Nucleic Acids Research, 1993, Vol. 21, No. 25 5817–5823

E1 protein of human papillomavirus is a DNA helicase/ ATPase

Fiona J.Hughes and Michael A.Romanos*
Department of Cell Biology, Wellcome Research Laboratories, Langley Park, Beckenham, Kent BR3 3BS, UK

Received October 28, 1993; Revised and Accepted November 24, 1993

ABSTRACT

Replication of human papillomavirus (HPV) DNA requires the viral proteins E1 and E2. Amino acid similarities to SV40 large-T antigen had suggested that E1 is a DNA helicase/ATPase involved in initiating viral DNA replication, and this has recently been shown for bovine papillomavirus type 1 (BPV-1) E1 protein. However, *in vitro* analysis of HPV E1 has been hampered by the inability to produce purified protein using heterologous expression systems. We have succeeded in demonstrating ATPase and DNA helicase activities in purified HPV E1, expressed in *E.coli* as a maltose-binding protein fusion (MBP-E1), for the first time. As further confirmation that the ATPase and DNA helicase activities are due to E1 and not contaminating *E.coli* enzymes, we have shown that a fusion protein containing an amino acid change (E1 Pro-479 to Ser), predicted to inactivate ATP-binding, has impaired activities. We have carried out a structure prediction analysis which suggests that E1 may form two domains: a relatively open N-terminal domain (residues 1–125), and a highly structured C-terminal domain (170–649), with an intermediate region (125–170) predicted to form an inter-domain linker. This is consistent with the proteolytic susceptibility of MBP-E1 at a site 15–20 kD from the N-terminus of E1, and the accumulation of a 58 kD C-terminal fragment of E1. We speculate that the N-terminal domain is involved in DNA-binding, while the C-terminal 58 kD may constitute a distinct enzymatic domain. HPV E1 is of interest as a therapeutic target and the availability of pure enzyme will be invaluable in the search for anti-viral compounds.

INTRODUCTION

Human papillomaviruses (HPVs) are small DNA viruses which infect epithelial cells inducing the formation of benign tumors or papillomas. HPVs cause a range of disease from benign hand warts, to genital warts, and life-threatening laryngeal papillomas. Since there are no effective anti-viral treatments, current therapy relies mainly on removal or destruction of infected tissue and must often be repeated. The HPV types associated with anogenital lesions can be divided into two groups: 'low-risk' HPVs, such as HPV-6b and HPV-11, which are associated with genital warts, and the 'high-risk' HPVs, such as HPV-16 or HPV-18, which are associated with lesions that can progress to malignancy. There is now compelling evidence from both epidemiological and biological studies to suggest that high-risk HPVs are involved in anogenital cancers, particularly cervical cancer (reviewed in 1).

HPVs show extreme tissue specificity for epithelial keratinocytes, and their life-cycle is closely coupled to keratinocyte differentiation. Thus in wart tissue the viral genome is maintained as a low-copy-number episome in the proliferating basal epithelial cells. As the infected cells undergo progressive differentiation, episome copy number and viral gene expression increase, until late gene expression and virion assembly occur in the terminally-differentiated cells at the surface. Due to the inability to propagate HPV in tissue culture, the inability to analyse it genetically, and the lack of suitable HPV animal models, much of our knowledge of HPV molecular biology has come from studies with surrogate viruses, eg. bovine papillomavirus type 1 (BPV-1), or from the study of isolated HPV genes.

In BPV-1, early genetic evidence had indicated that the E1 and E2 genes are essential for episome maintenance (reviewed in 2). Amino acid similarities suggested that E1, like SV40 large-T antigen, may be a DNA helicase/ATPase involved in initiating viral DNA replication (3). Subsequently, BPV-1 E1 protein was shown to be a nuclear 68–72 kD ATP-binding phosphoprotein (4, 5, 6). A mutation in E1 (Pro-434 to Ser), analogous to that in the SV40 large-T antigen mutant *tsA209* that impairs ATP-binding, renders BPV-1 defective for DNA replication and abolishes ATP-binding by E1 (6). More recently, it was established that E1 and E2 are necessary and sufficient for the transient replication of plasmids containing the BPV-1 replication origin (*ori*) in transfected cells (7). The *ori* maps to a 60 bp fragment centered on a unique *HpaI* site in BPV-1 DNA and contains binding sites for E1 and E2 proteins (8, 9). E1 and E2 also bind each other directly, and a major role of E2 in replication may be its ability to strengthen the E1–*ori* interaction (4, 10, 11). Yang *et al.* (12) have demonstrated the replication of *ori*-containing plasmids in cell-free extracts in the presence of BPV-1

* To whom correspondence should be addressed

E1 and E2 proteins. Recently, BPV-1 E1 has been purified and shown to have the predicted ATPase and DNA helicase activities (13, 14).

Until recently little was known about HPV DNA replication. However, transient replication experiments have demonstrated the essential roles of E1 and E2 for a number of HPV types (15, 16, 17). Although HPV and BPV-1 E1 clearly have analogous functions, there are significant sequence differences, particularly in the N-terminal region (18), which would suggest that they are not identical. However, *in vitro* biochemical studies of HPV E1 have been hampered by the difficulty in producing purified protein; pure protein is particularly important in the analysis of ATPase activity, since cellular ATPases are ubiquitous. Bream *et al.* (19) expressed HPV-11 E1 and E2 using the baculovirus system, demonstrated that the unpurified proteins bind to each other, and reported that E1 immunoprecipitated from cell extracts had ATPase activity several-fold above control. In this study we have produced sufficient E1 protein from HPV type 6b, as a fusion to maltose-binding protein (MBP-E1) in *E.coli*, to demonstrate DNA helicase and ATPase activity in purified preparations. As further confirmation that ATPase and DNA helicase activity are not due to contaminating *E.coli* enzymes, we have shown that a fusion protein containing an amino acid change predicted to inactivate ATP-binding has impaired ATPase and DNA helicase activity. Finally, MBP-E1 expressed in *E.coli* shows a susceptibility to cleavage at a site within a region of E1 predicted to be an inter-domain linker, generating a stable 58 kD C-terminal domain fragment.

