Minireview

The Papillomavirus E2 Regulatory Proteins

Alison A. McBride, Helen Romanczuk, and Peter M. Howley

From the Laboratory of Tumor Virus Biology, National Cancer Institute, Bethesda, Maryland 20892

Papillomavirus Life Cycle

The papillomaviruses are small DNA viruses that induce benign proliferative squamous epithelial and fibroepithelial lesions (warts and papillomas) in their natural hosts. They have been isolated from a variety of animal species, and over 60 human papillomavirus types have been described so far (reviewed in Ref. 1). The papillomavirus life cycle is closely linked to keratinocyte differentiation. In the basal epithelial cells and dermal fibroblasts of a fibropapilloma the viral genome is maintained as a multicopy, latent, non-productive infection is thought to be analogous to the pattern observed in filropapillomaviruses such as BPV-1 (Fig. 1). In the cancer-associated human papillomaviruses HPV-16 and HPV-18, the major promoter (P$_{7940}$ and P$_{99}$, respectively) is negatively regulated by E2 (24-26). This repression is mediated through two E2-binding sites immediately adjacent to the TATA box of the promoters (26, 27). Occupation of these sites by E2 is thought to interfere with assembly of the transcription preinitiation complex (see Fig. 3A and below). E2 regulation of the E6 promoter of HPV-11 also provides evidence for both positive and negative transcriptional modulation (28, 29).

In the majority of HPV-containing cervical carcinomas the viral genome is often integrated such that the E1 and/or E2 ORFs are disrupted. It has been proposed that the absence of these polypeptides, the E6 and E7 oncproteins are expressed in a deregulated manner which in turn may contribute to tumor development. In support of this model, it has been shown that mutation of the promoter proximal E2-binding sites in the HPV-16 genome results in a slight but reproducible increase in the ability of the viral DNA to immortalize human keratinocytes. Moreover, mutations which disrupt the HPV-16 E1 or E2 ORF cause a further augmentation of the immortalization potential, suggesting that these ORFs play additional roles in regulating the levels or activities of the viral oncoproteins.

Structural and Functional Domains of the E2 Polypeptide

A comparison of the predicted amino acid sequences of a number of papillomavirus E2 ORFs reveals that these polypeptides probably consist of three domains (30, 31). Two of the domains, consisting of about 220 amino acids at the N terminus and 90-100 amino acids at the C terminus, are relatively well conserved among papillomaviruses. However, the internal region varies both in length and amino acid sequence (Fig. 2).

The N-terminal region, which is unique to the full-length transactivator protein, encodes the activation domain (32-34). This domain is large compared with many other transactivators, and this may reflect the multifunctional role of E2 in transcription and replication. The activation domain contains two regions within the first 85 amino acids that can be predicted to form acidic amphipathic helices. These structures have been shown to be important for the activation function of a number of eukaryotic transactivators (35). Deletion analysis of BPV-1 E2 demonstrated that the region containing the potential amphipathic helices is necessary but not sufficient for transactivation (34). DNA-bound dimers of the BPV-1 transactivator protein can interact in vitro to form DNA loops, and this association requires the N-terminal domain of E2 (36). In addition, as described below, the transactivator can associate with the BPV-1 E1 polypeptide and is necessary for viral DNA replication (37-39). It is likely that sequences in the N-terminal domain are required because the E2
**E2 DNA-binding Sites**

The E2 proteins bind specifically to a 12-base pair palindromic sequence, ACCN4GGT, which is repeated several times in the papillomavirus genomes (11, 12, 46). BPV-1 has 12 sites that correspond to this consensus and an additional five closely related sequences that are also bound by E2 (47). The majority of these sites are located in the LCR, but there are several situated close to the promoters in the early region (Fig. 1). The affinity of the BPV-1 E2 protein for the various DNA-binding sites ranges over several orders of magnitude. The higher affinity sites are clustered in the LCR in regions that have been designated E2-dependent enhancer elements E2RE5 and E2RE6 (13).

