig. 3. Hydropathic profile of the protein deduced from the E2 ORF. A frame size of nine amino acids was used for comparing the average hydropathy values of Kyte and Doolittle (1982) and the plot was compressed 5-fold. Hydrophobic and hydrophilic regions are plotted above and below the centre line, respectively. The position of the full-length E2 ORF and the shorter E2-TR polypeptide are indicated above the plot and the amino acid numbers below. Two putative domains which are relatively well conserved among the predicted E2 polypeptides of other papillomaviruses are shown (Giri *et al.*, 1985). The positions from which the synthetic peptides used to generate antisera were derived are also indicated. The peptide DTSW extends from amino acid 96–116, STTG from 181–203, SRQE from 290 to 310 and NQVK from 336 to 355. The letters represent the first four amino acids of each peptide.

adjacent to the 2443 and 3080 promoters. The position of these potential E2 binding sites are shown in Figure 1. Progressive deletion analysis has shown that the presence of these motifs in the E2 responsive element, $E2RE_1$, correlates with its ability to act as an E2-dependent enhancer (Spalholz *et al.*, 1987).

A model by which the transcriptional repressor, E2-TR, could inhibit transactivation could be through direct competition for the DNA binding sites in the E2 responsive elements. In this manuscript we employ the approach used by Hope and Struhl (1985) to demonstrate that *in vitro* translated polypeptides corresponding to both the transactivator and repressor specifically bind to the $ACCN_6GGT$ motif. We map the DNA binding domain to the carboxy-terminal 101 amino acids of the E2 gene product.

Results

Synthesis of E2 proteins and generation of E2 specific antisera

Proteins representing the full-length E2 gene product and the shorter E2 repressor protein were generated by *in vitro* transcription and translation. The coding region for each polypeptide was cloned at a position just downstream from the T7 RNA polymerase promoter in the pTZ series of vectors. Two constructs were used to express the full-length E2 protein. In the first, $pTZE2_g$, a region extending from the *Bst*EII site at nucleotide 2406 to the *Bam*HI site at nucleotide 4451 was inserted behind the T7 RNA polymerase promoter. This construct has 205 nucleotides of additional BPV-1 sequences, containing several AUG codons, upstream from the E2 initiation methionine codon. Therefore, a second plasmid was used from which these upstream sequences had been removed and the initiation codon replaced with one in an environment which would favour translational initiation according to the consensus derived by Kozak (1986a). The latter construct, $pTZE2_m$, was generously provided by Dr Moshe Yaniv. The E2-TR product was expressed from a plasmid, pTZE2-TR, in which a fragment containing nucleotides 3090–4200 of BPV-1 was cloned behind the promoter. Each of these expression plasmids is shown in Figure 2.

Templates were prepared by cleaving the plasmid DNAs downstream from the coding region and capped 'run-off' RNA was synthesized by transcription with T7 RNA polymerase in the presence of the cap analogue diguanosine triphosphate (m^7G-5′ppp5′-G). This RNA was translated in a rabbit reticulocyte lysate to generate E2-derived polypeptides. Figure 2 also shows the protein synthesized from the $pTZE2_m$ and pTZE2-TR templates. Translation of the full-length E2 RNA also resulted in some minor polypeptide products. These are due to initiation at internal methionines, a common property of translation in rabbit reticulocyte lysates (Kozak, 1986b).

The predicted mol. wts of the E2 transactivator and repressor are 45 400 and 26 900 daltons, respectively. However, the apparent mol. wts determined from relative electrophoretic mobility were 50 000 daltons for E2 and 32 000 daltons for E2-TR. This aberrant mobility is a common feature of many proteins, thought to be due to amino acid composition. From the specific activity of the [^{35}S]-methionine and the predicted amino acid sequence of the E2-derived products it could be calculated that ~17 ng of E2 product was synthesized per 100 μl translation reaction.