MATERIALS AND METHODS

Construction of recombinant plasmids

The oligonucleotides ACGTACGGATCCACGATGGCGG-ACGATTCAGGTAC and ACGTGAATTCTTAGAACTCCTC-TAAAGTTCTAACAACTGTTCCTGG were used in a polymerase chain reaction (Perkin-Elmer Cetus kit) to amplify the E1 gene from an HPV-6b genomic clone (20), with a 5′ *Bam*HI site and 3′ *Eco*RI sites. The 3′ primer contained additional codons for the C-terminal amino acids Glu-Glu-Phe, since these bind the tubulin monoclonal antibody YL1/2 and may be used for affinity purification (21). Since the E1 gene contains an internal *Eco*RI site, the 2 kb PCR product was digested with *Bam*HI and *Hind*III to isolate the 5′ 1458 bp fragment and with *Eco*RI and *Hind*III to isolate the 3′ 484 bp fragment. These two fragments were joined with pIC-20H (22), digested with *Bam*HI and *Eco*RI, to generate the plasmid pIC-E1. In order to eliminate the possibility of base changes from PCR, the central 1430 bp *Ppu*MI to *Pac*I fragment was replaced with the equivalent fragment from the original HPV-6b DNA, and the authenticity of the 5′ and 3′ ends of the gene was confirmed by nucleotide sequence analysis.

As a source of a gene encoding E1 protein without the Glu-Glu-Phe tag, the 1980 bp *Fsp*I−*Sca*I fragment from HPV-6b DNA was cloned into the *Sma*I site of pUC18. Recombinant plasmids containing E1 in the same orientation as *lacZ′* were designated pUC18-E1a.

For expression of mature E1 with the Glu-Glu-Phe tag in *E.coli*, under control of the *tac* promoter, the *Bam*HI−*Xho*I fragment from pIC-E1 was ligated between the *Bam*HI and *Sal*I sites of pTrc99A (Pharmacia), to generate pTrc-E1. An expression vector for E1 fused to the C-terminus of glutathione S-transferase (GST), pGEX-E1, was constructed by joining the 1458 bp *Bam*HI−*Hind*III and 484 bp *Hind*III−*Eco*RI fragments from pIC-E1 to pGEX-2T (Pharmacia) digested with *Bam*HI and *Eco*RI.

In contrast to the plasmids described above, expression vectors for E1 containing an N-terminal His_6 affinity tag (pRSET-E1) and E1 fused to the C-terminus of maltose-binding protein (pMAL-E1), did not contain the C-terminal Glu-Glu-Phe tag. pRSET-E1 and pMAL-E1 were constructed by joining the 1677 bp *Bam*HI−*Pac*I fragment from pGEX-E1 and the 306 bp *Pac*I−*Sal*I fragment from pUC18-E1a to *Bam*HI and *Xho*I-digested pRSETA (Invitrogen) or *Bam*HI and *Sal*I-digested pMAL-c2 (New England Biolabs), respectively.

pMAL-E1Δ contained a mutation leading to the amino acid substitution Pro-479 to Ser (numbering from the initiating Met of E1). It was constructed by ligating a pair of synthetic oligonucleotides containing the mutation (CCT to TCT) between the *Eco*RI and *Pfl*MI sites of E1 in pMAL-E1; the plasmid was confirmed by nucleotide sequence analysis.

Induction and purification of E1 proteins

pRSET-E1 was introduced into *E.coli* host strains BL21(DE3) and HMS174(DE3) or the equivalent strains containing pLysS, and induced as described previously (23). pMAL-E1 and pMAL-E1Δ were introduced into the *E.coli* strain MC1061. Cultures were grown in L broth containing carbenicillin (50 μg/ml), and cells from starter cultures were pelletted and resuspended in fresh medium to reduce levels of secreted β-lactamase. 4-litre cultures were grown at 37°C to an A_{650} of 0.8−1.0 then induced at 25°C for 2 h by addition of isopropyl β-D-thiogalactoside (IPTG) to 1 mM. Induced cells were harvested by centrifugation at 4°C and resuspended in 80 ml of ice-cold amylose column buffer (10 mM tris HCl pH 7.4, 200 mM NaCl, 1mM EDTA, 1 mM dithiothreitol, 0.2 mM phenylmethylsulfonyl fluoride, 2 μg/ml leupeptin, 1 μg/ml pepstatin, 2 μg/ml antipain and 2 μg/ml aprotinin). The cells were lysed by three or four rounds of sonication (30 s of 2.5 s pulses with 2.5 s rests) with cooling on ice. Insoluble protein and membranes were removed from the lysate by ultracentrifugation for 1 h at 40,000 rpm in a Beckman 70Ti rotor (100,000 *g*), and the extract was loaded onto a 5-ml amylose resin column (New England Biolabs). The column was washed with 100 ml of column buffer + 5% glycerol + protease inhibitors, and maltose-binding protein (MBP)-fusions eluted with wash buffer containing 10 mM maltose. A 4- litre culture yielded up to 0.9 mg of eluted protein of which typically 25% was full-length MBP-E1.

In order to purify MBP-E1 further, maltose-eluates from a single preparation were concentrated to 200 μl by ultrafiltration using Centricon 30 filters (Amicon). The concentrate was loaded onto a 30-ml Superose 12 column, in a Pharmacia FPLC system, pre-equilibrated with 10 mM sodium phosphate pH 7.4, 0.3 M NaCl at room temperature, and eluates were collected on ice.

Protein analysis

Proteins were analysed by electrophoresis in 7.5% or 10% SDS-polyacrylamide gels followed by staining with Coomassie blue or Western blotting. Soluble proteins could be distinguished from insoluble ones by centrifugation for 15 min at 12000 g at 4°C. For Western blotting, proteins were transferred to Immobilon P membrane (Millipore). The E1 peptide antiserum was raised in rabbits against the HPV-11 E1 peptide TTQPKKVKRRLF-ETRELTDSGYGYS (amino acids 116−140; reference 19); it cross-reacted with HPV-6b E1 protein. Antiserum to MBP was

from New England Biolabs. Detection was with alkaline phosphatase IgG conjugates (Sigma) and BCIP/NBT chromogenic reagents (Sigma).

