Recruitment and loading of the E1 initiator protein: an ATP-dependent process catalysed by a transcription factor

Cyril M.Sanders and Arne Stenlund¹

Cold Spring Harbor Laboratory, P.O. Box 100, Cold Spring Harbor, New York, NY 11724, USA

¹Corresponding author
e-mail: Stenlund@cshl.org

Initiation of DNA replication critically depends on *ori* recognition as well as on catalytic activities of the initiator complex. For replication of papillomaviruses the catalytic activities for initiation are provided by the E1 protein. Here, we show that the transcription factor E2 acts to assemble E1 into a complex active for *ori* distortion in two steps. First, cooperative DNA binding of E1 and E2 generates a sequence-specific *ori* recognition complex. In the second ATP-dependent step, E2 is displaced and additional E1 molecules are incorporated. The net result is a final complex with low sequence specificity deposited onto a specific sequence in the DNA. This may be a general strategy to accomplish specific positioning of protein complexes with low sequence specificity.

Keywords: DNA replication/initiator/papillomavirus/replicator/transcription factor

Introduction

The original replicon hypothesis (Jacob *et al.*, 1963) postulated the existence of defined genetic elements (replicators) that engage protein or protein complexes (initiators) in order to govern the specific and timely onset of DNA replication. Studies in recent years have largely upheld the central tenets of this hypothesis. The initiation of replication is believed to occur in several steps: specific recognition of the origin precedes origin unwinding, helicase delivery and processive DNA synthesis. In some cases, such as initiation of replication at *Escherichia coli oriC*, these functions are performed by several different proteins. In others, exemplified by the DNA tumour viruses, a single protein is responsible for all these functions. The replicators of simple eukaryotic genomes, such as viral chromosomes, are modular in nature, being comprised of core and auxiliary components (DePamphilis, 1993, 1996). The core components include a binding site for the initiator protein, an A/T-rich region and a DNA-unwinding element. The latter two are regions of low thermodynamic stability, prone to distortion or strand separation. The auxiliary components are generally binding sites for sequence-specific transcription factors. These were first recognized and subsequently shown to act as potentiators of core activity in viral systems (Bergsma *et al.*, 1982; de Villiers *et al.*, 1984; Guo *et al.*, 1989).

In addition to DNA replication, multiprotein complexes assembled on DNA are required for other specialized

processes such as transcription and recombination. One of the requirements for understanding how these complexes function is to understand their assembly, including the basis for their recognition of specific sequences in DNA. These complexes may either exist as multiprotein complexes in solution or assemble on the DNA from individual components in an ordered fashion. In either case, the sequence complexity of mammalian organisms presents a great challenge to recognition of specific sites in DNA.

This consideration also affects the parasites that infect mammalian cells. Papillomaviruses represent an interesting group of viruses in this regard since in their latent state they behave similarly to cellular genes, directing only very low levels of gene expression from a small number of genomes. At early stages after infection, the viral sequences are present at single-copy levels in a vast excess of cellular DNA. Two obvious alternative strategies exist to achieve occupancy of the appropriate sequences in the viral genome by viral proteins. One alternative is the expression of high concentrations of a factor with low sequence specificity. The other possibility is to generate binding with a high degree of sequence specificity. Papillomaviruses provide an interesting example of how sequence specificity can be achieved through cooperative DNA binding of two different viral proteins. The bovine papillomavirus (BPV) origin of DNA replication contains adjacent binding sites for the two viral proteins E1 and E2, both of which are absolutely required for replication *in vivo* (Ustav and Stenlund, 1991; Ustav *et al.*, 1991, 1993; Yang *et al.*, 1991). The viral initiator E1 binds DNA but with low sequence specificity. In the presence of the viral transcription factor E2, E1 and E2 bind cooperatively to the *ori*, resulting in a complex with greatly increased sequence specificity and affinity (Mohr *et al.*, 1990; Lusky *et al.*, 1993; Seo *et al.*, 1993a; Sedman and Stenlund, 1995; Sedman *et al.*, 1997). Although many other E2-binding sites are present in the viral genome, replication initiates at this particular site, demonstrating that E1, in spite of its low specificity, is the ultimate determinant for where this complex forms.

