Mutations in the human papillomavirus type 16 E2 protein identify a region of the protein involved in binding to E1 protein

Alan Storey,* Antonella Piccini, Paola Massimi, Véronique Bouvard and Lawrence Banks

International Centre for Genetic Engineering and Biotechnology, Padriciano 99, I-34012 Trieste, Italy

Papillomavirus DNA replication is primarily dependent upon two viral gene products, E1 and E2. Work with bovine papillomavirus has shown that the E2 protein can bind directly to the E1 protein and enhance the binding of E1 to the viral origin of replication. However, little is known about the mechanism of interaction between E1 and E2 proteins. In this study we have analysed in detail the association between human papillomavirus type 16 (HPV-16) E1 and E2 proteins. Using a purified glutathione S-transferase–HPV-16 E1 fusion protein from Escherichia coli and E2 proteins produced by in vitro transcription–translation, we have developed a rapid and simple method for investigating the association between E1 and E2 in vitro. The binding of E2 to E1 was found to be dependent on sequences in the N-terminal activation domain of the E2 protein. Truncated forms of E2, including a putative repressor form of E2 encoding the DNA binding domain, failed to associate with E1 in this assay. The region of E2 required for efficient binding to E1 was then localized using mutants in the activation domain of E2. These results demonstrated that only a short region of E2 was required for association with E1. This region of E2 was found to be highly conserved amongst all papillomaviruses, suggesting a conservation of E2 function and a common mechanism of interaction between these virally encoded proteins.

Introduction

Human papillomaviruses (HPVs) are epitheliotropic viruses which infect a wide variety of anatomical sites. Of the over 65 different types which have now been identified, a small subset are associated with anogenital lesions (de Villiers, 1989). These anogenital-associated HPV types, in particular types 16, 18, 31, 33, 35 and 45, are linked to the development of cervical intraepithelial neoplasias (CIN) which may progress to malignancy (Dürst et al., 1983; Boshart et al., 1984; van den Brule et al., 1989). These CINs usually harbour many copies of episomal viral DNA. In contrast, invasive carcinomas often possess only integrated HPV sequences. Integration of the HPV DNA into the host genome frequently results in the loss of part of the early region including the E1, E2 and E5 genes. However, the upstream regulatory region (URR) and the two major viral oncogenes, E6 and E7, are always retained and expressed (Schwarz et al., 1985; Androphy et al., 1987a; Smotkin & Wettstein, 1986; Banks et al., 1987).

Studies on the replication of HPVs have until recently been limited by the lack of an in vitro method of virus propagation and the production of viral particles. Transient assays have proved useful, however, enabling the viral origin to be mapped, first for bovine papillomavirus type 1 (BPV-1) and then more recently for other HPVs (Yang et al., 1991; Chiang et al., 1992a; Del Vecchio et al., 1992; Sverdrup & Kahn, 1994). These studies have also demonstrated that viral replication has an absolute requirement for two virally encoded proteins, E1 and E2 (Ustav & Stenlund, 1991; Yang et al., 1991; Chiang et al., 1992b). The E2 protein, as well as E1, is required in its entirety (Knight & Botchan, 1991). The BPV-1 E1 protein is a phosphoprotein of about 68 kDa which binds ATP (Blitz & Laimins, 1991; Sun et al., 1990) and possesses an intrinsic ATPase and helicase activity (Seo et al., 1993). The BPV-1 E1 protein contains a consensus nucleotide binding sequence in the C-terminal region of the protein, and binds to specific sequences in the viral origin of replication (ori) located in the URR (Wilson & Ludes-Meyers, 1991; Yang et al., 1991). Formation of the complex between BPV-1 E1 and E2 can be modulated by phosphorylation of the E2 protein, and mutations within the C terminus of the E1 protein abolish the binding of E2 (Lusky & Fontaine, 1991). More recently, it has also been shown that HPV-11 E1 and E2 proteins expressed in a baculovirus system form heteromeric complexes (Bream et al., 1993),...
and that like BPV-1 E1 the HPV-11 E1 protein possesses an ATPase activity (Bream et al., 1993). The viral ori appears to be composed of a short DNA segment containing an E1 binding site and a short A+T-rich segment located in the vicinity of one or more sequences which are recognized by the E2 protein (Yang et al., 1991; Chiang et al., 1992a; Holt et al., 1994; Svedrup & Kahn, 1994). The BPV-1 E1 protein was shown to form complexes with the virally encoded E2 protein when these proteins were expressed together in baculovirus expression systems (Mohr et al., 1990). In addition, the binding of E1 to its recognition sequence was enhanced in the presence of E2 (Mohr et al., 1990).

