Human papillomaviruses (HPVs) are small DNA tumor viruses that infect stratified epithelium from skin or mucosa. Among the 150 or more HPV types identified so far, there are at least 25 that infect the anogenital tract. Those types are classified either as low risk if they cause only benign lesions, such as warts, or as high risk if they can lead to the development of cancer. Infection by a high-risk HPV type is a necessary cause for the development of cervical cancer, the second most common malignancy among women in the world (42).

The papillomavirus E1 helicase, with the help of E2, assembles at the viral origin into a double hexamer that orchestrates replication of the viral genome. The N-terminal region (NTR) of E1 is essential for DNA replication in vivo but dispensable in vitro, suggesting that it has a regulatory function. By deletion analysis, we identified a conserved region of the E1 NTR needed for efficient replication of viral DNA. This region is predicted to form an amphipathic α-helix (AH) and shows sequence similarity to portions of the p53 and herpes simplex virus (HSV) VP16 transactivation domains known as transactivation domain 2 (TAD2) and VP16C, which fold into α-helices upon binding their target proteins, including the Tfb1/p62 (Saccharomyces cerevisiae/human) subunit of general transcription factor TFIIH. By nuclear magnetic resonance (NMR) spectroscopy and isothermal titration calorimetry (ITC), we found that a peptide spanning the E1 AH binds Tfb1 on the same surface as TAD2/VP16C and with a comparable affinity, suggesting that it does bind as an α-helix. Furthermore, the E1 NTRs from several human papillomavirus (HPV) types could activate transcription in yeast, and to a lesser extent in mammalian cells, when fused to a heterologous DNA-binding domain. Mutation of the three conserved hydrophobic residues in the E1 AH, analogous to those in TAD2/VP16C that directly contact their target proteins, decreased transactivation activity and, importantly, also reduced by 50% the ability of E1 to support transient replication of DNA in C33A cells, at a step following assembly of the E1-E2-ori preinitiation complex. These results demonstrate the existence of a conserved TAD2/VP16C-like AH in E1 that is required for efficient replication of viral DNA.

Received 27 August 2010/Accepted 19 March 2011

* Corresponding author. Mailing address: Institut de Recherches Cliniques de Montréal (IRCM), 110 Pine Avenue West, Montréal, Québec, Canada H2W 1R7. Phone: (514) 987-5739. Fax: (514) 987-5741. E-mail: jacques.archambault@ircm.qc.ca.
† Supplemental material for this article may be found at http://jvi.asm.org/.
‡ Both authors contributed equally to this work.
§ Present address: Roche Research Center, Hoffman-La Roche Inc., Nutley, NJ 07110.
¶ Published ahead of print on 30 March 2011.
support transient replication of DNA and abolished its ability to maintain the viral genome in immortalized keratinocytes (8, 19). Finally, the E1 NTR also contains two overlapping caspase 3/caspase 7 cleavage sites (DxxDxxD), located between the p80-binding region and the shuttling module (Fig. 1A), whose cleavage is required for viral genome amplification in differentiated keratinocytes (37, 38).

A previous study by Demeret et al. (10) suggested that HPV18 E1 can function as a transcriptional transactivator when interacting with E2 or when fused directly to the DNA-binding domain (DBD) of E2. Similarly, we and others observed in the course of yeast two-hybrid studies that E1 can activate transcription when fused to the heterologous DNA-binding domain of Gal4 or LexA. This ability of E1 to activate
transcription in Saccharomyces cerevisiae could be mapped to the NTR, more precisely to sequences located between amino acids 1 and 190 of HPV16 E1 and amino acids 1 and 72 of HPV11 E1 (54, 61). However, the precise domain of E1 responsible for this "transactivation activity" has not been identified, and as a result, its function in DNA replication has never been tested.

Here we performed a systematic deletion analysis of the HPV31 E1 NTR and found that residues 40 to 50, located between the p80-binding site and shuffling module, are essential for efficient replication of DNA. This region of E1 is predicted to form an amphipathic α-helix (AH) and shows sequence similarity to the transactivation domain 2 of the tumor suppressor p53 (p53 TAD2) and to the related transactivi-

transactivation subdomain C of herpes simplex virus protein 16 (HSV VP16C), which both fold into an AH upon binding to their target proteins. These TADs are characterized by their acidic nature and the presence of three hydrophobic residues that make direct contact with their target proteins, such as the Tfb1/p62 (yeast/human) subunit of general transcription factor TFIIH (transcription factor IIH) (12, 30, 59). Accordingly, we determined by nuclear magnetic resonance (NMR) spectroscopy and isothermal titration calorimetry (ITC) that a peptide spanning the E1 AH binds to Tfb1 with a comparable affinity and through the same binding surface as p53 TAD2 and VP16C. Consistent with these results, we found that the E1 NTRs from several HPV types activate transcription in yeast and, to a lesser extent, in mammalian cells when fused to a heterologous DNA-binding domain. Furthermore, this activity was reduced by mutation of the three hydrophobic residues in E1 that correspond to those in p53 TAD2 and VP16C which mediate their interaction with target proteins, including Tfb1/p62. Importantly, we show that mutation of these three key residues in E1 also reduce by half its ability to support transient DNA replication in C33A cervical carcinoma cells, at a step after the assembly of the E1-E2 complex at the origin.

For purification of HPV31 E1(2-332) in bacteria, glutathione and EcoRI restriction sites of pCMV-3Tag-1a (Stratagene) (CMV stands for coding for Flag-tagged E1 (Flag-E1) and truncated derivatives were constructed. HPV11 E1(1-80), HPV16 E1(1-84), and HPV18 E1(1-83), and HPV31 E1(1-83) were inserted in pM between the BamHI and EcoRI sites and the codon-optimized sequence coding for HPV31 E2 TAD (aa 2 to 208) was inserted between the BamHI and HindIII sites. Site-directed mutagenesis was performed using the QuikChange mutagenesis kit (Stratagene). All DNA con-

structures were verified by sequencing. Further details on their constructions will be provided upon request.