A low level of E2-dependent activation can be obtained with one E2 binding motif in Saccharomyces cerevisiae and in mammalian cells. However, two E2 DNA-binding sites cooperate to constitute a strong enhancer (48-52). Enhancer elements such as E2RE5, multiple binding sites cooperate to form a highly responsive element (53). Some cooperativity in DNA binding has been observed in vitro with the BPV-1 E2 protein, but this is not sufficient to explain the strong synergism seen in vivo, suggesting that additional mechanisms are probably involved (36, 42, 48).

E2 transactivates the P443 and P3930 promoters located in the early region, and activation depends on the E2-dependent enhancers in the LCR (15-17). It was postulated that the E2-binding sites adjacent to these promoters might cooperate with those in the LCR to bring about activation. DNA-bound E2 transactivator molecules have been demonstrated to associate and form DNA loops, providing a model as to how activation could occur at a distance (36). Alternatively, binding of E2 to these sites could repress expression from the promoters in a manner analogous to the situation with the HPV-16 P37 and HPV-18 P30 promoters. However, mutation of the E2 site adjacent to the P443 promoter (which has the lowest affinity of all the BPV-1 sites) has no effect on E2 regulation of this promoter (17, 54), and mutation of the two E2 sites adjacent to the P3930 promoter results in only a minimal decrease in E2-dependent promoter activity (17). Nevertheless, it is possible that these sites could have an important regulatory role under different circumstances. For example, the DNA binding specificity of E2 for the various motifs could change during the cell cycle, with the state of cellular differentiation, or at a particular stage of the viral life cycle.

The number and distribution of the E2-binding sites differ in the human papillomaviruses as compared with the animal papillomaviruses. In the genital papillomaviruses HPV-16 and HPV-18, only two E2-binding sites have been identified, all of which are located in the LCR (Fig. 1). As noted above, two of these sites are situated immediately upstream from the respective TATA boxes of the E6/E7 viral promoters (designated P37 in HPV-16 and P30 in HPV-18), and binding of E2 to these sites represses rather than activates transcription from these promoters (24-27). Mutation of these proximal E2-binding sites so that they no longer bind E2 results in E2-mediated activation of transcription.
of the P7 or P185 promoters through interaction with the remaining upstream E2-binding sites (26, 27). These results indicate that E2 modulation of HPV transcription can involve both positive and negative regulation. The final outcome may be a function of the different binding affinities of E2 for each of the sites (27, 55). Alternatively, E2 may positively regulate some other, as yet unidentified, HPV promoters.

The precise contact points of the E2 proteins on the ACCN/GGT motif have been determined (41, 55). The protein contacts the guanosine residues on either strand of the motif, which are predicted to be present on the same face of the DNA helix. The nucleotides in the non-conserved core and immediately outside the binding site determine the affinity for the E2 protein. Binding of both BPV-1 and HPV-16 E2 proteins has been shown to bend DNA, and the HPV-16 E2 preference for an A-T rich core sequence may reflect the flexibility of such a sequence (55, 56).