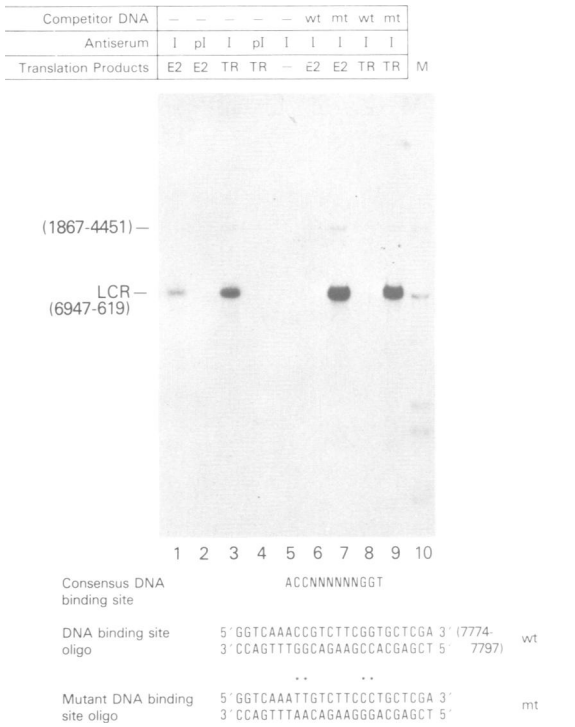

Fig. 4. Immunoprecipitation of E2−DNA complexes. The end-labelled DNA fragments used in this assay were generated by digesting a plasmid, p142-6, containing the full-length BPV-1 genome cloned at the *Bam*HI site of pML2D with *Bam*HI, *Xho*II and *Eag*I (shown in **lane 10**). In **lanes 1** to **5** the standard DNA binding assay was carried out with either E2 (**lanes 1** and **2**), E2-TR (**lanes 3** and **4**) or extract generated without the addition of exogeneous RNA (**lane 5**). DNA−protein complexes were immunoprecipitated with either SRQE antisera (**lanes 1, 3** and **5**) or with pre-immune antisera (**lanes 2** and **4**). Two DNA fragments which contain E2 binding sites are indicated; one of these contains the LCR (nucleotides 6947 to 619) and the other spans nucleotides 1867 to 4451. In **lanes 6** to **9**, 10 μg of synthetic oligonucleotide was added to the assay as competitor. The sequence of the oligonucleotides used are shown below the autoradiograph. The oligonucleotide designated wt represents nucleotides 7774−7797 of the BPV-1 genome which contains an E2 binding site. The second oligonucleotide, mt, contains 4 bp substitutions in the consensus binding motif. The assay was carried out with either E2 (**lanes 6** and **7**) or E2-TR (**lanes 8** and **9**) translation products. wt competitor was added to the experiments shown in **lanes 6** and **8** and mt competitor to those in **lanes 7** and **9**.

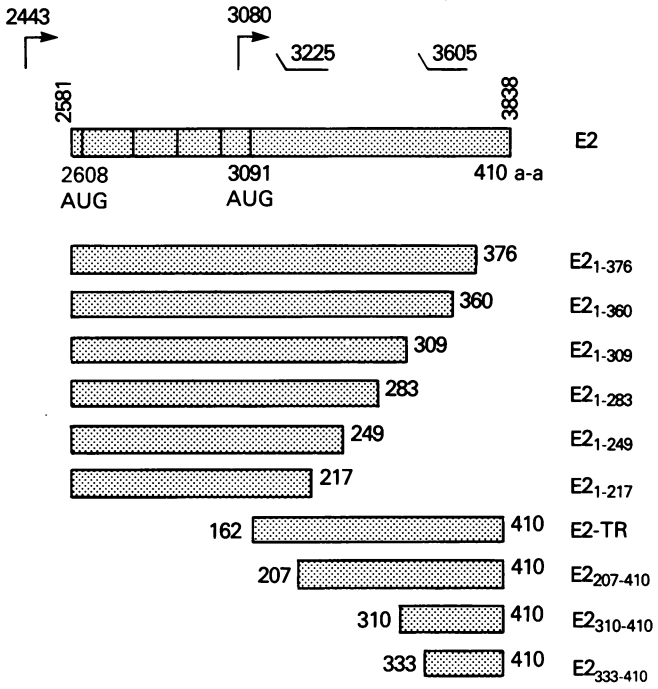

Fig. 5. Diagram of the region of the BPV-1 genome encoding the E2 gene and the position of truncated polypeptides assayed for DNA binding activity. The full-length E2 ORF, which extends from nucleotides 2581 to 3838 of the viral genome, is shown at the top of the figure. The position of AUG codons are represented by vertical bars and those thought to be used for initiation of E2 and E2-TR are indicated below. Two promoters which have been mapped to this region are represented by arrows, and splice acceptor sites, at nucleotide positions 3225 and 3605, are also indicated. The truncated E2 derived polypeptides are shown, each protein designated according to which amino acids from the full-length E2 are represented. The first six polypeptides were expressed from templates which contained translation termination linkers inserted at various positions within the E2 gene. Synthetic oligonucleotides were inserted into the templates to provide initiation codons for the translation of proteins E2$_{207-410}$ and E2$_{310-410}$. E2$_{333-410}$ was expressed from a cDNA which contains two amino acids from the late leader spliced to the acceptor at nucleotide 3605 (Baker and Howley, 1987).

To further characterize these products, antisera was generated against synthetic peptides predicted from the sequence of the E2 ORF (Chen *et al.*, 1982). Four peptides, corresponding to regions predicted to be hydrophilic, were synthesized and used to immunize rabbits. Figure 3 shows a hydropathy profile of the E2 ORF, determined by the methods of Kyte and Doolittle (1982), and indicates the regions to which the synthetic peptides correspond. Antisera generated against peptides designated as DTSW, STTG and SRQE were able to immunoprecipitate efficiently the appropriate translation products, but the antisera designated NQVK could only precipitate a small amount of the E2-derived products.