Factor Xa cleavage of MBP-fusion proteins was carried out according to the manufacturer's instructions (New England Biolabs) at a Factor Xa to protein ratio of 1/50 for 2 to 24 hours at room temperature.

For N-terminal sequence analysis, proteins were transferred to Immobilon P, bands identified by brief staining in 0.005% Coomassie blue and the bands excised. This method was also used to locate stained polypeptide bands during Western blotting. The membranes were destained in methanol and analysed in an Applied Biosystems ABI 477A pulse liquid sequencer using the Normal-1 program.

ATPase and DNA helicase assays

ATPase activity in protein samples was detected by the release of ^{32}P-phosphate from [γ-^{32}P]ATP. Reactions were in 20 μl 50 mM tris HCl pH 8.0, 100 mM NaCl, 10 mM $MgCl_2$, 1 mM dithiothreitol, 0.5 μM ATP and contained 0.1 μCi [γ-^{32}P]ATP (Amersham, 3000 Ci/mmol) each. After incubation at 37°C for 1–2 h, 1 μl was spotted onto a polyethyleneimine (PEI) thin-layer chromatography plate, allowed to dry, and the plate developed with 0.375 M potassium phosphate pH 3.5. Plates were then subjected to autoradiography with Kodak XAR-5 film for 1 h.

DNA helicase activity was detected by the release of labelled oligonucleotide annealed to M13 single-stranded DNA (ssDNA). The substrate was prepared by annealing excess M13-40 sequencing primer (GTTTTCCCAGTCACGAC) to 1 μg M13 mp18 ssDNA, then 3′-end labelling the hybrid by extension with dTTP, dGTP and [α-^{32}P]ATP (Amersham, 3000 Ci/mmol) using Klenow fragment. Excess labelled nucleotide was removed by gel filtration (NICK columns, Pharmacia), and the substrate eluted in 400 μl 10 mM tris HCl pH7.4, 0.1 mM EDTA. Helicase assays were carried out using 10 μl of substrate and up to 30 μl of protein in a final volume of 50 μl containing 20 mM tris HCl pH 7.4, 10 mM $MgCl_2$, 0.5 mM dithiothreitol, 2 mM ATP, 0.1 mg/ml bovine serum albumin, incubated at 37°C for 2–3 h. Reactions were stopped by addition of 5 μl 3.3% SDS and 5 μl 0.5 M EDTA; samples were loaded onto 20% polyacrylamide gels (non-denaturing) which were run at 200 V for 3–4 h, fixed and autoradiographed wet at room temperature. Helicase activity was detected by release of labelled 23-base oligonucleotide.

Computer predictions

Predictions of physico-chemical parameters from the amino acid sequence of E1 were based on published methods (24). Three parameters, location of core or non-core residues, coil prediction by three-state secondary structure methods, and hydrophilicity, are selected for illustration.

RESULTS

Expression of E1 and MBP-E1 fusion proteins in *E.coli*

Preliminary experiments attempting to produce mature HPV-6b E1 protein in *E.coli*, using the vector pTrc-E1, were unsuccessful due to the toxicity of the product. We were able to produce E1 containing an N-terminal His_6 tag using the T7 expression vector pRSET-E1, but the product (approx. 80 kD) was insoluble and could not be extracted in native form for purification (Fig.1A,