**Mechanisms of E2 Transcriptional Regulation**

As mentioned above, there are several ways in which the E2 proteins can either activate or repress transcription, depending on the position of the E2-binding sites and the nature of the E2 polypeptides. These are shown diagrammatically in Fig. 3. In one case, such as for the BPV-1 P740 and P50 promoters, binding of the full-length E2 transactivator to a number of sites located upstream from the promoter activates transcription (13, 14). E2 can also stimulate transcription from heterologous promoters which contain several E2-binding sites located at some distance either upstream or downstream from the promoter (51). In these cases, the E2 sites behave like classical conditional enhancer elements. However, the E2 transactivator can also repress transcription when the E2 motifs are situated close to the transcriptional start site, overlapping the binding sites for essential cellular transcription factors. In the case of HPV-16 (and HPV-18), two E2 sites are located just upstream from the TATA box of the P7 promoter, and binding of E2 to these sites may sterically hinder the binding of the cellular factor, TFIIID, and/or interfere with the assembly of the transcriptional initiation complex (26, 27, 57). There is also an essential Sp1-binding site located just upstream from the E2 motifs, but it is not known whether occupation of the E2 sites interferes with binding of Sp1 to this region of the promoter (58). In a similar situation, an E2-binding site located downstream from the BPV-1 P185 promoter overlaps an Sp1-binding site. This Sp1 site is essential for P185 activity, and binding of E2 polypeptides can repress expression from this promoter, probably by interfering with Sp1 binding (22, 23). Finally, as described above, the BPV-1 P1440 promoter has a low affinity E2-binding site located just upstream from the TATA box, but this site does not appear to play a role in the regulation of this promoter. This E2 site is adjacent to an Sp1 site which is critical for both basal and E2-dependent promoter activity, and E2 transactivation of this promoter depends on the E2 sites located 2.5 kilobases upstream in the LCR (17, 54).

The mechanism by which E2 stimulates transcription has not yet been elucidated, but it is likely that activation results from a direct or indirect interaction of E2 with some component of the basic transcriptional machinery. E2 is unable to activate minimal promoters containing only a TATA box but requires additional promoter proximal elements which bind factors such as Sp1 (59). There are Sp1 sites located close to many papillomavirus promoters, and several of these sites have been shown to be essential for promoter activity (17, 23, 54, 58). Many of these sites are adjacent to or overlap E2 binding motifs and may play a role in E2-mediated regulation. Sp1 can interact directly with the full-length E2 polypeptide and can enhance E2 binding to weak affinity DNA sites (17). Binding of Sp1 and E2 to sites spaced widely apart on a DNA fragment results in the formation of DNA loops in vitro and provides a conceivable mechanism by which the LCR E2-binding sites can activate transcription from promoters located far downstream in the early region, such as P2440 and P15090. However, it seems unlikely that the interaction between E2 and Sp1 is exclusive because E2 is able to synergize with other promoter proximal elements (59).

Modulation of papillomaviral transcription also depends on the nature and type of interaction of the E2 polypeptides (Fig. 3B). There are several mechanisms by which the E2 repressor proteins could inhibit transcriptional activation mediated by the full-length transactivator. The repressors could bind and titrate out essential cellular factors required for activation. Alternatively, the repressors could block E2 transactivation by binding competitively to the ACCN/GGT DNA-binding site or by forming potentially inactive heterodimers with the transactivator species. Since all three E2 regulatory factors share the C-terminal domain required for DNA binding and dimerization, the latter two models are likely mechanisms for E2 repression (Fig. 3B).

**E2 Polypeptides and Viral DNA Replication**

It has recently been shown conclusively that the full-length products of the E1 and E2 ORFs are both necessary and sufficient for BPV-1 transient DNA replication (39). The E1 polypeptide has some homology with SV40 and polyomavirus T antigens and binds ATP and binds DNA nonspecifically (38, 60, 62). In addition, the full-length E1 and E2 polypeptides form a complex capable of binding to E2 DNA sites, suggesting that E2 might direct the E1 replication protein to the viral origin of replication (37, 38). In support of this hypothesis, the minimal origin required for transient replication has been mapped to a 105-base pair fragment spanning the Hpol restriction site at nucleotide 1 of the