Transactivation assays in Saccharomyces cerevisiae. Saccharomyces cerevisiae strain EGY48 (MATa leu2-3,112 his3-11,15 trpl-1 ura3-1 6aXLexAop-LEU2) or BY4741 (MATa his3-1 leu2-3 met5-5 ura3-0) and its derivatives carrying a deletion of ADA1, GCN5, or SPT7 were transformed with the LexA operator-lexA fusion plasmid pH818-34 and a vector encoding the indicated LexA DBD fusion proteins. For each transformation, β-galactosidase activity was determined from three independent samples, and average values are reported with standard deviations. To measure β-galactosidase activity, the transformed cells were pregrown overnight in Synthetic defined (SD) liquid medium lacking uracil and histidine and then used to inoculate new cultures in fresh medium. These cultures were grown at 30°C until they reached an optical density at 600 nm (OD600) of approximately 0.6. The cells were then harvested, washed, and per-

mutation by three cycles of freezing and thawing. β-Galactosidase activity was measured spectrophotometrically (at 578 nm) with the substrate chlorophenol red-β-D-galactopyranoside (CPRG) (Roche) as previously described (7a). Enzym-

atic activity was calculated by the following equation: Miller units = (1,000 × OD578)/(elapsed minutes × V × OD600 where V = 0.1 × concentration factor of the culture. One Miller unit is defined as the amount of β-galactosidase that hydrolyzes 1 μmol of CPRG to chlorophenol red andgalactose per minute per cell.

Transactivation assays in mammalian cells. C33A cells were plated in white flat-bottom 96-well plates at a density of 25,000 cells/well and transfected 20 h later. For each well, 100 ng of pG5luc was transfected along with 25, 50, and 100 ng of vector expressing the indicated Gal4 DBD fusion protein (see Fig. 5) or with pM as a control, and 0.5 ng of pRL as a Renilla luciferase internal control. The culture medium was changed 4 h after transfection. Firefly and Renilla luciferase activities were measured 48 h posttransfection using the Dual-Glo luciferase assay system (Promega).

Western blotting and antibodies. Proteins fused to enhanced yellow fluores-

cent protein (EYFP) were detected using a mixture of two mouse monoclonal antibodies against green fluorescent protein (GFP), purchased from Roche (cat-

alogue no. 1181446001). Proteins fused to the LexA or Gal4 DBD were detected using mouse monoclonal antibodies against Santa Cruz Biotechnology (catalog no. sc-7544 or sc-s10, respectively). Proteins fused to the triple-Flag epitope were detected using the mouse monoclonal antibody Flag M2 from Sigma-Aldrich (catalog no. F1804). Human β-tubulin, yeast β-actin, and human caspase 3 were detected using a mouse monoclonal antibody from Sigma-Aldrich (catalog no. T0426), a mouse monoclonal antibody from Abcam (catalog no. ab8224), and a rabbit polyclonal antibody from Cell Signaling Technology (catalog no. 9626), respectively. The rabbit antisera against the C-terminal domain of p80 was kindly provided by Jae U. Jung (University of Southern California). For Western blot analysis, proteins were transferred onto polyvinylidene difluoride mem-

branes and detected using horseradish peroxidase-conjugated secondary anti-

bodies from GE Healthcare, either sheep anti-mouse IgG (catalog no. NA931) or donkey anti-rabbit IgG (catalog no. NA934V), and an enhanced chemilumi-

nescence detection kit (GE Healthcare).

NMR spectroscopy. A synthetic peptide containing amino acids 44 to 63 of HPV11 E1 [HPV11 E1(44-63)] was purchased from the Sheldon Biotech-

ology Centre (McGill University, Montreal, Canada) and purified to homoge-

neity by reverse-phase high-performance liquid chromatography (HPLC) using a C18 column (Vydac). The amino acid sequence of the peptide is EEEEVDGSY DMDVFIDDRG. Tfb1 (amino acids 1 to 115) was expressed as a GST fusion protein and purified as previously described (13). A sample containing 0.5 mM 15N-labeled Tfb1(1-115) in nuclear magnetic resonance (NMR) buffer (20 mM sodium phosphate buffer [pH 6.5], 1 mM EDTA, and 1 mM dithiothreitol [DTT] in 90% H2O and 10% D2O) was used. The unlabelled HPV11 E1(44-63) peptide was added to a final ratio of 1:1. NMR spectra were collected at 300 K on a Varian Unity Inova 600-MHz NMR spectrometer equipped with a z-pulsed-field gradient unit and a triple resonance probe. For the chemical shift mapping, the changes were determined from both the 1H and 15N chemical shifts (ΔΔ > 0.1 ppm; ΔΔ = [(0.17Δν1H)2 + (Δν15N)2]1/2) as observed in the two-dimensional 2D 1H-15N heteronuclear single-quantum coherence (HMQC) spectra of 15N-labeled Tfb1(1-115) upon binding to a molar equivalent of HPV11 E1(44-63). Figures were generated with PyMol (www.pymol.org).
**RESULTS**

**Deletion analysis of the HPV31 E1 N-terminal region.** The N-terminal region (NTR) of E1 is required for viral DNA replication in *vitro* but not *in vivo*, suggesting that it has a regulatory function. The E1 NTR contains a binding site for the cellular protein p80 and a “shuttling module” that mediates the nuclear import and export of E1 (Fig. 1A) (8, 11, 18, 32, 63). Both domains are separated by a short region containing two overlapping caspase 3/caspase 7 cleavage sites (Fig. 1A) that are required for the differentiation-dependent amplification of the viral genome (37). To examine in more detail the role of the E1 NTR in viral DNA replication, we constructed a set of N-terminal deletions in HPV31 E1 at 10-amino-acid intervals (C10, C20, C30, and C40). We also created two deletions that mimic the form of E1 cleaved by caspase 3/caspase 7 at either of the two overlapping caspase sites (C50 and C53).