Sequence-specific binding by E2 and E2-TR in vitro translation products

To determine whether the E2 and E2-TR translation products had sequence-specific DNA binding activity, an immunoprecipitation assay was used which was adapted from that first described by McKay (1981). For this assay the plasmid p142-6, which contains the entire BPV-1 genome, was digested with *Bam*HI, *Xho*II and *Eag*I and the resulting fragments end-labelled. The DNA fragments were incubated with reticulocyte lystate containing either E2 or E2-TR translation products, and the DNA−protein complexes were immunoprecipitated with E2 specific antisera. As shown in Figure 4, two specific BPV-1 fragments were precipitated when incubated with either E2 or E2-TR products but no fragments were precipitated when the lysate contained no translational products, or when pre-immune antiserum was used. The fragment precipitated most efficiently extends from nucleotide 6947 to nucleotide 619 and contains the LCR which includes 10 copies of the consensus E2 DNA binding motif. The second fragment (nucleotides 1867−4451) also contains two consensus E2 DNA binding sites located at positions 2396 and 3088. The position of these motifs in the BPV-1 genome is shown in Figure 1. At very high concentrations of E2 protein a 768-bp fragment of the vector pML2d was also weakly precipitated; this fragment contains a consensus E2 DNA binding motif at nucleotide position 1664. Each of the four E2-specific antisera was able to immunoprecipitate the E2 protein−DNA complexes, indicating that the binding of antibodies to each of these polypeptide