The E2 protein binds as a dimer to the viral promoter–enhancer region at ACCGN4CGGT motifs which are found repeated in the viral URR (Androphy et al., 1987b; Lambert et al., 1987; Phelps & Howley, 1987; Gauthier et al., 1991) and this occurs in the absence of added E1. The E2 protein serves as the major trans regulator of viral gene expression. The structure of the full-length E2 protein can be divided into three distinct regions: an N-terminal activation domain; a central hinge region; and a C-terminal domain which mediates both DNA binding and dimerization (Giri & Yaniv, 1988). HPV-11 E2 proteins which lack the activation domain repress replication of the viral DNA in transient assays (Chiang et al., 1991). An HPV-16 E2 protein generated by alternative splicing may function as a repressor of viral gene expression and replication in a similar fashion to truncated forms of the HPV-11 E2 protein (Doorbar et al., 1990).

Although it is apparent that BPV and HPV E1 and E2 proteins form complexes and are interchangeable in DNA replication assays (Del Vecchio et al., 1992), there is as yet no detailed information on the mechanism of this association. Further, the HPV-16 E1 protein is uncharacterized in terms of enzyme activities and ability to complex with E2. Using a bacterial expression system we have been able to express and purify the HPV-16 E1 protein. The purified protein is shown to possess a potent ATPase activity. In addition, using both N- and C-terminally truncated as well as mutant forms of the E2 protein generated by in vitro mutagenesis, we have identified a small region in the activation domain which appears to be required for efficient binding to the HPV-16 E1 protein. This small region of E2 is highly conserved amongst papillomaviruses which suggests a common mechanism of interaction between these proteins.

Methods

Recombinant DNA. The full-length HPV-16 E1 gene from the plasmid pW12 (Storey et al., 1992) was amplified by PCR using Vent polymerase (Biolabs) and cloned between the BamHI and EcoRI sites of the plasmid pGEX2T (Smith & Johnson, 1988) to generate the plasmid pGEX2T16E1 which produces a fusion protein between glutathione S-transferase (GST) and E1. For expression of proteins in vitro each of the HPV genes were individually subcloned into the plasmid pSP64. The full-length E2 sequence was derived from the plasmid pJ4Ω 16E2, and a truncated form of E2 encoding only the activation domain was derived from the plasmid pJ4Ω ACE2 (Storey et al., 1992). The putative spliced repressor form of E2 encoding the DNA binding domain (nucleotides 1264–1301 A 3357–4222) was from the pTZ18uE2C, E5 plasmid described by Doorbar et al. (1990). Both the N-terminal ACE2 and carboxyl spliced repressor form of E2 were cloned into pGEX2T to generate GST fusion proteins. Mutants of E2 were generated by cloning the full-length gene into M13mp19 and mutagenesis reactions were performed using an in vitro system (Amersham). For single point mutations 21 base oligonucleotides were used. Where larger deletions were required, 28-mers were used such that a 14 base anchor was present on each side of the sequence which was to be deleted.

Production and purification of GST fusion proteins. The protease deficient (lon-) bacterial strain KC1101 harbouring either the pGEX2T16E1 or pGEX2T plasmid was grown overnight at 37 °C in 100 ml of LB medium supplemented with 50 μg/ml ampicillin. The next day the culture was added to 700 ml of fresh medium and incubated for 1 h. IPTG was then added to a final concentration of 1 mM and incubation was continued for a further 2 h. Bacteria were harvested by centrifugation in a Sorvall GSA rotor at 4000 r.p.m. for 10 min. The pellet was resuspended in 10 ml of PBS and MgCl2 was added to a final concentration of 15 mM and Triton X-100 to 1%. The bacteria were then lysed by sonication and insoluble material was removed by centrifugation at 10000 r.p.m. Glutathione-agarose resin was then added to the clarified supernatant and then mixed by rotation for 1 h at 4 °C. Unbound proteins were removed by washing the resin three times in PBS containing 1% Tween 20. The purity of the GST–HPV-16 E1 fusion protein (denoted GST–16E1) was then checked by SDS–PAGE, after the gels had been stained with Coomassie blue.