Further truncations into the shuttling module were not created, as those would be expected to affect the replication activity of E1 by preventing its nuclear import. Each E1 truncation was tested using our recently described luciferase-based transient DNA replication assay (19). This assay is based on a plasmid containing the viral origin of replication (ori) that also carries a firefly luciferase reporter gene in *cis* of the ori, such that replication of this plasmid by E1 and E2 results in increased expression of firefly luciferase (19). Each fragment of E1 was tested at three different amounts (1.25, 2.5, and 5 ng of Flag-tagged E1 [Flag-E1] expression plasmid) in duplicate experiments, and luciferase activities were determined 72 h post-transfection. As shown in Fig. 1B, truncation C10 was as active as wild-type E1. In contrast, truncations C20, C30, and C40 reduced viral DNA replication by 50 to 70%. These truncations remove part or all of the p80-binding site (aa 10 to 40) (8). Interestingly, the effects of these truncations are very similar to the effects of specific amino acid substitutions in E1 that prevent p80 binding. Indeed, we previously reported that three 2-amino-acid substitutions in E1 that prevent p80 binding also reduce viral DNA replication by 50 to 70% (8, 19). Interestingly, the two larger deletions C50 and C53, which mimic the forms of E1 resulting from caspase cleavage at the first or second site, were both inactive (Fig. 1B), thus indicating that sequences located between amino acids 40 and 50 play an important role in viral DNA replication. For a control, we verified that these E1 truncations were expressed at comparable levels by Western blotting against their triple-Flag epitope and thus that their replication defects were not due to a reduction in protein expression or stability (Fig. 1C).

The findings that C40 retains some replication activity whereas C50 is completely inactive suggested that the region located between amino acids 40 and 50 plays an important role in viral DNA replication, independently of the p80-binding site (which is completely lacking in C40). Although this region contains the two overlapping caspase 3/caspase 7 cleavage sites, it is unlikely that caspase cleavage of E1 plays a significant role in our assay, since little to no cleavage of E1 is detected in the C33A cells used for the DNA replication assay (37) (see Fig. S1 in the supplemental material). However, to conclusively rule out an effect of caspases in our DNA replication assay, we mutated the three aspartates of the two overlapping (D46xxD49xxD52) caspase cleavage sites and confirmed that the resulting three E1 mutants were as active as wild-type E1. In contrast, truncations C20, C30, and C40 reduced viral DNA replication by 50 to 70%. These truncations remove part or all of the p80-binding site (aa 10 to 40) (8). Interestingly, the effects of these truncations are very similar to the effects of specific amino acid substitutions in E1 that prevent p80 binding. Indeed, we previously reported that three 2-amino-acid substitutions in E1 that prevent p80 binding also reduce viral DNA replication by 50 to 70% (8, 19). Interestingly, the two larger deletions C50 and C53, which mimic the forms of E1 resulting from caspase cleavage at the first or second site, were both inactive (Fig. 1B), thus indicating that sequences located between amino acids 40 and 50 play an important role in viral DNA replication. For a control, we verified that these E1 truncations were expressed at comparable levels by Western blotting against their triple-Flag epitope and thus that their replication defects were not due to a reduction in protein expression or stability (Fig. 1C).

The findings that C40 retains some replication activity whereas C50 is completely inactive suggested that the region located between amino acids 40 and 50 plays an important role in viral DNA replication, independently of the p80-binding site (which is completely lacking in C40). Although this region contains the two overlapping caspase 3/caspase 7 cleavage sites, it is unlikely that caspase cleavage of E1 plays a significant role in our assay, since little to no cleavage of E1 is detected in the C33A cells used for the DNA replication assay (37) (see Fig. S1 in the supplemental material). However, to conclusively rule out an effect of caspases in our DNA replication assay, we mutated the three aspartates of the two overlapping (D46xxD49xxD52) caspase cleavage sites and confirmed that the resulting three E1 mutants were as active as the wild-type protein (see Fig. S2 in the supplemental material).

Collectively, the results presented above indicate that residues 40 to 50 of HPV31 E1 contribute to viral DNA replication independently of the p80-binding domain and caspase cleavage sites. As such, they also raise the possibility that residues 40 to 50 of HPV31 E1 encode a previously unrecognized activity required for efficient replication of viral DNA.
Residues 40 to 50 of HPV31 E1 span a predicted amphipathic α-helix with sequence similarity to the p53 TAD2 and VP16C transactivation domains. As a first step toward understanding the role of amino acids 40 to 50 of E1 in HPV DNA replication, we analyzed their potential secondary structure in silico. Several bioinformatics analyses of the E1 NTR predicted that amino acids 40 to 50 lie within an intrinsically disordered domain (IDD). Figure 2A shows the predictions obtained with three different protein disorder prediction programs, MetaPrDos (http://prdos.hgc.jp/meta) (27) and the two GeneSilico programs metadisorder binCons and floatCons (http://iimcb.genesilico.pl/metadisorder) (29), which use several individual analyses to make a consensus prediction. Each program provides a score of disorder probability between 0 and 1 for every amino acid of the protein; a region is predicted to be disordered if the scores are higher than 0.5. All three programs suggested a highly unstructured region located between amino acids 27 and 48 of HPV31 E1 (Fig. 2A).