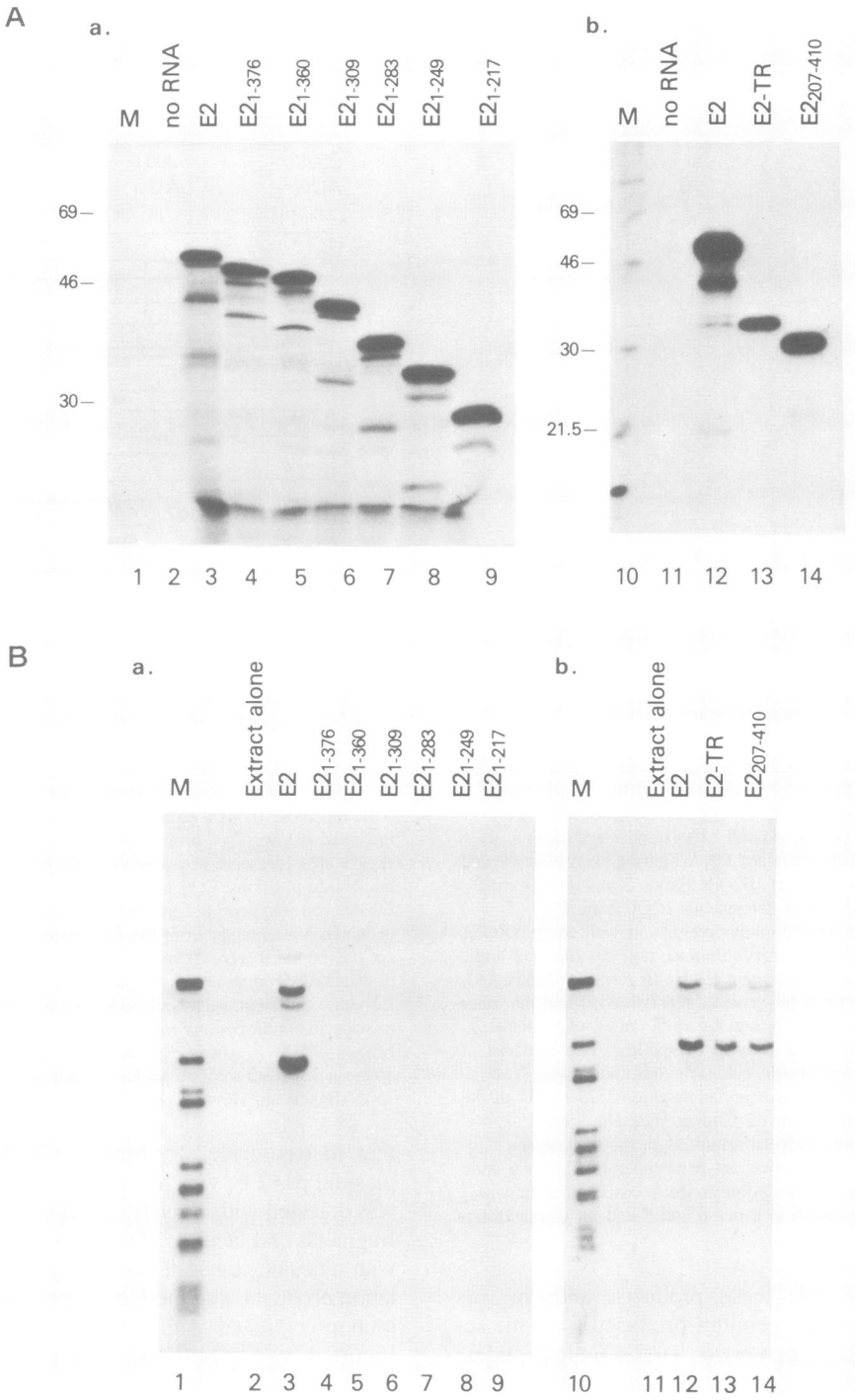

Fig. 6. (**A**) Autoradiograph of the translation products represented in Figure 5. **Lanes 1**, and **10** contain standard mol. wt markers and **lanes 2** and **11** contain translational products generated in the absence of exogenous RNA. The E2-derived polypeptides in **lanes 3–9** and **12–14** are indicated above the autoradiograph. The proteins in **panel a** were separated on a 12% SDS polyacrylamide gel and were not immunoprecipitated. **Panel b** shows a 15% gel and the translational products were immunoprecipitated by the SRQE antisera. (**B**) DNA–protein immunoprecipitation assay using truncated E2 polypeptides. The end-labelled DNA fragments used in this assay, shown in **lanes 1** and **10**, were as described in Figure 4. **Lanes 2** and **11** show DNA recovered in the immunoprecipitation assay using lysate without E2 translation products and all other lanes show DNA precipitated in the presence of the E2-derived polypeptides, as indicated above. The DNA–protein complexes in **panels a** and **b** were immunoprecipitated with DTSW and SRQE antisera, respectively.

regions does not interfere with the DNA–protein interaction.

To establish the specificity of the E2 and E2-TR *in vitro* translated proteins for the E2 binding consensus motif, $ACCN_6GGT$, (Androphy *et al.*, 1987; Moskaluk and Bastia, 1987) a synthetic double-stranded oligonucleotide containing the site was made. This 24 bp oligonucleotide, which contains the BPV-1 DNA sequence from nucleotide 7774 to nucleotide 7797 of the viral genome, was added to the DNA–protein incubation and, as shown in Figure 4 (lanes 6 and 8), completely abolished binding of the DNA fragments to either the E2 or E2-TR proteins. The specificity of this competition for the double-stranded oligonucleotide containing the motif, $ACCN_6GGT$, was determined by using a double-stranded oligonucleotide in which the motif

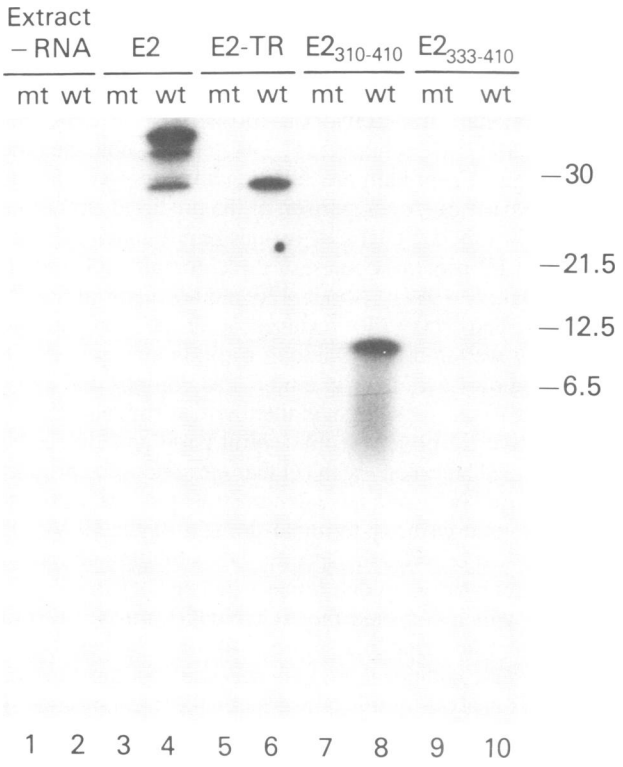

Fig. 7. Oligonucleotides containing a wild type or mutated E2 binding site (as shown in Figure 4) were bound to sepharose and used to isolate E2 derived *in vitro* translation products. The bound polypeptides were eluted and analysed on an 18% SDS polyacrylamide gel. **Lanes 1, 3, 5, 7** and **9** show proteins bound to the mutant oligonucleotide−sepharose and **lanes 2, 4, 6, 8** and **10** to wild type. The translation products were as follows: **lanes 1** and **2**, lysate without the addition of exogenous RNA; **lanes 3** and **4**, E2; **lanes 5** and **6**, E2-TR; **lanes 7** and **8**, $E2_{310-410}$; and **lanes 9** and **10**, $E2_{333-410}$.

was changed to $ATTN_6CCT$. As shown in lanes 7 and 9 of Figure 4, this oligonucleotide was unable to compete for binding of the BPV-1 fragments to either protein. Thus, both E2 and E2-TR polypeptides can bind DNA with the same specificity.