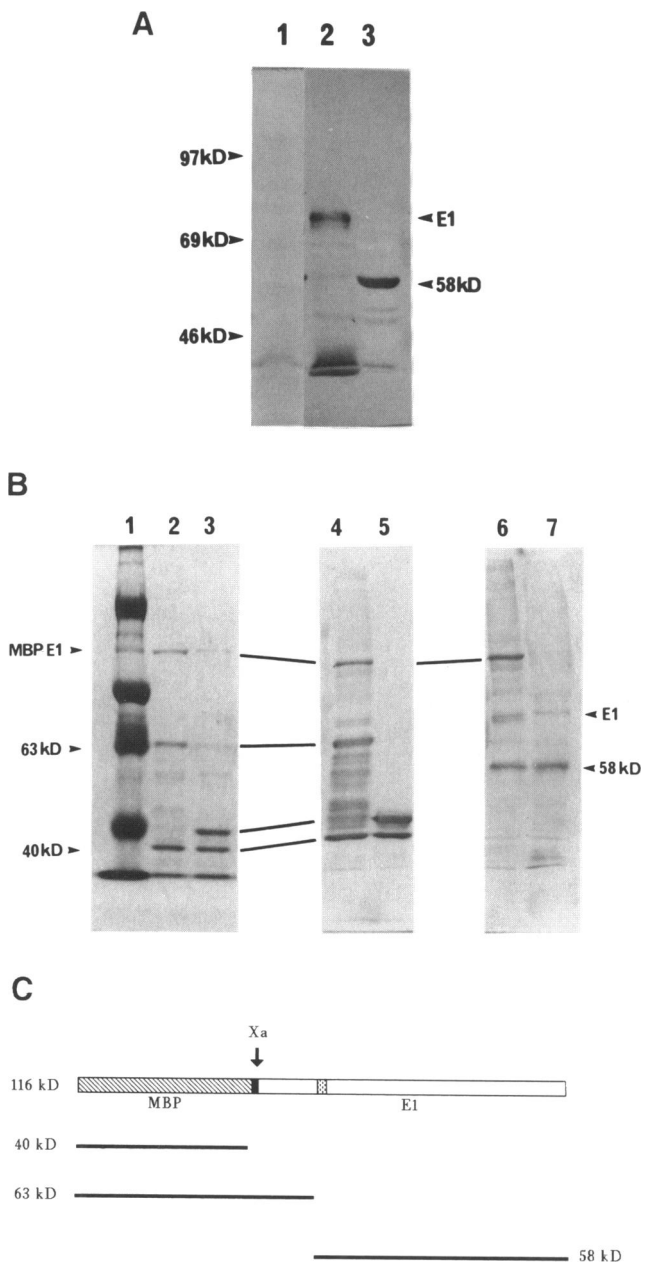

Figure 1. (A) Western blot of T7-induced E1 protein. *E.coli* containing pRSET-E1 was induced, and 50 μg of protein extract analysed by Western blotting using E1 peptide antiserum. Lane 1, induced *E.coli* (BL21(DE3)) lacking pRSET-E1; lane 2, induced BL21(DE3) containing pRSET-E1, insoluble fraction from 50 μg; lane 3, induced BL21(DE3) containing pRSET-E1, soluble fraction from 50 μg. Molecular masses of size markers are indicated on the left. Full-length E1 and an immunoreactive fragment of apparent mass 58 kD are indicated by arrows on the right. (B) Analysis of MBP-E1 purified by amylose resin affinity chromatography as described in Materials and Methods. 35 μl samples (14 μg) were separated on 7.5% SDS-polyacrylamide gels. Lanes 1, 2, and 3, Coomassie-stained gel: (1) size markers (200 kD, 97 kD, 69 kD, and 46 kD; Rainbow markers, Amersham); (2) affinity-purified MBP-E1; (3) affinity-purified MBP-E1 treated with Factor Xa for 24 h. Lanes 4 and 5, and 6 and 7, contained the same samples as lanes 2 and 3, but were analysed by Western blotting using anti-MBP serum and E1 peptide antiserum, respectively. The major products in affinity-purified MBP-E1 are indicated on the left. The minor 58 kD band (detectable only by Western blotting), and mature E1 generated by Factor Xa cleavage are indicated on the right. (C) Major products in affinity-purified MBP-E1. Full-length MBP-E1 is shown as a box, with MBP hatched, the Asn_{10} linker region in black, the region reacting with the peptide antiserum dotted, and the Factor Xa cleavage site indicated.

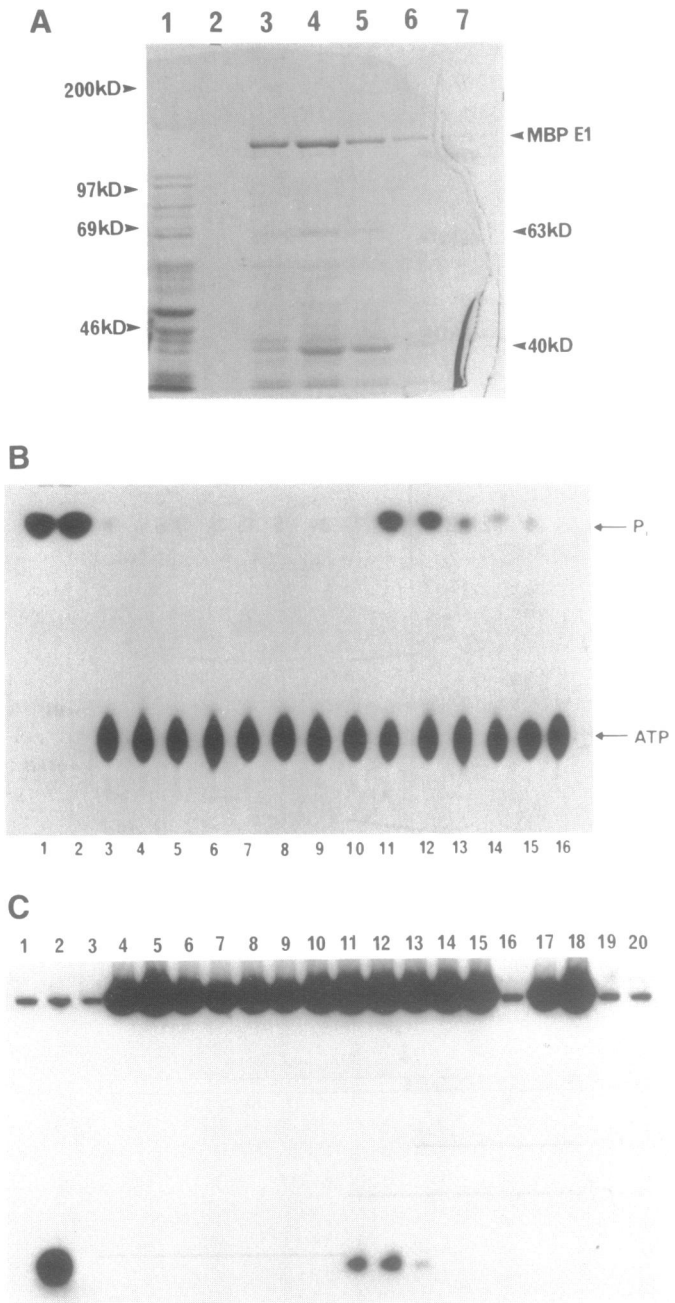

Figure 2. (**A**) SDS-gel electrophoresis of MBP-E1 purified using amylose resin. Soluble protein extract from a 4-litre induction was applied to a 5-ml amylose resin column. The column was washed with 100 ml of column buffer, and elution was with 10 mM maltose in 1-ml fractions. 30 μl of selected fractions were separated on a 7.5% SDS-polyacrylamide gel which was stained with Coomassie blue. Lane 1, soluble protein applied to column (50 μg); lanes 2–7, maltose-eluates (fractions 6–11 , respectively). (**B**) ATPase assay of maltose-eluate fractions. 2 μl of each fraction was used per reaction. Lane 1, flow-through of applied protein extract; lane 2, sample after 10 ml column wash; lane 3, 20 ml column wash; lane 4, 100 ml column wash; lanes 5–16, maltose-eluates (fractions 1–12, respectively). Peak ATPase activity was in fractions 7 and 8 (lanes 11 and 12), corresponding to the peak of MBP-E1. (**C**) DNA helicase assay of maltose-eluate fractions. 30 μl of each fraction was used in the reaction. Helicase activity was detected by release of labeled oligonucleotide using native polyacrylamide gel electrophoresis. Lane 2, positive control (boiled substrate); lanes 4, negative control (no protein); lanes 5–15, maltose-eluates (fractions 1–11, respectively); lane 17, no ATP; lane 18, peak maltose-eluate fraction from purification of MBP-E1Δ. DNA helicase activity is seen in fractions 7 and 8 (lanes 11 and 12), corresponding to the peak of MBP-E1 and of ATPase activity.

lane 2). However, an immunoreactive product of approximately 58 kD was seen in the soluble protein fraction (Fig.1A, lane 3). We were also able to produce intact E1 to levels of 2–6% of cell protein using the baculovirus system but only a very small proportion of this was soluble (data not shown). Finally, E1 could be produced as a fusion protein to glutathione S-transferase (GST-E1) or to maltose-binding protein (MBP-E1) with reduced toxic effects on the host cells; due to a persistent degradation problem with GST-E1, we have concentrated on the purification and biochemical activities of MBP-E1.