A closer examination of the interval of E1 from amino acids 40 to 50 showed that it has sequence similarity to two other and well-characterized IDDs, namely, to the transactivation domains of p53 and VP16 (Fig. 2B). p53 contains two separable transactivation subdomains in its N terminus, located between residues 1 to 40 (TAD1) and residues 40 to 73 (TAD2) (7, 48).
The VP16 TAD also contains two similar autonomous subdomains, named VP16N and VP16C (28). Both the p53 TAD2 and VP16C are largely unstructured in solution but fold into a 9-residue α-helix upon binding to their target proteins (Fig. 2C) (5, 12, 28, 56). These TADs contain many acidic residues surrounding three critical hydrophobic residues that make direct contacts with target proteins and that are essential for transactivation activity. Interestingly, these three hydrophobic residues (M47, F50, and I51 for HPV31 E1) are highly conserved among E1 from different HPV types (Fig. 2B, residues on a black background). Furthermore, the short region containing these conserved residues (47MVDFI51; referred herein as the MxxFI motif) has a lower disorder score (around 0.35) and is predicted by different secondary structure programs, including the Porter program (http://distill.ucd.ie/porter/) (44), to lie within a short α-helix (boxed region in Fig. 2A). Interestingly, the Porter program also predicts helices for the analogous regions of p53 TAD2 and VP16C (data not shown). Finally, the p53 TAD2 also contains a phosphorylation site (S46) that enhances its transactivation activity by increasing its binding to target proteins (6, 12, 41). A potential CK2 phosphorylation site is present and conserved at a similar position in E1 (T43 for HPV31 E1, Fig. 2B, boxed) but is absent in VP16C. Collectively, the bioinformatics predictions described above suggest that the role of E1 amino acids 40 to 50 in viral DNA replication may involve formation of a short amphipathic α-helix (AH) over the MxxFI motif. Importantly, the E1 proteins from other anogenital HPV types are also predicted to contain this short amphipathic α-helix within a similar IDD (see Fig. S3 in the supplemental material).

The E1 amphipathic helix is structurally related to p53 TAD2 and HSV VP16C. The TADs of p53 and HSV VP16 have been shown to make direct contacts with several components of the RNA polymerase II transcription machinery. One of our laboratories (J. G. Omichinski) has previously reported that the p53 and VP16 TADs bind with high affinity to the Tfb1/p62 subunit of the yeast/human general transcription factor IIH (TFIIH), a common target of many acidic TADs (12, 13, 30). In addition, we have determined that both p53 TAD2 and VP16C form 9-residue α-helices upon binding to the PH (pleckstrin homology) domain of Tfb1 (amino acids 1 to 115) (12, 30). The structures of these two complexes have revealed a simple recognition code for the Tfb1/p62 subunit of TFIIH and made it possible to identify putative acidic TADs in other proteins based on sequence alignments. Given that the E1 AH resembles p53 TAD2 and VP16C based on this recognition code for Tfb1/p62 (Fig. 2B), we were interested in determining whether a peptide corresponding to this region of HPV11 E1 (aa 44 to 63) was in fact capable of interacting with the PH domain of Tfb1.

To determine whether HPV11 E1(44-63) could interact with Tfb1 in a manner similar to p53 TAD2 and VP16C, we performed NMR chemical shift mapping studies. Additions of HPV11 E1(44-63) to 15N-labeled Tfb1(1-115) resulted in changes in 1H and 15N chemical shifts for several signals of Tfb1(1-115) in the 1H,15N heteronuclear single-quantum coherence (HSQC) spectra (Fig. 3A). When mapped onto the NMR solution structure of free Tfb1, the residues displaying significant chemical shift changes (Fig. 3B) clustered on strands β5, β6, and β7 (Fig. 3C). On the basis of these changes...
in chemical shifts, the HPV11 E1(44-63)-binding site on Tfb1 is similar to that of p53 TAD2 (Fig. 3D). It is also comparable to that of VP16C, because the latter uses the same binding surface as p53 (data not shown). In addition, isothermal titration calorimetry (ITC) analysis demonstrated that HPV11 E1(44-63) bound to Tfb1 (apparent dissociation constant \(K_D = 380 \text{nM}\)) with virtually the same affinity as p53 TAD2 \((K_D = 390 \text{nM})\) and VP16C \((K_D = 360 \text{nM})\) (data not shown). From these studies, we conclude that HPV11 E1(44-63) is structurally related to p53 TAD2 and VP16C and thus likely to form an amphipathic \(\alpha\)-helix.

The E1 N-terminal region activates transcription in yeast.

The finding that E1 could interact with yeast Tfb1 prompted us to determine whether the NTR of E1 could activate transcription in yeast. HPV11 E1 and HPV16 E1 have been previously observed to activate transcription of reporter genes in yeast when fused to the DNA-binding domain (DBD) of Gal4 and LexA, respectively, and for both proteins, transactivation activity could be mapped to the NTR (54, 61). We found that the N-terminal 83 amino acids of HPV31 E1, when fused to the LexA DBD, could also activate transcription in yeast, in this case in a LacZ reporter gene under the control of 8 LexA operators (Fig. 4A; also see Fig. S4A in the supplemental material). Note that the transactivation activity of each LexA fusion protein used in this study is reported in Miller units in Fig. S4A and S5A in the supplemental material. (C) Transactivation activity of the indicated LexA fusion proteins, either WT or mutants, are reported as fold activation of the reporter gene relative to that of the LexA DNA-binding domain alone (control [ctl]), which was set at a value of 1. The MFI mutation (LFI mutation for HPV16) changes the three hydrophobic residues of the E1 MxxFI motif (LxxFI motif for HPV16) to three alanines. E1 proteins of HPV6, HPV11, HPV16, HPV18, and HPV31 are shown (e.g., 31 E1, HPV31 E1). (B) Pictures of yeast colonies grown on solid selective medium 4 days after their transformation with plasmids coding for the indicated LexA-DBD fusion proteins. (C) Transactivation activity of the indicated LexA fusion proteins measured in yeast strain BY4741 or its isogenic derivatives lacking the GCN5, ADA1, or SPT7 gene.