The DNA binding activity of the E2 and E2-TR proteins is encoded by a C-terminal domain

Since both the E2 transactivator and repressor proteins bind DNA in a sequence-specific manner, the DNA binding domain must be encoded in the shared carboxyl-terminal two-thirds of the E2 ORF. To map this domain further, E2-derived polypeptides were synthesized which were truncated at the carboxy-terminus. This was done by inserting a translation termination linker at various restriction sites within the E2 ORF in the $pTZE2_m$ plasmid. Six polypeptides were generated which were truncated at the C-terminus: the smallest encoded amino acids 1−217 of the full-length E2 polypeptide (410 amino acids) and the largest was 376 amino acids long. The region of E2 ORF represented in each truncated polypeptide is shown in Figure 5. The *in vitro* translation products of each of these truncated E2 genes is shown in Figure 6A. As with the full-length E2 and E2-TR, each polypeptide migrates at a slightly higher mol. wt than that predicted from the amino acid sequence. Each of the truncated translation products was incubated with end-labelled BPV-1 DNA fragments and immunoprecipitated with the DTSW antisera. As shown in Figure 6B, none of the truncated polypeptides bound DNA. Even the smallest truncation, removing only 34 amino acids from the C-terminus of E2, eliminated DNA binding. This indicated that either the DNA binding domain was at the C-terminus of the E2 ORF or that the tertiary structure of such a domain had been disrupted.

To map further the DNA binding domain of E2, truncated polypeptides were synthesized which contained intact C-terminal regions of the ORF. Plasmids $pTZE2_{207-410}$ and $pTZE2_{310-410}$ were constructed by inserting a synthetic oligonucleotide encoding an in-frame initiation codon at the *Tth*III1 and *Sty*I sites, respectively. In the case of $pTZE2_{207-410}$, the oligonucleotide also reconstructed the two authentic amino acids downstream of the 3′ splice junction at nucleotide 3225, so that, with the exception of the methionine, this polypeptide represented the domain encoded by the 3′ exon of an mRNA downstream from the splice acceptor at nucleotide 3225. A third polypeptide, $E2_{333-410}$, was synthesized from a cDNA isolated from a fibropapilloma library (Baker and Howley, 1987). This cDNA contains a 5′ exon from the LCR (nucleotides 7217−7385) spliced to the acceptor at nucleotide 3605 and extending to the poly(A) addition site at nucleotide 4203. This cDNA could encode a polypeptide with two amino acids from the 5′ leader and 78 amino acids from the E2 ORF. These *in vitro* translation products are shown in Figure 6A (panel b) and Figure 7.

As shown in panel b of Figure 6B, the $E2_{207-410}$ polypeptide retained the E2-specific DNA binding activity, thus mapping this property to the C-terminal 203 amino acids. However, we were unable to determine the DNA binding properties of the $E2_{310-410}$ and $E2_{333-410}$ polypeptides using the immunoprecipitation assay because of the low affinity of the antisera against the C-terminus (NQVK). To circumvent this problem, oligonucleotides containing either the wild-type or mutated E2 binding site (as described above) were linked to sepharose and used to determine the binding properties of these polypeptides. As shown in Figure 7, the full-length E2, E2-TR and $E2_{310-410}$ polypeptides were bound by oligonucleotides containing the wild-type E2 binding site but not to the mutated site when linked to sepharose. We were not able to detect binding of the smallest polypeptide, $E2_{333-410}$, using this assay, suggesting that the DNA binding domain was not intact in this polypeptide. The interpretation of this negative result must be considered preliminary. However, the data clearly establish that the DNA binding domain of the E2 polypeptide is contained in the C-terminal 101 amino acids.

Discussion

The E2 ORF of BPV-1 encodes both a transcriptional transactivator and repressor which share a common C-terminal domain (Lambert *et al.*, 1987a). It has previously been shown that bacterially generated E2 fusion proteins can bind to specific sequences in the viral genome (Androphy *et al.*, 1987; Moskaluk and Bastia, 1987) and these binding sites appear to be important for E2-dependent transcriptional activation (Spalholz *et al.*, 1987). There are several mechanisms by which such factors, with common domains, could possess antagonistic functions. The shared domain could be important for interaction with cellular factors and, therefore, the repressor could act by titrating out such factors. Alternatively, it is possible that E2 could function as

a multimer in transactivation and thus interactions between E2 and E2-TR could, by subunit mixing, result in inactive complexes. Perhaps the simplest model is that both factors share a DNA binding domain and that repression results from direct competition for the E2 DNA binding sites. We demonstrate here that *in vitro* translated polypeptides corresponding to the full-length E2 transactivator and the shorter repressor protein can both bind specifically to the same consensus DNA sequence, thus supporting the latter model.