Transcription and translation of proteins in vitro. Proteins were produced in vitro using a coupled transcription–translation system (Promega) as directed by the manufacturer. Radiolabelled proteins were generated by inclusion of [35S]lysine in the reaction mixtures.

Association of E1 and E2 proteins in vitro. Glutathione-agarose beads (25 μl) containing about 2 μg of either the GST or GST–16E1 protein were washed several times in PBS to remove the detergents used in the purification. The volume of beads used was kept constant for each binding reaction by the addition of fresh beads. The GST/GST–16E1 beads were then incubated directly with 2 μl of radiolabelled E2 protein produced in vitro for 1 h at 4 °C. Non-specific proteins were removed by washing the beads three times in PBS containing 1% Tween 20. Proteins specifically bound to the beads were removed by boiling in an equal volume of 2 × SDS sample buffer. Eluted proteins were then electrophoresed on 10% SDS–polyacrylamide gels, fluorographed, dried and exposed to Kodak XAR X-ray film for 20 h.

Reciprocal binding experiments were performed by mixing directly the in vitro translated radiolabelled E1 and E2 proteins. After allowing the proteins to associate by incubation at 4 °C for 1 h, 0.5 μg of a biotinylated oligonucleotide containing two E2 binding sites was added and the incubation was allowed to continue for a further 20 min. As a control, an identical but unbiotinylated oligonucleotide was added in place of the biotinylated oligonucleotide. Proteins bound to the biotinylated oligonucleotide were recovered by the addition of 20 μl of avidin-agarose beads (Sigma). The beads were washed and proteins were eluted as described above. The sequence of the E2 oligonucleotide used was: 5' GCTTCAACCGAATCCGTTGACCGAACCAGGT TGATGG, and its complement, which were hybridized together prior
to incubation with E2. The E2 binding sites are underlined. The oligonucleotide was biotinylated at the 5' end where required.

ATPase assays. The ATPase activity of the purified E1 fusion protein was detected by release of P_i from [γ-32P]ATP. Reactions were performed in a total volume of 20 μl in 50 mM-Tris-HCl pH 8.0, 100 mM-NaCl, 10 mM-MgCl₂, 1 mM-dithiothreitol, 0.1 μCi [γ-32P]ATP. About 1 μg per 10 μl of either the GST-16E1 or the GST-cyclin D proteins were used per assay. After incubation at 37 °C for 1 h, 1 μl was spotted onto a polyethyleneimine thin-layer plate and chromatographed in 0.5 M formic acid-0.5 M-LiCl. The plate was then dried and exposed to Kodak XAR film for 20 min to 1 h. Radioactive spots were then excised and quantified by liquid scintillation counting.

Results

Purification of active GST-HPV-16 E1 fusion protein

The GST-16E1 fusion protein was purified to near homogeneity from bacteria as shown in Fig. 1(a). The GST-16E1 protein migrated on 10% polyacrylamide gels with an apparent molecular mass of 98 kDa, in good agreement with the predicted size of the fusion protein. Under the conditions used a small amount of protein breakdown was usually observed as faster migrating bands. Western blotting showed that these faster migrating products were derived from the GST-16E1 fusion protein as they reacted with anti-GST antibodies (data not shown). Since the GST-16E1 fusion protein was intended for use in protein binding studies, it was necessary to determine if a proportion of the E1 fusion protein was in a native and active conformation. To assess this we employed an ATPase assay as a measure of E1 activity. As can be seen in Fig. 1(b), the GST-16E1 fusion protein possessed a potent ATPase activity in the absence of the E2 protein. No ATPase activity was found to be associated with similar extracts prepared from bacteria over-expressing a GST-human cyclin D protein. These results demonstrate that at least a proportion of the purified E1 fusion protein was folded correctly.