Next, we investigated whether the E1 proteins from other anogenital HPV types could also activate transcription in yeast when fused to LexA. The E1 N-terminal regions from the two low-risk types HPV6 and HPV11, as well as those from the oncogenic types HPV16 and HPV18, were all able to activate transcription (Fig. 4A; see Fig. S4A in the supplemental material) and were expressed to comparable levels (see Fig. S4B in the supplemental material). Thus, the E1 proteins from both high- and low-risk HPV types can activate transcription in yeast. The E1 N-terminal regions from HPV6 and -18 displayed the most robust activity, being able to transactivate the reporter gene by 186- and 195-fold, respectively, similar to the levels obtained with p53 TAD2. Importantly, the transcriptional activities of the HPV6, -11, -16, and -18 NTRs were abolished by mutations of the three conserved hydrophobic.
residues to alanine (Fig. 4A; see Fig. S4A in the supplemental material), similar to what was observed for HPV31 E1. From these results, we conclude that the ability of the N-terminal region of E1 to activate transcription in yeast is a conserved feature of this viral helicase from both low- and high-risk HPV types.

Finally, we also investigated whether the transactivation activity of the E1 NTR in yeast was dependent on some of the same factors as those required for p53 TAD2. The ability of the full-length p53 TAD, or each of its subdomains, to function in yeast was previously shown to require the Spt7-Ada1-Gcn5-acetyltransferase (SAGA) chromatin-modifying complex, which is needed for the transcription of about 10% of yeast genes (7, 62). Some of the components of the SAGA complex are encoded by nonessential genes (22). We therefore tested whether the SAGA complex was needed for transcriptional activation by the E1 NTR by performing transactivation assays in yeast strains lacking the Gcn5, Ada1, or Spt7 component of the SAGA complex. Gcn5 is the subunit that has acetyltransferase activity, while Ada1 and Spt7 are required for the structural integrity of the complex (reviewed in reference 3). The transactivation activity of both E1 and p53 was severely reduced in the Ada1 and Spt7 deletion strains, while the effect of deleting Gcn5 was more moderate (Fig. 4C; see Fig. S5A in the supplemental material), perhaps because the acetyltransferase activity of SAGA is not absolutely necessary for transactivation by these LexA fusion proteins. However, as detected by Western blotting, all LexA proteins were expressed at lower levels in these deletion strains to various extents (see Fig. S5B in the supplemental material). Nevertheless, the finding that transactivation by the E1 NTR is affected to an extent similar to that of p53 TAD2 by alteration of the SAGA complex provides further evidence that E1 AH is similar to p53 TAD2.

The E1 N-terminal region activates transcription weakly in human cervical carcinoma cells. Next we tested the ability of the HPV31 E1 NTR(1-83) to activate transcription in C33A cells when fused to the Gal4 DNA-binding domain. For comparison, we also measured in parallel the activities of the full-length p53 TAD(1-73), p53 TAD2(40-73), and TAD of HPV31 E2(2-208). Each Gal4 fusion was tested at three concentrations for its effect on expression of a firefly luciferase reporter gene under the control of five Gal4-binding sites. Luciferase activities were measured from C33A cervical carcinoma cells transfected with increasing amounts of a plasmid encoding the indicated Gal4 fusion protein and a constant amount of pRL, encoding Renilla luciferase, which was used as an internal control. Luciferase activity values are presented as the fold activation of the reporter gene above that measured with the empty Gal4 DBD vector, which was set at 1. (B) Transactivation activities of the E1 NTRs from different HPV types. Transactivation activities of the Gal4-E1 NTR from the indicated papillomavirus types were measured as described above for panel A. (C) Effects of mutations in the E1 a-helix (AH) on transactivation activity. Transactivation activities of the Gal4-E1 NTR mutant derivatives were measured as described above for panel A. The MFI mutation changes the three hydrophobic residues of the E1 MxxFI motif to three alanines. Western blots showing the expression of the different Gal4 fusion proteins are shown to the right of each graph. The position of the full-length Gal4-E2 TAD protein, which was consistently found to show evidence of degradation products, is indicated by an asterisk. Tubulin was used as a loading control.

HPV31 E1 NTR. In several independent experiments, we found that its transactivation activity was reduced by approximately 40% by mutation of the three hydrophobic residues MFI (Fig. 5C), a significant but smaller reduction than that observed in yeast. We also observed that mutation of T43 to alanine (T43A) had little to no effect, while mutation to glutamate (T43E) was always slightly inhibitory (Fig. 5C). By Western blotting, we found that that the different Gal4-E1 NTR fusion proteins were always produced at very low levels in C33A cells, significantly less than those of p53 TAD2 and the
HPV E2 TAD (Fig. 5A to C, right panels). This was especially true for the NTRs of the two low-risk HPV types, HPV6 and HPV11, which were expressed at lower levels than those of the high-risk types (Fig. 5B). It is thus possible that the lower expression of the E1 NTR contributes to its lower transactivation activity. From these studies, we conclude that the E1 NTR can activate transcription weakly in C33A cells. However, given this low level of transactivation, we surmised that the primary role of the E1 AH may be in viral DNA replication rather than in transcription per se (see Discussion). The experiments below were designed to address the role of the E1 AH in viral DNA replication.