The LCR contains E2-responsive elements; $E2RE_1$ has been defined as a 200-bp sequence with two copies of the E2 binding motif at either end. Deletion and mutation analyses have shown that all four copies are required for full enhancer activity and that at least two copies of the motif are necessary for any activity (Spalholz *et al.*, 1987; Spalholz and Howley, in preparation). It has been postulated that the E2 transactivator may function as a multimer in binding to these sites, possibly causing the DNA to loop, either directly through protein−protein interaction or indirectly via other factors (Spalholz *et al.*, 1987). As shown here, the E2 repressor protein contains the domain for binding to the $ACCN_6GGT$ sequences and therefore can bind competitively to the same sites as the E2 transactivator. The repressor, however, lacks the amino-terminal domain of E2 which is necessary for transactivation. This N-terminal domain is highly conserved among all the papillomaviruses sequenced to date (Giri *et al.*, 1985; Baker and Howley, 1987) and thus probably forms a functionally important domain. This domain could be involved in homodimer or multimer formation, or possibly in the interaction with other cellular or viral factors. The precise mechanism by which the transactivator and repressor regulate transcription *in vivo* is unknown at this time but it is possible that their relative levels may control transcription from the LCR promoters. The level of the P_{3080} mRNA, believed to be the E2-TR transcript, is ~10-fold higher than the level of the unspliced P_{2443} mRNA which could represent the full E2 transcript (Baker and Howley, 1987). Thus if these mRNAs were translated with equal efficiencies *in vivo* one could predict that the E2 repressor would be in large excess over the E2 transactivator. Such a ratio could account for the low levels of viral transcription in BPV-1 transformed cells. The presence of E2 binding sites adjacent to the 2443 and 3080 promoters (Figure 1), suggests that these DNA binding factors may somehow also autoregulate their own expression.

To map the domain of the E2 polypeptides which encodes this specific DNA binding property a range of polypeptides were generated, truncated at either the amino or carboxy termini. The removal of 34 amino acids at the C-terminus was sufficient to abolish DNA binding activity. However, proteins which encode amino acids 207−410 or 310−410 still retain the specific DNA binding property. The former polypeptide represents the region of the E2 ORF downstream from the splice acceptor at nucleotide 3225. Genetic evidence suggests that a function important in transformation may be expressed from a message which utilizes this 3′ exon of E2 (Rabson *et al.*, 1986; Hermonat and Howley, 1987). One potential message which would use a 5′ splice donor at nucleotide 1234 and the 3225 nucleotide acceptor would fuse the E8 and E2 ORFs and genetic evidence is consistent with an E8/E2 fusion protein also coding for a repressor function (Lambert *et al.*, 1987b). Experiments are in progress to determine if this mRNA encodes a factor involved in BPV-1 transcriptional regulation.

The E2 binding motif, $ACCN_6GGT$, is found in the genome of each of the papillomaviruses which have been sequenced and the E2 gene products of other papillomaviruses that have been examined have been shown to have transcriptional transactivation properties (Phelps and Howley, 1987; Hirochika *et al.*, 1987). Thus, both the binding site and E2 protein are functionally conserved among papillomaviruses. A comparison of the predicted amino acid sequence of the E2 ORFs of various papillomaviruses suggests that E2 proteins consist of three domains (Giri *et al.*, 1985; Baker, 1987). About 220 amino acids at the N-terminus and 100−120 residues at the C terminus are relatively well conserved among papillomaviruses; the internal region varies both in amino acid composition and in length. It has also been noted that the C-terminal region of the putative E2 protein from several papillomaviruses has some limited homology with cellular *mos* oncogene products (Giri *et al.*, 1985). We have demonstrated in this paper that the conserved carboxy-terminal domain of the BPV-1 E2 gene product encodes the specific DNA binding activity. It will be informative to determine whether C-terminal polypeptides which can bind the $ACCN_6GGT$ motif are sufficient for transcriptional repression or whether E2-TR encodes some additional property necessary for this function.

Further analysis of this transcriptional control mechanism should provide insight not only into papillomavirus gene regulation but into the processes involved in eukaryotic gene expression.

Materials and methods

Plasmid constructions

All BPV-1 fragments were cloned into the pTZ series of plasmids (United States Biochemcial Corporation). For the E2 construct, $pTZE2_g$, p142-6 DNA [which contains the entire BPV-1 genome cloned at the *Bam*HI site of pML2D (Sarver *et al.*, 1982)] was digested with *Bst*EII, the ends made blunt by the Klenow fragment of DNA polymerase I and the site modified by the addition of *Hin*dIII linkers. The *Hin*dIII (*Bst*EII) to *Bam*HI fragment of BPV-1 was then cloned into the multiple cloning site of pTZ19R. The modified E2 construct, $pTZE2_m$, was generously provided by Dr Moshe Yaniv (Pasteur Institue, Paris). This was constructed by cleaving $pTZE2_g$ with *Hin*dIII and *Sph*I and replacing the *Hin*dIII−*Sph*I fragment with a synthetic oligonucleotide of the sequence 5′-AGCTTGGACCATG-GAGACAGCATG-3′. The coding sequence for the E2-TR construct was subcloned from a cDNA clone, p1140 (Lambert *et al.*, 1987a), which contains BPV-1 sequences from nucleotide 845 to the poly(A) addition site at nucleotide 4203, cloned in the Okayama−Berg vector (Okayama and Berg, 1983). Additionally, this cDNA contained an *Hpa*I linker at the *Nco*I site (nucleotide 3089) which enabled an *Hpa*I to *Xho*II (a site present in the vector) fragment to be inserted into the *Sma*I−*Sal*I sites of the pTZ18R vector.