The plasmid pMAL-E1 encodes a polypeptide containing MBP (40 kD), followed by an Asn_{10} linker, Factor Xa cleavage site, and mature E1 (73 kD). *E.coli* cells containing the plasmid gave rise upon induction to a polypeptide of the expected size (116 kD) which reacted with both MBP and E1 peptide antisera. Extracts from 4-litre inductions were prepared and subjected to affinity purification using amylose resin. Upon elution with maltose, full-length MBP-E1 (116 kD) was recovered, along with major polypeptides of 63 kD and 40 kD, detectable by Coomassie blue staining in SDS-polyacrylamide gels (Fig.1B, lane 2). The proportions of these proteins varied in different preparations. They all reacted with MBP antiserum in Western blots (Fig.1B, lane 4). MBP-E1 (116 kD) and a minor band of approximately 58 kD reacted with E1 antiserum (Fig.1B, lane 6).

N-terminal sequence analysis of the 116 kD, 63 kD and 40 kD species isolated from gels showed that all had the same N-terminus as MBP (Met-Lys-Ile-Glu-Glu-Gly-Lys-Leu...). This indicates that the 40 kD protein was generated by cleavage at the junction of MBP and E1, and the 63 kD protein by cleavage within E1 (approximately 20 kD from its N-terminus). Maltose-eluates also contained a number of minor bands (40–63 kD) which reacted with antiserum to MBP (Fig.1B, lane 4). We conclude that the N-terminal 20 kD region of E1 contains several sites susceptible to proteolytic cleavage in *E.coli*, and that amylose affinity chromatography separates a nested set of MBP fusion proteins ending in this region, in addition to full-length MBP-E1.

We were not able to isolate sufficient amounts of the 58 kD protein to determine the N-terminus, but the fact that it did not react with MBP antiserum suggested that it is derived entirely from E1. The observation that a 58 kD polypeptide accumulated in *E.coli* cells containing either pRSET-E1 or pMAL-E1 suggests that this represents a stable C-terminal region of E1.

These conclusions, summarised in Fig.1C, were consistent with the results of Factor Xa cleavage. Full-length MBP-E1, the 63 kD fragment, and most of the minor bands were cleaved, while the 58 kD fragment remained intact (Fig.1B, lanes 3, 5, and 7). An MBP band of 43 kD was generated, in addition to the 40 kD MBP band present before digestion. Determination of the exact masses of these species by electrospray mass spectroscopy (data not shown) indicated that the 40 kD protein had been generated by clipping between a Ser residue and the Asn_{10} linker, while the 43 kd protein was the expected Factor Xa cleavage product. Unfortunately, only very small amounts of mature E1 were generated by Factor Xa cleavage (Fig.1B, lane 7).

MBP-E1 has ATPase and DNA helicase activities

Since MBP-E1 purified by amylose resin affinity chromatography was relatively pure, the only detectable contaminants being related degradation products (Fig.2A), we assayed column fractions for ATPase and DNA helicase activities. High-level ATPase activity, representing *E.coli* ATPases, was found in the flow-through and

first wash fractions, but this dropped to undetectable levels in later wash fractions (Fig.2B). Upon elution with maltose, a peak of ATPase activity was seen in fractions 7 and 8, corresponding to the fractions containing MBP-E1 (Fig.2B). These fractions also contained GTPase activity (data not shown).

Fractions from amylose resin chromatography were also assayed for DNA helicase activity (Fig.2C). DNA unwinding was detected by the release of a 3'-^{32}P-labelled 23-mer from M13 mp18 ssDNA which was then separated in non-denaturing gels. DNA helicase activity was seen, in prolonged incubations with $>$ 5 μg of protein, mainly in fractions 7 and 8, corresponding to the MBP-E1 and ATPase peak. No activity was detected if ATP was omitted from the reaction.

MBP-E1 was further purified by FPLC gel filtration using Superose 12 resin: in buffer containing 0.3 M NaCl, two major sharp peaks were obtained (Fig.3A). Peak A contained highly purified, full-length MBP-E1 (Fig.3B). On assaying the Superose 12 fractions, almost all of the ATPase activity co-migrated with MBP-E1 (Fig.3C). We could not detect significant DNA helicase activity in any of the Superose 12 fractions, presumably because the protein concentrations were much lower.

Substitution of E1 Pro-479 with Ser impairs MBP-E1 ATPase

As further confirmation that the MBP-E1 ATPase activity was not due to association with contaminating *E.coli* proteins, we expressed a fusion protein (MBP-E1Δ) containing the amino acid substitution (E1 Pro-479 to Ser) corresponding to the BPV-1 mutation Pro-434 to Ser, which eliminates ATP-binding and renders BPV-1 DNA unable to replicate (Fig.4A). *E.coli* cells harboring the mutated plasmid, pMAL-E1Δ, were induced, extracts prepared and affinity-purified on amylose resin columns. Eluates were found to contain the same pattern of polypeptides as with MBP-E1, that is a full-length 116 kD product and similar degradation products (not shown). MBP-E1Δ had no significant ATPase activity when compared to similar amounts of MBP-E1 under our assay conditions (Fig.4B). The protein also lacked detectable DNA helicase activity (Fig.3C, lane 18).