**Mutation of the E1 amphipathic helix reduces transient replication of HPV DNA.** As the primary function of E1 is to replicate the viral genome, we investigated whether the E1 AH was required for HPV DNA replication. Specifically, we tested the effects of the MFI, T43A, and T43E mutations in HPV31 E1 on its ability to support transient DNA replication using our luciferase-based assay (19). These experiments were performed with EYFP-HPV31 E1, which we showed previously is competent for DNA replication (18). Each E1 mutant was tested at three different amounts (1.25, 2.5, and 5 ng of EYFP-E1 expression plasmid) in duplicate experiments, and luciferase activities were determined 72 h posttransfection. As can be seen in Fig. 6A, the T43A and T43E E1 mutations did not impair the ability of E1 to support transient DNA replication, indicating that phosphorylation of this residue is not required for this process. In contrast, the DNA replication activity of the MFI mutant was decreased by approximately 40 to 50%. All four E1 proteins were expressed to comparable levels, as determined by Western blotting (Fig. 6B) and by flow cytometry (see Fig. S6 in the supplemental material).

Because the E1 AH is located between the shuttling module and the p80-binding site, we wanted to rule out the possibility that the reduced DNA replication activity of the MFI mutant was due to a defect in nuclear accumulation or a defect in interaction with p80. To do this, we first determined the cellular localization of the EYFP-E1 MFI mutant in the absence of E2. For sake of completeness, we also determined the localization of the two T43 mutants. All mutant E1 proteins accumulated in the nuclei of transfected cells, similar to wild-type (WT) E1 (Fig. 6C), indicating that none of the mutations tested has a negative effect on the function of the nearby “shuttling module.” We then investigated the interactions of these E1 mutants with p80 by GST pulldown experiments using purified GST-HPV31 E1(1-83) and endogenous p80 from a cell extract. Results presented in Fig. S7 in the supplemental material confirmed that all three mutant proteins were as competent as WT E1 in binding p80. Collectively, these results suggest that the E1 AH plays an important role in viral DNA replication that is independent of nuclear accumulation and p80 interaction.

**DNA replication defect of the E1 MFI mutant as a function of time and E1 concentration.** In order to ascertain the transient DNA replication defect of the E1 MFI mutant, we performed several additional mechanistic experiments. First, we investigated whether the replication defect of the MFI mutant would be affected by lengthening the duration of the transient DNA replication assay. To do so, we measured DNA replication catalyzed by the E1 MFI mutant, and by WT E1 as a control, every 24 h for a period of 5 days. Maximal signals were obtained after approximately 72 h for both proteins. Importantly, the signals obtained with the E1 MFI mutant remained 40 to 50% lower than those obtained with E1 WT at all time points tested (Fig. 7A). Thus, regardless of the length of the assay, the E1 MFI mutant consistently showed a 40 to 50% reduction in transient DNA replication activity. Next, we tested whether the DNA replication defect of the E1 MFI mutant could be rescued by overexpression of the protein, as might be anticipated if this E1 mutant is slightly unstable or has a lower affinity for E2. To do so, we performed transient DNA
The finding that the DNA replication defect of the E1 MFI mutant is independent of the tag fused to E1.

Similar results were obtained with vector and, importantly, was lower than that obtained with WT E1 under all conditions tested suggests that the MFI mutation affects the formation of the E1-E2 ori complex would provide an important clue as to when the E1 AH is required during DNA replication.

In an initial attempt to determine whether the MFI mutation affects the initiation phase of DNA replication, we tested whether it had an effect on the interaction of E1 with E2 by coimmunoprecipitation. Specifically, C33A cells were cotransfected with an expression vector for EYFP-E1, either the wild-type, MFI mutant, or ATPase-deficient K463A protein, together with a plasmid encoding Flag-tagged E2. Twenty-four hours posttransfection, E1 was immunoprecipitated with an anti-GFP antibody, and the presence of E2 in the precipitate was probed by Western blotting using an anti-Flag antibody. Results shown in Fig. 8A demonstrate that E2 could be coimmunoprecipitated with wild-type E1 as well as with the two mutant proteins, suggesting that neither the MFI mutation nor the K463A mutation significantly affect the interaction of E1 with E2. To provide additional and more quantitative evidence that the MFI mutation does not affect the interaction of E1 with E2, we made use of the ATPase-deficient K463A mutant E1. We previously showed that the corresponding mutation in HPV11 E1 (K484A) inhibits the ATPase activity of E1 but not its ability to bind cooperatively to the origin with E2 (54, 58). Furthermore, we determined in the coimmunoprecipitation studies presented above that the HPV31 K463A E1 can still interact with E2. Thus, we rationalized that the HPV31 K463A E1 would have a dominant-negative effect on viral DNA replication and, furthermore, that this inhibitory effect would be reduced by the MFI mutation if it compromises the interaction of E1 with E2 at the origin. The results presented in Fig. 8B indicated that both the E1 K463A mutant and the E1 MFI/K463A double mutant could inhibit DNA replication catalyzed by WT E1 in a dominant-negative and dose-dependent manner. Analysis of this inhibitory data by nonlinear regression confirmed that the E1 MFI/K463A double mutant was as good as E1 K463A at inhibiting viral DNA replication; both mutants were able to inhibit DNA replication by 50% (EC50) when present at a ratio of approximately 1:1 relative to wild-type E1. If anything, the E1 MFI/K463A double mutant appeared to be slightly more potent than E1 K463A in this experiment, although this may simply reflect the fact that it was expressed at somewhat higher levels, as determined by Western blotting (data not shown). Thus, we conclude that the MFI mutation has little to no effect on the dominant-negative activity of E1 K463A, and by extension, we infer that the MFI mutation does not alter the ability of E1 to interact with E2. Consistent with the notion that the MFI mutation...
affects a replication activity of E1 other than its interaction with E2, we showed that the deleterious effect of the MFI mutation on transient DNA replication was additive, rather than synergistic, with that of the E39Q mutation in E2, which weakens the E1-E2 interaction (see Fig. S8 in the supplemental material). In summary, the results presented above provide in vivo evidence that the MFI mutation does not alter the interaction of E1 with E2.