Here we have asked how the E2 transcription factor functions in the assembly of a DNA replication pre-initiation complex. The surprising result is that E2 functions transiently and 'catalytically', providing sequence specificity for the formation of an E1-*ori* complex that can distort the *ori* DNA and therefore is likely to represent an early step in the initiation of DNA replication. This activity is provided in a two-step process where the first step involves the binding of E1 and E2 to *ori*, forming a highly sequence-specific complex. The second step involves the assembly of additional E1 molecules onto this complex and the displacement of E2 in a process that requires ATP hydrolysis. The net result is that a complex with very limited sequence specificity is

Fig. 1. E1 and E2 binding to the origin of replication. (A) Structure of the BPV-1 minimal origin (replicator). The minimal *ori* spans 59 bp (BPV-1 genome coordinates 7914–27) and consists of an A/T-rich region and the E1- and E2-binding sites separated by 3 bp. (B) E1-*ori* and E1E2-*ori* complex formation: effect of ATP/Mg²⁺ and complex stability in the presence of ethidium bromide (EtBr). Parallel binding reactions with wild-type origin probe were assembled with and without ATP/Mg²⁺. After 30 min, reaction products were cross-linked and divided in two. One sample was adjusted to 25 µg/ml EtBr prior to electrophoresis (lower panel). The E1 titration was from 50, 30, 20, 12.5 to 2.5 nM (lanes 1–5 and 11–15). For formation of the E1E2-*ori* complex (lanes 6–8 and 16–18), E1 was at 2.5 nM and E2 was increased from 0.5 to 2.5 nM. At the highest concentration of E2 alone, no complex formation is observed (lanes 9 and 19).

deposited onto a specific sequence in the DNA. We suggest that this 'seeding' may be a general method to accomplish specific positioning of complexes with a low degree of sequence specificity. Thus, the results presented here provide an interesting example of how sequence-specific recognition can occur and how this process can be coupled to the assembly of a multiprotein–DNA complex at specific sites.

Results

Formation of the E1E2-*ori* and E1-*ori* complexes

Two complexes, E1E2-*ori* and E1-*ori*, can form on the BPV minimal origin of replication (Figure 1A) *in vitro* (Sedman and Stenlund, 1995; Sedman *et al.*, 1997). In the E1E2-*ori* complex, E1 is bound to DNA as a dimer, and in E1-*ori* it is likely that four molecules of E1 interact with binding sites in the DNA (Chen and Stenlund, 1998; unpublished observations). The sequence specificity and affinity of E1 and E2 bound cooperatively to DNA (E1E2-*ori*) are much greater than that of E1 bound alone in the E1-*ori* complex. Since E1 alone possesses all the activities required for replication *in vitro* (Yang *et al.*, 1991, 1993; Seo *et al.*, 1993b), but initiates replication with low sequence specificity, we hypothesized that E2 may function

to generate *ori* specificity *in vivo*. E1E2-*ori* could be a precursor required to form the E1-*ori* complex that is in turn active for replication. This would be consistent with the observation that the ability to replicate *in vivo* correlates with the ability to form both complexes *in vitro*. If this were the case, several predictions should be met: (i) we would expect E2 to stimulate the formation of the E1-*ori* complex; (ii) E1E2-*ori* should be a preferred substrate for E1-*ori* complex formation compared with naked *ori* DNA; and (iii) E1 molecules recruited to *ori* by E2 should be incorporated preferentially into the final E1-*ori* complex. To address these three questions, we first investigated the effects of the nucleotide cofactor adenosine triphosphate (ATP/Mg²⁺) on *ori* complex formation since E1 activities required for replication are ATP dependent. The results are shown in Figure 1B.

E1E2-*ori* and E1-*ori* complexes form in the absence of ATP/Mg²⁺ and can be observed by agarose gel electrophoresis, after protein cross-linking with glutaraldehyde (Lusky *et al.*, 1993; Sedman and Stenlund, 1995). For E1 alone, binding to DNA is highly cooperative and of low sequence specificity. At low E1 concentration in the presence of E2, where binding of E1 alone is no longer observed, a highly sequence-specific E1E2-*ori* complex forms (Sedman and Stenlund, 1995, 1996; Sedman *et al.*, 1997). As shown in Figure 1B, formation of the E1-*ori* complex was not altered significantly in the presence of nucleotide cofactor (Figure 1B, top panel, lanes 1–5 compared with 11–15), but the half-life of the complex increased significantly from ~18 min to very much greater than 60 min when ATP/Mg²⁺ was present (data not shown). ATP/Mg²⁺ may alter the association rate for E1–DNA binding since the extent of complex formation was similar under both conditions. Formation of the E1E2-*ori* complex in the presence of ATP/Mg²⁺ was reduced (Figure 1B, lanes 6–8 compared with 16–18), consistent with a 3-fold decrease in the half-life of the complex from ~95 to 35 min in the presence of ATP/Mg²⁺ (data not shown). Under these conditions, binding of E2 alone to the low affinity E2-binding site is not observed by gel shift analysis (Figure 1B, lanes 9 and 19) or DNase I protection (see below; Figure 4). DNase I and dimethyl sulfate (DMS) footprinting assays indicated that the nucleotide cofactor had little effect on the structure of the E1E2-*ori* complex (data not shown). Therefore, the same complexes can form on the minimal BPV-1 origin of replication, in the presence or absence of ATP/Mg²⁺.