Binding of HPV-16 E2 to GST-16 E1 protein

Having expressed and purified the HPV-16 E1 protein we then proceeded to investigate its ability to associate with E2. The HPV-16 E2 protein was transcribed in vitro and translated in the presence of [35S]cysteine and mixed directly with either beads containing the GST-16E1 fusion protein or with GST alone. The mixtures were then incubated at either room temperature or at 4 °C for 1 h. The beads were then washed and bound proteins were eluted and analysed by SDS-PAGE. The efficiency of binding was assessed by comparing the level of recovered E2 to a sample of the input translation mix. The E2 protein was found to bind strongly to the GST-16E1 fusion protein at 4 °C (Fig. 2), where the majority of the E2 protein was recovered. Since a reduced level of binding, about 25–30% of the input of E2, was observed if the incubation was carried out at room temperature, all further binding experiments were performed at 4 °C.

A faster migrating band, which was probably derived from part of the E2 sequence, was observed in all E2 transcription–translation reactions and also appeared to interact strongly with the E1 protein.
Fig. 2. Binding of E2 to GST-16E1. Radiolabelled full-length E2 protein (4 μl) synthesized in vitro was mixed with either GST or GST-16E1 bound to glutathione–agarose resin. Following incubation at either 4 °C or 20 °C for 1 h the resin was washed and bound proteins were eluted as before. A 1 μl sample of the translation mix (input) was also electrophoresed to assess the recovery of E2.

**E1 and E2 form complexes which bind DNA**

To perform the reciprocal binding experiment we made use of the DNA binding properties of the E2 protein. Radiolabelled E1 and E2 proteins were mixed together as described earlier. After incubation at 4 °C and addition of either biotinylated or unbiotinylated oligonucleotide, the E2 protein bound to the biotinylated oligonucleotide was recovered by the addition of avidin–agarose resin which was gently mixed with the proteins for 20 min. The resin was then washed as before and bound proteins were analysed by SDS–PAGE. Fig. 3 shows that the full-length E2 protein was only recovered if a biotinylated oligonucleotide was used (lanes 2 and 3) and that E1 protein was only detected if both E2 and the biotinylated oligonucleotide were present in the reaction (lane 2). The faster migrating species which was always generated in the E2 transcription–translation reactions was not recovered by this method, suggesting that if this protein was generated from the E2 coding sequence then it did not possess any specific DNA binding capacity. These results demonstrated that E1 and E2 can form protein–protein complexes that do not block the ability of E2 to bind to its cognate recognition sequence.

**Localization of a region of the HPV-16 E2 protein important for E1 binding**

We then proceeded to identify the domain of E2 responsible for binding to E1 by using first N- and C-terminally truncated forms of E2, and then specific E2 mutants generated in vitro. The regions of E2 encoded by these proteins are shown in Fig. 4(a). We first tested the ability of GST fusion proteins either containing the full-length E2, the N-terminal activation domain or the C-terminal DNA binding domain to bind to radiolabelled in vitro translated E1. The relative mobility and purity of these fusion proteins is shown in Fig. 4(b) (top panel). In agreement with the observation that the full-length E2 was shown to be required to support viral replication, we observed no binding of E1 to the short C-terminal repressor form of E2 (GST–E2Ct in Fig. 4b, bottom panel) which encodes the DNA binding domain but which lacks the activation domain. In addition, no binding of E1 to the GST–ACE2 fusion protein (GST–E2Nt) was observed. This construct has only the capacity to encode the first 140 amino acids of the N-terminal activation domain of E2, since a small deletion of 2 base pairs which results in a stop codon is located immediately after the last in-frame initiation codon in the E2 sequence (Storey et al., 1992). No non-specific binding or co-sedimentation of the internal luciferase (LUC) standard to the GST–16E2 protein was observed.