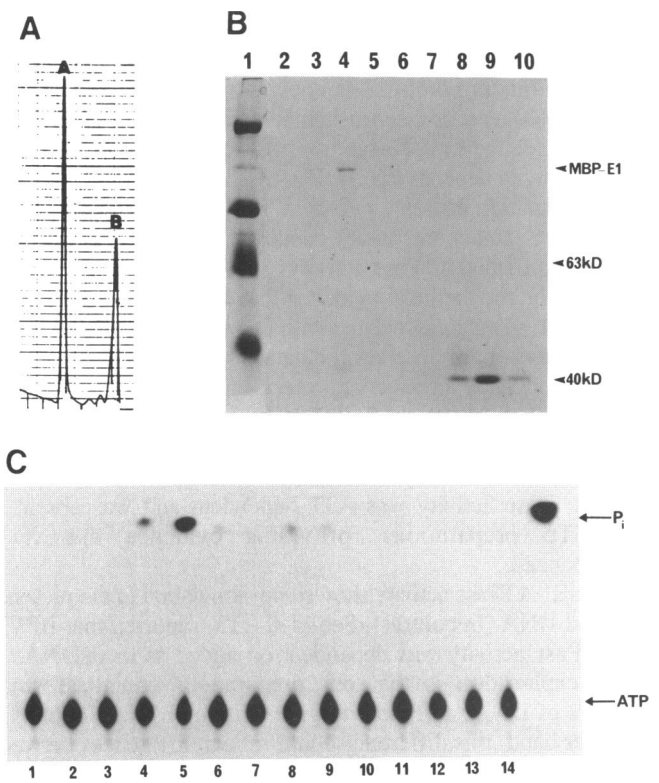

Figure 3. (**A**) Further purification of MBP-E1 by Superose 12 FPLC. A_{260} trace of FPLC fractions showing two major peaks. (**B**) Proteins present in FPLC peaks. Coomassie blue-stained gel of 35 μl samples of each fraction. Lane 1, size markers (200 kD, 97 kD, 69 kD, and 46 kD; Rainbow markers, Amersham); lanes 2–10, fractions 4–12 , respectively. Fraction 6 (lane 4) contains purified full-length MBP-E1 (traces in fractions 5 and 7), while the smaller fragments are in later fractions. The 63 kD band is in fraction 9 (lane 7, not visible in reproduction). (**C**) ATPase assay of FPLC fractions. Lanes 1–13, FPLC fractions 2–14, respectively; lane 14, no protein; lane 15, positive control. Most of the ATPase activity co-purifies with MBP-E1 in fractions 5–7 (lanes 4–6).

Domain structure of E1

Examination of a series of aligned traces for various scaled parameters along the amino acid sequence of E1 revealed a qualitative difference before residue 125 and after residue 170 (Fig.5). The C-terminal part showed a consistent density of core-forming residues (see Fig.5), whereas in the N-terminal part core-forming clusters were separated by appreciable surface segments. The region between these parts (residues 125–170) was predicted to be essentially superficial, with a number of short coil segments, probably indicating a series of exposed loops or turns. This intervening region also had the most sustained high level of hydrophilicity in the entire sequence. Therefore residues 125–170 could contain an exposed inter-domain linker, and the N- and C-terminal parts could form structurally and functionally distinct domains. The putative inter-domain linker may be accessible to proteolytic cleavage, as may be the flanking run of basic residues (120–125, KKVKRR).

The pattern of proteolytic degradation seen with MBP-E1 in *E.coli* is consistent with the suggested model of E1 structure. Cleavage occured throughout the N-terminal 20 kD of E1 (Fig.1),

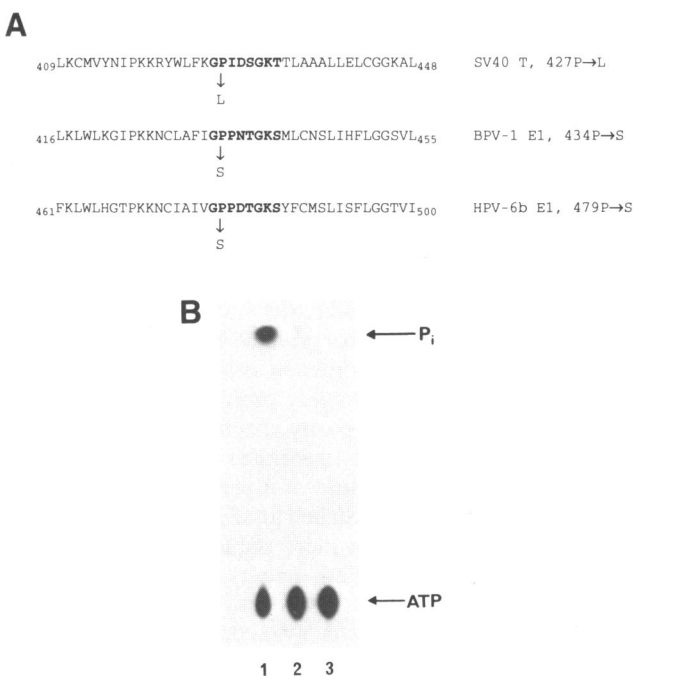

Figure 4. (**A**) Amino acid sequences around the nucleotide-binding motif (shown in bold) in SV40 large-T antigen, BPV-1 E1, and HPV-6b E1, showing homologous mutations, including that made in MBP-E1Δ. (**B**) ATPase assay of MBP-E1Δ. Equivalent amylose resin-purified preparations of MBP-E1 (lane 1) and MBP-E1Δ (lane 2) were assayed; lane 3, no protein.