---

**Fig. 3. Mechanisms of E2 regulation.** A, regulation of transcription by position of E2 DNA-binding sites. Four examples of papillomavirus promoters are shown. Transcriptional activation or repression depends on the position and proximity of the E2 motifs with respect to the binding sites of cellular factors. B, regulation of transcription by different E2 polypeptide species. Two mechanisms by which the E2 repressors could inhibit transcription by the full-length E2 protein are shown.
BPV-1 genome. This fragment also contains E2-binding sites 11 and 12. However, subsequent studies have shown that the E1 polypeptide does in fact bind specifically to a region spanning the HpaI site (63) and that the adjacent E2 sites are not required for transient DNA replication. The role of E2 in replication is therefore less clear. It may function by stabilizing the interaction of E1 with DNA or by altering the chromatin structure around the origin of replication thus allowing greater accessibility of cellular replication factors. In support of the former hypothesis, it has recently been shown that cellular extracts supplemented with purified BPV-1 E1 and E2 proteins can support replication of exogenously added papillomavirus DNA. The E2 transactivator stimulates binding of the E1 replication protein to the origin of replication and thereby activates DNA replication (64).

Only the full-length E2 protein is able to complex with E1, but it is possible that the E2 repressors could regulate replication by competitively binding to E2 sites adjacent to the replication origin or by complexing with the E2 transactivator. In support of this model, a BPV-1 mutant that is unable to express the E2-TR repressor replicates at a high copy number (65, 66).

Regulation of E2 Function

The abundance and activities of the E2 polypeptides can be regulated at several levels. As described above, the promoters of cellular replication factors. In support of the former hypothesis, El mutants exhibit increased replication and thereby activates DNA replication (64).

The E2 proteins may also be regulated by post-translational modifications. The E2 polypeptides are phosphorylated primarily on two serine residues at amino acids 298 and 301 in a region of the hinge immediately adjacent to the DNA-binding domain and present in all three E2 polypeptide species (45). Substitution of serine 301 with an alanine residue results in a virus that replicates to a much higher copy number than that of the wild-type virus (67). This suggests that E2 phosphorylation may play a direct role in viral DNA replication. In support of this hypothesis, E2 binds preferentially to the underphosphorylated form of E2, suggesting that phosphorylation may modulate the E1-E2 protein interaction (68). Alternatively, E2 phosphorylation may modulate replication indirectly by regulating viral gene expression.

While it seems likely that an E1-E2 protein complex is required for viral DNA replication, there is little data on the effect of this complex on the transcriptional properties of the E2 transactivation. It has been postulated that the E1 ORF encodes a transcriptional repressor because BPV-1 E1 mutants exhibit increased viral transcription (69, 70); however, the mechanism by which this phenotype arises has not yet been elucidated.

REFERENCES

1. De Villiars, E. M. (1989) J. Virol. 63, 4895-4903
2. DiMaio, D., and Neary, K. (1988) in Papillomaviruses and Human Cancer (Pfister, H., ed) pp. 133-144, CRC Press, Inc., Boca Raton, FL
3. Lamport, P. F. (1981) J. Virol. 46, 3417-3430
4. Münker, K., Windsch, P., and Howley, P. M. (1990) in The Cellular and Molecular Biology of Human Carcinogenesis (Boutwell, R. K., and Riegel, L. L., eds) pp. 223-254, Academic Press, San Diego
5. Rabson, M. S., Yee, C., Yang, Y. C., and Howley, P. M. (1986) J. Virol. 60, 629-634
6. DiMaio, D. (1986) J. Virol. 57, 476-480
7. Groffen, D. E., and Lancaster, W. D. (1986) Virology 150, 211-218
8. Additional references are: P. M. (1981) J. Virol. 41, 3889-3895
9. DiMaio, D., and Neary, K. (1988) EMBO J. 7, 1197-1204
10. Spalholz, B. A., and Howley, P. M. (1984) Cell 22, 183-191
11. Moskvihuk, C., and Bastis, D. (1987) Proc. Natl. Acad. Sci. U.S.A. 84, 1210-1213
12. Androphy, E. J., Lown, R. L., and Schiller, J. T. (1987) Nature 325, 70
13. Spalholz, B. A., Lambert, P. F., Yee, C. L., and Howley, P. M. (1987) J. Virol. 61, 2128-2137

* A. Ustav, E. Ustav, P. Symanek, and A. Stehling, submitted for publication.