These results appeared to localize the domain of E2 responsible for binding to E1 between residue 140 at the end of the ACE2 protein and residue 202 at the start of the sE2 protein. We then generated mutants of the full-length E2 protein in conserved regions of the activation
Binding of HPV E1 and E2 proteins

Fig. 4. (a) E2 proteins used in this study. The position of the stop codon present in the ACE2 sequence is indicated. Also shown is the putative E1 \( ^{\wedge} \) E2 repressor. This protein encodes a short stretch of amino acids at the N terminus which are derived from the E1 region of the HPV genome and spliced to the C-terminal DNA binding domain of E2. The locations of mutations generated in the full-length E2 protein are also indicated. (b) Binding of \textit{in vitro} translated HPV-16 E1 protein to GST-E2 fusion proteins. Top panel: purification of GST-E2 proteins. Purified fusion proteins were resolved on a 12\% polyacrylamide gel and stained with Coomassie blue. Lane 1, GST fused to full-length E2; lane 2, GST fused to the N-terminal domain of E2 (GST-E2Nt) derived from pJ4\( \Omega \) ACE2; lane 3, GST fused to the E2 C-terminal DNA binding domain (GST-E2Ct) derived from pTZ18uE2C,E5. Bottom panel: \textit{in vitro} translated E1 (2 \( \mu \)l) was incubated with equal amounts of either the full-length GST-E2 or E2 truncated proteins. Binding of E1 was only observed when the full-length E2 protein was fused to GST (lane 1, GST-E2). No binding of E1 to either GST-E2Nt (lane 2) or GST-E2Ct (lane 3), or of the luciferase control (LUC) to GST-E2WT (lane 4) was observed. The input levels of E1 and luciferase are shown in lanes 5 and 6 respectively (2 \( \mu \)l of the respective translation mix).

domain missing in the shorter ACE2 protein on the basis that the binding function might be highly conserved (Fig. 4\( a \)). These mutants were generated in M13mp19 and then subcloned into pSP64 for \textit{in vitro} transcription-translation. All the E2 mutants were translated to the same efficiency as the wild-type (WT) protein (data not shown). Equal amounts of either labelled wild-type or mutant E2 proteins were then mixed with GST–16E1 and assayed for binding. Of the four E2 mutants tested in these assays only one mutant, which had a deletion of amino acids 156–159, completely failed to bind to the GST–16E1. All the other mutants tested bound strongly to GST–16E1 (Fig. 5\( a \)). As before, no non-specific binding of either wild-type or mutant E2 proteins to the GST resin was observed (b).

Fig. 5. Binding of HPV-16 E2 mutants to GST–16E1 resin. All the E2 mutants were translated to the same efficiency and were as stable as the wild-type protein in the translation mixture. Each radiolabelled E2 protein (2 \( \mu \)l) was mixed with glutathione resin which had bound to it either GST–16E1 (a) or GST (b). E2 input (1 \( \mu \)l) is shown in (a). The position of the full-length E2 protein is indicated.

Fig. 6. Alignment of papillomavirus E2 sequences showing conserved residues. The deleted region of HPV-16 E2 implicated in binding to E1 is highlighted.

Although no binding was observed between the GST–16E1 and the sE2 protein, we were interested to determine if heteromeric complexes between the full-length E2 and the sE2 proteins were capable of binding to E1. This is of particular interest since a truncated form of HPV-11 E2 has been shown to inhibit viral DNA replication and a possible mechanism may be by
inhibiting the association of E1 and E2. This would be most likely if E2 bound to E1 as a multimer. We have recently shown that co-translation of sE2 with full-length E2 results in heteromeric E2 complexes which are able to bind DNA (Bouvard et al., 1994). Heteromeric complexes of E2 and sE2 were formed by transcribing and translating the two proteins in the same reaction mixture, and these were then tested for their ability to bind both GST and GST–16E1 as before. The sE2 protein contains only five cysteines as compared to the 11 cysteines present in the full-length E2 protein, and this probably accounted for the fact that it always appeared as a weaker band in the co-translation mixtures. Fig. 7 shows that a proportion of sE2 was retained on the GST–16E1 resin when it had been co-translated with the full-length E2 protein. A densitometer scan of the autoradiogram revealed that the absorbance ratio of E2:sE2 in the input was 1.8 (Fig. 7, lane 6), and 4:1 in the proteins bound to GST–16E1 (Fig. 7, lane 1), suggesting that just under half of the sE2 produced in the in vitro translation mix was bound to full-length E2. As before, no binding of the sE2 alone was observed in the absence of the full-length protein. In addition, no binding of heteromeric E2/sE2 complexes to GST resin was observed (not shown). This observation suggests that heteromeric E2 complexes are able to interact to some degree with E1.