In the C-terminal truncated mutants, a translation termination linker (TTL) with the sequence 5′-TTAGTTAACTAA-3′ (Phelps and Howley, 1987) was inserted at various sites in the E2 ORF of $pTZE2_m$. In the plasmids $pTZE2_{1\text{-}360}$, $pTZE2_{1\text{-}249}$ and $pTZE2_{1\text{-}217}$, the linker was inserted at the *Pfl*MI (nucleotide 3683), *Aat*I, (nucleotide 3353) and *Dra*II (nucleotide 3259) sites, respectively. The linker was also inserted at the *Bcl*I (nucleotide 3737) and *Kpn*I (nucleotide 3460) sites, but in these cases a *Dra*II to *Bam*HI fragment was isolated from cDNAs, p1136 and p1142 (Lambert *et al.*, 1987a), which already contained the linker at the appropriate sites, and exchanged with the analogous fragment in $pTZE2_m$. This gave rise to plasmids $pTZE2_{1\text{-}376}$ (TTL at the *Bcl*I site) and $pTZE2_{1\text{-}283}$ (TTL at the *Kpn*I site). Plasmid $pTZE2_{1\text{-}309}$ contains a TTL at the *Sty*I site at nucleotide 3535 The linker was initially inserted into pTZE2-TR and then transferred to $pTZE2_m$ as described above. To generate the mutants truncated at the amino-terminus of the E2 product, synthetic oligonucleotides were generated which provided an in-frame initiation codon just upstream from the *Tth*III1 and *Sty*I restriction sites. Plasmid $pTZE2_{207\text{-}410}$ was constructed by cleaving $pTZE2_m$ with *Hin*dIII and *Tth*III1 and replacing the released fragment with an oligonucleotide of sequence 5′-AGCTTGGACCATGGATCGCCCAG-ACG-3′. To generate $pTZE2_{310\text{-}410}$, the *Eco*RI (blunted) to *Sty*I fragment

of pTZE2-TR was replaced by an oligonucleotide of sequence 5′-AGCTT-CCACCATGC-3′. $pTZE2_{333-410}$ contains a cDNA isolated from a bovine fibropapilloma library (Baker and Howley, 1987). This cDNA contains a 5′ exon which ranges from nucleotide 7217 to 7385, is spliced to the acceptor at nucleotide 3605 and extends to the poly(A) addition site at nucleotide 4203. A *Pst*I–*Bam*HI fragment containing this cDNA was inserted into the multiple cloning site of pTZ19U.

In vitro transcription and translation

DNA templates were prepared by cleaving the plasmid at an appropriate restriction site downstream from the ORF. Capped RNA was generated by transcription of 10 μg of DNA in a 100 μl reaction containing 40 mM Tris–HCl pH 8.0, 15 mM $MgCl_2$, 0.25 mg/ml bovine serum albumin, 10 mM DTT, 0.5 mM ATP, 0.5 mM UTP, 0.5 mM CTP, 75 nM GTP, 0.5 mM me^7G-5′-ppp5′-G (Pharmacia), 20 units RNasin (Promega Biotec) and 250 units of T7 RNA polymerase (United States Biochemical Corporation). The reaction was incubated at 37°C for 60 min and 10 units of DNase I were added for a further 10 min. After separation from unincorporated nucleotides and cap analogue by chromatography on G50 sephadex, the RNA was extracted with phenol and precipitated by ethanol. This provided enough RNA for several translation reactions. *In vitro* translation was carried out in a reaction volume of 100 μl with 150 μCi of [^{35}S]methionine (>800 Ci/mmol) using rabbit reticulocyte lysate, as described by the manufacturers (Promega Biotec).

Translational products were analysed by SDS-PAGE as described by Laemmli (1970). The gel was fixed, treated with Enlightning (New England Nuclear) and autoradiographed.

Generation of E2 specific antisera

Hydrophilic regions of the E2 protein were determined using the methods of Kyte and Doolittle (1982). Peptides corresponding to four of these regions were synthesized by solid phase methodology at Peninsula Laboratories (San Carlos, California). New Zealand White rabbits were injected with 500 μg synthetic peptide every 3 weeks for a total of three injections. Antigen was dispersed in Complete Freund's adjuvant and injected at four different subcutaneous sites along the flank. Serum was tested for reactivity with synthetic peptide by ELISA as described previously (Schlegel and Wade, 1984).

DNA–protein immunoprecipitation assay

DNA fragments used in the assay were end-labelled with [^{32}P]dNTPs using the Klenow fragment of *Escherichia coli* DNA polymerase after digestion with appropriate restriction enzymes. In the binding assay, 20 ng DNA was mixed with between 25 and 100 μl amounts of translation products in a buffer containing 10 mM Tris–HCl, pH 7.4, 5 mM Mg acetate, 20 mM KCl, 0.1 mM EDTA and 10 μg poly dIdC. After 1 h on ice, 20 μl antisera were added to the assay and incubated for a further 60 min. The complexes were isolated by binding to Protein A sepharose and washed four times in 1 ml NET-100 (50-mM Tris–HCl, pH 7.4, 100 mM NaCl, 5 mM EDTA, 0.1% Triton X-100). The sepharose was suspended in 200 μl NET-100 and 20 μl slurry removed for protein analysis. The complex was dissociated with 1% SDS and after the addition of 20 μg tRNA carrier, DNA was isolated by phenol extraction and ethanol precipitation. Fragments were analysed by electrophoresis on agarose gels which were dried before autoradiography.