Next, we investigated whether the reduced replication activity of the E1 MFI mutant could be explained by a defect in origin binding by measuring its DNA-binding affinity in vitro using our previously described fluorescence polarization assay (20, 52). First, we produced and purified from bacteria a polypeptide spanning amino acids 2 to 332 of HPV31 E1, either wild type or containing the MFI mutation (see Fig. S9A in the supplemental material). These polypeptides encompass the complete amino-terminal region and origin-binding domain (OBD) of E1. The affinities of these E1 fragments for a DNA probe containing two E1-binding sites (E1BS) (5'-H11032-ATACTT-3') in an inverted orientation and separated by 3 bp were then measured. We previously determined that the E1 OBD can bind and dimerize efficiently on this combination of E1BS (19, 52). A similar probe in which the two E1BS were mutated was used as a specificity control. As can be seen in Fig. 8C, both E1 fragments were able to bind specifically to the 2E1BS probe with comparable affinities. Thus, the MFI substitution does not impair the function of the OBD. Similar results were obtained with E1 fragments containing the T43A and T43E substitutions (see Fig. S9B in the supplemental material).

Collectively, the results presented above indicate that the MFI mutation does not affect the interaction of E1 with either E2 or the origin. From these data, we infer that the E1 AH plays a role in viral DNA replication at a step following the assembly of the E1-E2-ori preinitiation complex.

**DISCUSSION**

The E1 N-terminal regulatory region contains a conserved amphipathic α-helix similar to p53 TAD2 and VP16C. Our deletion analysis of the HPV31 E1 N-terminal region (NTR) confirmed the importance of the p80-binding domain (aa 1 to 40) for transient HPV DNA replication and, importantly, identified the region located immediately next to it (aa 40 to 50) as also being required for this activity. In this study, we show that this region is highly conserved among E1 from several HPV types, lies within a predicted intrinsically disordered domain (IDD), and shows similarity to the p53 TAD2 and VP16C transactivation domains, which are also IDDs. IDDs are protein domains that have no intrinsic secondary structure in their free form but adopt a specific conformation when they bind to their target. This inherent flexibility allows them to interact with several different binding partners. It is therefore not surprising that IDDs are very common among eukaryotic proteins involved in regulatory and signaling processes, especially among transcription factors (reviewed in reference 21). The predicted E1 IDD is conserved in different HPV types (Fig. 2A; see Fig. S3 in the supplemental material) and, as is typical of these unstructured domains, is of low sequence complexity, being largely devoid of hydrophobic amino acids (except for the MxxFI motif) and containing a high proportion of polar...
(Ser and Thr) and negatively charged (Asp and Glu) amino acids. Several lines of evidence support the existence of this IDD in E1. First, our preliminary characterization by one-dimensional (1D) NMR of a recombinant polypeptide spanning HPV31 E1(1-83) indicates that it is very flexible in solution (J.G.O., unpublished data). Specifically, the 1H NMR spectra display very limited chemical shift dispersion, suggesting that E1(1-83) does not possess any regular elements of secondary structure in the unbound form. Second, the presence of this predicted IDD region makes the E1 NTR migrate slower on SDS-polyacrylamide gels than predicted on the basis of its molecular weight (data not shown). Third, when expressed in E. coli, it is accessible to cleavage by bacterial proteases (see Fig. S9 in the supplemental material; also data not shown). Similarly, in mammalian cells, it is accessible to cleavage by caspases 3 and 7 (see Fig. S1 in the supplemental material), thus indicating that this region is solvent accessible.

The p53 TAD2 and VP16C IDDs fold into a 9-residue amphipathic α-helix (AH) upon binding to their target proteins such as the Tfb1/p62 subunit of TFIH (12, 28, 30, 57). Three key hydrophobic residues (Φ) located in a ΨxxΦΨ motif on one face of this AH make direct contact with the target protein (Fig. 2C). In agreement with the E1 IDD being structurally similar to p53 TAD2 and VP16C, our ITC and NMR chemical shift mapping studies revealed that a peptide spanning the conserved ΨxxΦ hydrophobic motif of E1 binds the PH (pleckstrin homology) domain of Tfb1 with a comparable affinity and in a similar manner as p53 TAD2 and VP16C, indicating that it very likely folds into a similar AH. On the basis of these observations, we propose that the E1 NTR contains an IDD that folds into an AH to facilitate viral DNA replication.

The E1 NTR can activate transcription in yeast and, to a lesser extent, in mammalian cells. Consistent with the E1 NTR containing a region with similarity to p53 TAD2 and VP16C, we found that the E1 NTRs from five different HPV types (6, 11, 16, 18, 31) could activate transcription in yeast when fused to the LexA DNA-binding domain. This conserved transcription activity of E1 was found to require the SAGA histone acetyltransferase complex, similar to what was shown for p53 TAD2 (7) and, importantly, to be dependent on the integrity of the ΨxxΦ motif, underlying the role of the AH of E1 in this activity. When tested as a Gal4 fusion in mammalian cells, the E1 NTR was also capable of activating transcription, albeit to a lower extent than p53 TAD2, and mutation of the ΨxxΦ motif reduced this activity by 40%. Thus, relative to p53 TAD2, the E1 AH appears to be a stronger activator in yeast than in mammalian cells. It was previously shown that the full-length HPV18 E1 protein can activate transcription weakly in mammalian cells when tethered to DNA by fusion to the E2 DNA-binding domain (10). Our results are in complete agreement with this finding and point to the E1 NTR as being responsible for this activity.