Discussion

We have developed a simple and rapid in vitro method of assaying the association of the HPV-16 E1 and E2 proteins. We have used this to identify an E2 mutant which was completely unable to bind to the E1 protein. Our results suggest that only a small, but highly conserved, region of the E2 activation domain may be required for efficient binding of E2 to E1.

The E1 and E2 proteins are the only virally encoded factors that are required for replication of papillomavirus DNA in vivo, and both of the proteins are required in their entirety (Knight & Botchan, 1991; Ustav & Stenlund, 1991). Consistent with these observations, we did not observe binding of E1 to the spliced repressor form of HPV-16 E2 which encodes only the C-terminal region of the molecule. A small amount of repressor was able to associate with the GST–16E1 fusion protein, but only in the presence of the full-length E2 protein. This was most probably due to the formation of heteromeric complexes between full-length and sE2 proteins (Bouvard et al., 1994) through the dimerization domain, which is located in the C terminus of the E2 molecule. The factors controlling the intracellular levels of each of the two forms of E2 may therefore be important for the control of both viral transcription and replication, functions which are often intimately linked (Eckhart, 1990). It will also be interesting to determine if dimer formation and transactivation functions associated with the full-length E2 protein are formally required for DNA replication since a full-length E2 dimer was required for efficient transactivation (Gauthier et al., 1991), but our results suggest that only one full-length E2 molecule may be required for interaction with E1. Whether such heteromeric E2 complexes are able to stimulate HPV DNA replication in vivo remains to be determined.

We have also demonstrated that HPV-16 E1 possesses an intrinsic ATPase activity and, in addition, that it can associate with the E2 protein when E2 is bound to its cognate recognition sequence in the absence of an E1 binding site. Recent evidence has demonstrated that at low concentrations BPV-1 E1 stimulates E2 mediated transactivation, whereas in the presence of excess E1 transactivation by E2 is repressed (Le Moal et al., 1994). This modulation of E2 transactivation is dependent on the presence of E2 but not E1 DNA binding sites (Le Moal et al., 1994). Our results demonstrating that E1/E2 protein complexes are capable of binding to the E2 recognition sequence imply that a similar mechanism may therefore operate for HPV-16.

The simple assay described here will allow the rapid testing of reagents which may block the interaction between the two proteins, in particular small peptides spanning the region deleted in the E2\textsubscript{A154-158} mutant, since we cannot rule out at this stage that the conformation of this mutant protein is altered in such a way that it cannot bind to E1. Furthermore, we cannot rule out the possibility at this time that other regions of the E2 protein may also contribute to mediating the association between the two proteins.

Fig. 7. Binding of sE2/E2 complexes to GST–16E1. Proteins were either translated separately or together and mixed with GST–16E1 or GST resin. GST–16E1 resin was used in lanes 1–3. The input E2 proteins were: lane 1, 1 µl sE2 plus E2 cotranslated mixture; lane 2, 1 µl sE2 alone; and lane 3, 1 µl E2 alone. Lane 4, GST resin plus 1 µl E2 protein. Lanes 5–7, 1 µl of each translation reaction as follows: lane 5, E2 alone; lane 6, sE2 plus E2; and lane 7, sE2 alone.
It will also be of great interest to test the repressor and E2 mutants for their ability to support HPV replication in transient assays in vivo. This may implicate other regions of the molecule which may contribute towards replication but are not involved directly in mediating the association of E2 and E1. Transfection of wild-type E2 with an excess of the E2A159-159 mutant may block replication if the mutant acted in a dominant negative fashion by forming dimers with the wild-type protein.

Our study suggests that only a very small region of E2 may be involved in mediating the binding to E1. This region of E2 appears to be highly conserved, not only amongst other genital HPVAs including the benign-associated HPV types 6 and 11, but also with BPV-1. If this is the case then it may be possible to design common reagents to block papillomavirus replication. Future studies are aimed at further characterizing the contribution made by E2 to HPV replication in vivo.

We are grateful to Emma Lees for the gift of the pGEX2T-human cyclin D plasmid, and to Carlo Gregory for photographic work. This work was in part supported by a grant from the Association Italiana per la Ricerca sul Cancro.

References


(Received 16 June 1994; Accepted 26 October 1994)