DNA affinity binding assay

100 μg of double-stranded oligonucleotides were ligated and bound to sepharose as described previously (Kadonaga and Tjian, 1986). 100 μl (settled volume) DNA sepahrose was incubated with translation products as described above for the immunoprecipitation binding assay and washed three times with NET-100 buffer. Bound proteins were dissociated by heating in SDS-sample buffer and analysed by SDS–PAGE.

Acknowledgements

We would like to thank Moshe Yaniv for the $pTZE2_m$ plasmid and for helpful discussions. We are also grateful to Carl Baker and Paul Lambert for plasmids, discussions and for critical reading of the manuscript; to Steve Harrison and David Levens for helpful discussions; to Janet Byrne for synthesis of the oligonucleotides and to Anne Burkhardt and Marta Wade-Glass for the initial characterization of the antisera.

References

Androphy,E.J., Lowy,D.R. and Schiller,J.T. (1987) *Nature*, **325**, 70–73.
Amtmann,E. and Sauer,G. (1982) *J. Virol.*, **43**, 59–66.
Baker,C.C. and Howley,P.M. (1987) *EMBO J.*, **6**, 1027–1035.
Baker,C.C. (1987) In Salzman,N.D. and Howley,P.M. (eds), *The Papovaviridae 2*. Plenum Press, New York, pp. 321–385.
Chen,E.Y., Howley,P.M., Levenson,A.D. and Seeburg,P.H. (1982) *Nature*, **299**, 529–534.
Dvoretzky,I., Shober,R. and Lowy,D.R. (1980) *Virology*, **103**, 369–375.
Engel,L.W., Heilman,C.A. and Howley,P.M. (1983) *J. Virol.*, **47**, 516–528.
Giri,I., Danos,O. and Yaniv,M. (1985) *Proc. Natl. Acad. Sci. USA*, **82**, 1580–1584.
Haugen,T.H., Cripe,T.P., Ginder,G.D., Karin,M. and Turek,L.P. (1987) *EMBO J.*, **6**, 145–152.
Heilman,C.A., Engel,L., Lowy,D.R. and Howley,P.M. (1982) *Virology*, **119**, 22–34.
Hermonat,P.L. and Howley,P.M. (1987) *J. Virol.*, **61**, 3889–3895.
Hirochika,H., Broker,T.R. and Chow,L.T. (1987) *J. Virol.*, **61**, 2599–2606.
Hope,I.A. and Struhl,K. (1985) *Cell*, **43**, 177–188.
Kadonaga,J.T. and Tjian,R. (1986) *Proc. Natl. Acad. Sci. USA*, **83**, 5889–5893.
Kozak,M. (1986a) *Cell*, **44**, 283–292.
Kozak,M. (1986b) *Cell*, **47**, 481–483.
Kyte,J. and Doolittle,R. (1982) *J. Mol. Biol.*, **157**, 105–132.
Laemmli,U.K. (1970) *Nature*, **227**, 680–685.
Lambert,P.F., Spalholz,B.A. and Howley,P.M. (1987a) *Cell*, **50**, 69–78.
Lambert,P.F., Spalholz,B.A. and Howley,P.M. (1987b) *Cancer Cells*, **5**, 15–22.
Law,M.-F., Lowy,D.R., Dvoretzky,I. and Howley,P.M. (1981) *Proc. Natl. Acad. Sci. USA*, **78**, 2727–2731.
McKay,R.D.J. (1981) *J. Mol. Biol.*, **145**, 471–488.
Moskaluk,C. and Bastia,D. (1987) *Proc. Natl. Acad. Sci. USA*, **84**, 1215–1218.
Okayama,H. and Berg,P. (1983) *Mol. Cell. Biol.*, **3**, 280–289.
Phelps,W.C. and Howley,P.M. (1987) *J. Virol.*, **61**, 1630–1638.
Rabson,M.S., Yee,C., Yang,Y.-C. and Howley,P.M. (1986) *J. Virol.*, **60**, 626–634.
Sarver,N., Byrne,J.C. and Howley,P.M. (1982) *Proc. Natl. Acad. Sci. USA*, **79**, 7147–7151.
Schlegel,R. and Wade,M. (1984) *J. Biol. Chem.*, **259**, 4691–4694.
Spalholz,B.A., Yang,Y.-C. and Howley,P.M. (1985) *Cell*, **42**, 183–191.
Spalholz,B.A., Lambert,P.F., Yee,C.L. and Howley,P.M. (1987) *J. Virol.*, **61**, 2128–2137.

Received on October 5, 1987; revised on November 23, 1987