At this time, it is unclear whether E1 functions as a transcription factor during the viral life cycle. The weak transactivation activity of E1 in mammalian cells (10; this study) suggests that it may not regulate transcription per se. One possibility could be that E1 needs to interact with one or more transcription factors to facilitate viral DNA replication. An equally plausible interpretation would be that the transcription activity of E1 simply reflects the fact that the NTR contains a region capable of folding into an AH. Indeed, several years ago, the research group of Ptashne reported that in yeast, artificial transactivation domains, selected from random pieces of DNA fused to the coding region of a heterologous DNA-binding domain, are acidic in nature and often predicted to contain AHs (35, 46). Regardless of the exact significance of the E1 transactivation activity, our results indicate that the highly conserved AH is required for efficient replication of viral DNA, and we discuss below the possible roles that this AH may play in this process.

The E1 AH region is necessary for optimal viral DNA replication. As part of this study, we investigated the requirement for the E1 AH in viral DNA replication. We found that mutation of the MxxFI motif in the E1 AH reduced transient viral DNA replication by approximately 50%, irrespective of the duration of the assay or the level of expression of E1. Perhaps not surprisingly, this reduction in viral DNA replication brought about by the MFI mutation was of a similar magnitude as its effect in mammalian transactivation assays (40%; Fig. 5C). Further mechanistic studies revealed that the E1 MFI mutant was able to accumulate in the nucleus, to interact with p80, and to bind to E2 and to the origin of replication as well as to wild-type E1. The finding that the E1 AH is not required for interaction with E2 is consistent with previous studies indicating that the C-terminal helicase domain of E1, rather than the NTR, mediates this interaction (36, 39, 51, 54, 61). On the basis of these results, we propose that the E1 AH is implicated in a novel activity of E1 that is required for efficient replication of viral DNA at a step following assembly of the E1-E2-ori complex.

As part of this study, we also investigated the function of a highly conserved putative CK2 phosphorylation site, T43, located within the E1 AH. Mutation of T43 to alanine (to prevent its phosphorylation) or to glutamate (as a phosphomimetic) had little to no effect on the ability of HPV31 E1 to support transient DNA replication in C33A cells. Thus, the reason underlying the evolutionary conservation of this putative CK2 site remains unknown. It is also noteworthy that the sequence of the E1 AH overlaps the conserved caspase cleavage sites in E1, which were previously shown to be required for amplification of the viral genome in differentiated keratinocytes (37). This prompted us to investigate whether one of the two T43 mutations or the MFI mutation that changes residues within the caspase sites (DxxDxxD and DMVDFID) had an effect on the susceptibility of E1 to caspase cleavage. Under our experimental conditions, none of the mutations had an effect (see Fig. S1 in the supplemental material). The fact that the MFI mutation has no effect on caspase cleavage but impairs DNA replication indicates that these two processes are distinct events. This notion is further supported by the findings that activated caspases are present only at very low levels in C33A cells (see Fig. S1 in the supplemental material) and that mutations of the caspase cleavage sites in E1 have little to no effect on DNA replication (Fig. S2).

Potential functions of the E1 amphipathic helix in papillomavirus DNA replication. ΨxxΦΨ motifs (where Φ represents a hydrophobic amino acid) present within α-helices are often involved in mediating protein-protein interactions. In addition to those in the p53 and VP16 TADs, other well-characterized examples include the LxxLL motif found in several transcrip-
tional coactivators that bind to ligand-associated nuclear receptors (24, 31, 55) and the LxxLL motif present in many of the proteins that interact with the E6 protein of high-risk HPVs and bovine papillomavirus (BPV), such as E6AP whose LxxLL motif was shown to lie within an α-helix by NMR (4). Thus, the presence of the conserved MxxFI motif in the E1 AH raises the possibility that it also mediates a protein-protein interaction required for efficient replication of HPV DNA. A possible interaction partner may be E1 itself, for example if the AH region is required for proper assembly of a double hexamer at the origin. Alternatively, or in addition, this region of E1 could interact with cellular factors to promote viral DNA replication.

Given that the papillomavirus genome is packaged into nucleosomes (16, 50), E1 could interact with chromatin-remodeling complexes to facilitate progression of the replication fork. During the initiation phase of viral DNA replication, the transcription factor E2 relieves the inhibition exerted by nucleosomes at the origin. Interestingly, this antirepression mechanism can also be achieved by other transcription factors, including p53 and VP16, when tethered near the ori (33). The E1 AH region could assume this function together with E2 or E2/AH region could assume this function together with E2 or E2 may also interact with TFIIH. As TFIIH is involved in both transcription and DNA nucleotide excision repair, it may play similar roles in the context of the viral episome. The E1 AH, and more generally the entire E1 IDD, could mediate interaction with several different cellular factors, the ones mentioned above representing a few possibilities. It is noteworthy that genes carried by small DNA tumor viruses often encode proteins with IDDs, as exemplified by the E1A protein of adenovirus (43). The use of IDDs may thus represent an evolutionary solution for the need of these viruses to interact with a plethora of host factors to complete their life cycle despite the limited coding capacity of their genomes.

ACKNOWLEDGMENTS

We thank Joe Mymryk (University of Western Ontario) for the gift of the SAGA deletion yeast strains and Lou Laimins and Peter Bullock for critical reading of the manuscript. This work was supported by grants from the Canadian Institutes of Health Research (CIHR) and the Canadian Cancer Society Research Institute (CCSRI) to J. A. and J. G. O., respectively. A.-T. H. holds a studentship from the Fonds de la Recherche en Santé du Québec (FRSQ).

REFERENCES

papovavirus DNA replication initiator proteins, bovine papillomavirus type 1 E1 and simian virus 40 large T antigen, with human replication protein A. J. Virol. 78:8499–8497.