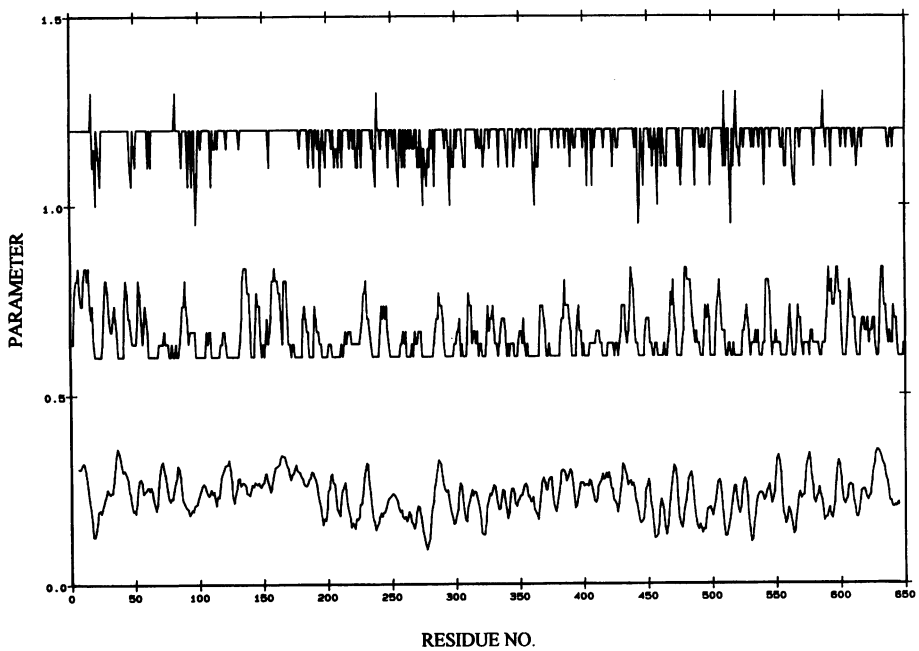

Figure 5. Computer predictions of physico-chemical parameters of HPV-6b E1 polypeptide. Top — core prediction, indicated by a downward spike proportional to the number of rules predicting core location. Middle — coil/turn prediction. Bottom — hydrophilicity.

consistent with the predicted high surface-accessibility of this region. The apparent molecular mass (63 kD) of the major product suggested cleavage near the predicted inter-domain region. The 63 kD fragment did not react with the E1 peptide antiserum (raised to HPV-11 E1 peptide 116–140), indicating cleavage within the peptide sequence (eg. in the basic run, KKVKRR) or N-terminal to it. Thus this product must have a maximum molecular mass of approximately 57 kD, and therefore migrates anomalously during SDS-gel electrophoresis.

DISCUSSION

We have demonstrated ATPase and DNA helicase activities in purified HPV E1 protein for the first time. Production of HPV-6b E1 has presented problems of toxicity to the host cells, insolubility, and instability due to proteolytic degradation *in vivo*. Thus E1 could only be produced using transient expression systems such as the T7 or baculovirus systems. Unfortunately, we were not able to extract these proteins in soluble form under non-denaturing conditions. Expression as a GST fusion in *E.coli* alleviated the toxicity and solubility problems, and indeed we have been able to express and purify enzymatically active GST-E1 fusion protein, and mature E1 generated by thrombin-cleavage (unpublished). However, because of a persistent problem with proteolytic degradation, we switched to MBP-E1 fusion proteins which appear to be somewhat more stable.

We have demonstrated ATPase activity that co-purifies with MBP-E1. MBP-E1, purified by amylose resin affinity chromatography and free of detectable contaminants other than degradation products, had ATPase activity. Almost all of this activity co-purified with full-length MBP-E1 upon further separation by gel filtration. This result is also consistent with the fact that none of the major degradation products found in the maltose-eluates contained large segments of E1. Since cellular ATPases are ubiquitous, we sought additional, genetic evidence that the activity is associated with E1. Therefore we constructed a vector, pMAL-E1Δ, to express a fusion protein (MBP-E1Δ) having the amino acid substition Pro-479 to Ser in the the nucleotide-binding motif. The equivalent mutation (Pro-427 to Leu) in SV40 renders the virus temperature-sensitive, and large-T antigen defective in ATP-binding; ATPase is impaired but the level of activity depends upon the assay conditions (25). In BPV-1 E1, the mutation causes a defect in DNA replication and ATP-binding (6). Under our assay conditions, the altered MBP-E1 had no significant ATPase activity, providing genetic evidence to add to the biochemical evidence that HPV E1 is an ATPase.

MBP-E1 was also associated with DNA helicase activity. This was detected in amylose resin eluates where it was possible to use relatively large amounts of protein, but not in the gel filtration fractions. The absence of detectable helicase activity in gel filtration eluates could have been due to a number of reasons, for example dilution or denaturation of the protein. However, DNA helicase activity was ATP-dependent and was absent in MBP-E1Δ preparations, providing evidence that it is E1-associated.

MBP-E1 ATPase activity showed no stimulation in the presence of added DNA. In contrast, Seo *et al.* (13) reported that BPV-1 E1 ATPase activity was dependent on added ss or dsDNA. A trivial explanation is that our preparations contained small amounts of DNA, sufficient to give maximum stimulation. On the other hand, this difference could reflect differences between BPV-1 and HPV. Alternatively, it could be due to a steric effect of the MBP fusion on E1 DNA binding which might also have a negative effect on DNA helicase activity. Unfortunately, we have not been able produce sufficient mature E1 from MBP-E1 using Factor Xa cleavage in order to test this possibility.

Several years ago significant sequence similarities were noted between 200 amino acids in the C-terminal half of papillomavirus E1 polypeptides and the large-T proteins of polyoma viruses (3). The strongest similarities were found in two blocks in the C-terminal portion of E1, which correspond in SV40 and polyoma virus to sites involved in ATPase and nucleotide-binding

activities. Although E1 is one of the most highly-conserved papillomaviral genes, a comparison of predicted polypeptide sequences from animal and human viruses shows that the N-terminal region (ie. the first 170 amino acids) is relatively divergent (18). This suggests the possibility that N-terminus is involved in non-enzymatic functions of E1 that are less well conserved, such as *ori*-binding. However, the fact that the E1 and E2 proteins of BPV-1 and different HPVs are interchangeable in transient replication assays argues that all E1 functions are qualitatively conserved.

We have carried out a detailed structure prediction analysis of E1 and found that the first 125 amino acids form a relatively open domain with few internal residues. In striking contrast, the C-terminal portion (approx. 450 residues, 58 kD) is predicted to be highly structured and could form a distinct domain. The region between these parts (residues 125−170) is predicted to form a series of exposed loops or turns, and could constitute a linker between the predicted N- and C-terminal domains. These predictions appear to correspond closely with the homology data. The pattern of proteolytic degradation that we see with E1 proteins expressed in *E.coli* appears to support the structural predictions. The N-terminal 20 kD of E1 contains multiple minor cleavage sites and a single major cleavage site near the predicted inter-domain linker. Additionally, using two different systems, we detected stable E1 fragments of the correct size (58 kD) for the predicted C-terminal domain. We are attempting to purify this fragment in order to determine the precise cleavage site and enzymatic properties. It would seem likely that the C-terminal fragment indeed constitutes a functional enzymatic domain.

In summary, we have provided biochemical and genetic evidence that the HPV E1 protein is a DNA helicase/ATPase. The E1 protein of BPV-1 has previously been shown to be multi-functional, with the following activities: DNA helicase/ATPase, *ori*-binding, binding to E2. The interchangeability of HPV and BPV proteins in transient replication assays (26) suggests that HPV E1 will have analogous activities. However, Bream *et al.* (19) have demonstrated E1−E2 binding, but were only able to show E1 binding to the *ori via* E2 protein to E2-binding sites. Thus efficient E1-*ori* interaction may be more dependent on E2 protein with HPV than with BPV-1. Other differences between BPV-1 and HPV may emerge. Our expression system should allow us to increase our understanding of the HPV E1 enzyme activity through further enzymological and genetic studies. The essential role of E1 in HPV replication, its high degree of conservation among different HPV types, and its enzymatic activity make it an attractive target for therapeutic intervention. The availability of suitable amounts of purified HPV enzyme will be invaluable in the search for anti-viral compounds.

ACKNOWLEDGEMENTS

We would like to thank Bill Phelps and Dave Lobe for the E1 antiserum, Chris Beddell for carrying out the structure prediction analysis, Mike Payne for N-terminal sequence analysis, and Peter Lowe and Ken Powell for helpful discussions.

REFERENCES

1. Gissmann,L. (1993) *Semin. Cancer Biol.*, 3, 253−261.
2. Lambert,P. F. (1991) *J. Virol.*, 65, 3417−3420.
3. Clertant,P., and Seif,I. (1984) *Nature*, 311, 276−279.
4. Blitz,I. L., and Laimins,L. (1991) *J. Virol,*. 65, 649−656.
5. Santucci,S., Androphy,E. J., Bonne-Andrea,C., and Clertant,P. (1990) *J. Virol.*, 64, 6027−6039.
6. Sun,S., Thorner,L., Lentz,M., MacPherson,P., and Botchan,M. (1990) *J. Virol.*, 64, 5093−5105.
7. Ustav,M., and Stenlund,A. (1991) *EMBO J.*, 10, 449−458.
8. Ustav,M., Ustav,E., Szymanski,P., and Stenlund,A. (1991) *EMBO J.*, 10, 4321−4329.
9. Wilson,V. G., and Ludes-Meyers,J. (1991) *J. Virol.*, 65, 5314−5322.
10. Mohr,I. J., Clark,R., Sun,S., Androphy,E. J., MacPherson,P., and Botchan,M. R. (1990) *Science*, 250, 1694−1699.
11. Seo,Y.-S., Muller,F., Lusky,M., Gibbs,E., Kim,H.-Y., Phillips,B., and Hurwitz,J. (1993) *Proc. Natl. Acad. Sci. USA*, 90, 2865−2869.
12. Yang,L., Li,R., Mohr,I., Clark,R., and Botchan,M. R. (1991) *Nature*, 353, 628−632.
13. Seo,Y.-S., Muller,F., Lusky,M., and Hurwitz,J. (1993) *Proc. Natl. Acad. Sci. USA*, 90, 702−706.
14. Yang,L., Mohr,I., Fouts,E., Lim,D. A., Nohaile,M., and Botchan,M. (1993) *Proc. Natl. Acad. Sci. USA*, 90, 5086−5090.
15. Chiang,C.-M, Dong,G., Broker,T. R., and Chow,L. T. (1992) *J. Virol.*, 66, 5224−5231.
16. Del Vecchio,A. M., Romanczuk,H., Howley,P. M., and Baker,C. C. (1992) *J. Virol.*, 66, 5949−5958.
17. Remm,M., Brain,R., and Jenkins,J. R. (1992) *Nucleic Acids Res.*, 20, 6015−6021.
18. Danos,O., Giri,I., Thierry,F., and Yaniv,M. (1984) *J. Invest. Dermatol.*, 83, 7s−11s.
19. Bream,G.L., Ohmstede,C.-A., and Phelps,W.C. (1993) *J. Virol.*, 67, 2655−2663.
20. Schwarz,E., Durst,M., Demankowski,C., Lattermann,O., Zech,R., Wolfsperger,E., Suhai,S., and zur Hausen,H. (1983) *EMBO J.*, 2, 2341−2348.
21. Stammers,D.K., Tisdale,M., Court,S., Parmar,V., Bradley,C., and Ross,C.K. (1991) *FEBS Lett.*, 283, 298−302.
22. Marsh,J.L., Erfle,M., and Wykes,E.J. (1984) *Gene*, 32, 481−485.
23. Studier,F.W., Rosenberg,A.H., Dunn,J.J., and Dubendorf,J.W. (1990) In Goeddel,D.V. (ed.), Gene Expression Technology, Methods in Enzymology, vol. 185.. Academic Press Inc., San Diego.
24. Riveros-Moreno,V., Beddell,C., and Moncada,S. (1993) *Eur. J. Biochem.*, 215, 801−808.
25. Ray,S., Anderson,M. E., Loeber,G., McVey,D., and Tegtmeyer,P. (1992) *J. Virol.*, 66, 6509−6516.
26. Chiang,C.-M., Ustav,M., Stenlund,A., Ho,T.F., Broker,T.R., and Chow,L.T. (1992) *Proc. Natl. Acad. Sci. USA*, 89, 5799−5803.