Our previous characterization of E1-*ori* and E1E2-*ori* has shown that E1 and E2 bind to one face of the DNA helix in the E1E2-*ori* complex, but that E1 molecules encircle the DNA in the E1-*ori* complex (Sedman and Stenlund, 1996). Consistent with these observations, the two complexes showed different sensitivities to the intercalating agent ethidium bromide (EtBr). When EtBr was added to 25 µg/ml to pre-formed, cross-linked complexes, the E1E2-*ori* complex dissociated from the *ori* probe (Figure 1B, lanes 6–8 and 16–18, compare the lower panel with the upper panel), while the E1-*ori* complex was stable (Figure 1B, lanes 1–5 and 11–15, compare lower and upper panels). One explanation is that in the cross-linked E1-*ori* complex, the E1-binding site is inaccessible to EtBr. Protein–DNA but not protein–protein interactions are disrupted by intercalation of EtBr that affects DNA

Fig. 2. Effect of E2 on E1-ori complex formation in the absence or presence of ATP/Mg²⁺. The additions of E1 (increasing from 2, 12, 20 to 30 nM) and E2 (1.5 nM) are indicated above each lane. Selected reactions (4, 6 and 8, and 17, 19 and 21) were analysed following addition of polyclonal anti-E2 antiserum (lanes 11–13, and 24–26, respectively). The polyclonal antibodies generate diffuse supershifted bands.

structure (Schroter *et al.*, 1985; Lai and Herr, 1992). We used this property to recover selectively the E1-ori complex from mixtures of E1E2-ori and E1-ori complexes.

E2 stimulates formation of an E1-ori complex only in the presence of ATP/Mg²⁺

To determine if E2 could stimulate E1-ori formation, we performed binding reactions in the presence and absence of ATP/Mg²⁺ and increasing E1 concentration (Figure 2). Under both conditions, E1 bound cooperatively to the origin (Figure 2, lanes 2–5 and 15–18). At low E1 concentration in the presence of E2, the E1E2-ori complex formed (Figure 2, lanes 6 and 19). In the absence of ATP/Mg²⁺, E2 inhibited E1-ori complex formation to a significant extent (Figure 2, compare lanes 2–5 and 6–9). Addition of excess polyclonal anti-E2 antibody to a sample of reactions 4, 6 and 8 (Figure 2, lanes 11–13, respectively) demonstrated that when E2 was present, in the absence of ATP/Mg²⁺, the majority of the ori complexes contained E2. In clear contrast to the results described above, in the presence of ATP/Mg²⁺, E2 stimulated the formation of the E1-ori complex (Figure 2, lanes 15–18 compared with 19–22). All the E1E2-ori complex that formed in the presence of ATP/Mg²⁺ could be supershifted, generating a diffuse band in the agarose gel (Figure 2, lane 25). Surprisingly, only a minor fraction of the E1-ori complex formed in the presence of E2 and ATP/Mg²⁺ could be supershifted with excess anti-E2 antibodies (Figure 2, compare lane 21 with 26). This indicates that either E2 is hidden from the polyclonal antisera in this complex or, more likely, that E2 is turned over during complex formation. A similar observation has been made by Lusky *et al.* (1994). Therefore, E2 inhibits E1-ori formation in the absence of ATP/Mg²⁺ but stimulates E1-ori formation in the presence of ATP/Mg²⁺. Also, E2 is not stably maintained in the E1-ori complex whose formation it promotes.

The sequence-specific E1E2-ori complex is a preferred precursor for E1-ori

To test directly whether the E1E2-ori complex could act as a precursor for E1-ori formation *in vitro*, we designed

a protocol based on molecule tagging. The experimental design is outlined in Figure 3A. We first assembled a substrate E1E2-ori complex with E1 that had been tagged at the N-terminus with two copies of the hemagglutinin (HA) epitope (^{HA}E1E2-ori). These reactions were performed at high probe concentrations (0.4 nM) in the absence of ATP/Mg²⁺, with a 4- to 5-fold molar excess of ^{HA}E1 and E2 proteins. Under these conditions, the ^{HA}E1E2-ori complex is stable and the concentrations of free ^{HA}E1 and E2 protein are low. After a 20 min incubation, we added an excess of unlabelled specific competitor E2-binding site oligonucleotide (high affinity E2BS9; Li *et al.*, 1989) and diluted the reaction 15-fold into buffer containing nucleotide cofactor, non-specific competitor DNA and a 100-fold molar excess of untagged E1 (relative to tagged E1)—the ‘assembly reaction’. The E2 oligonucleotide competes with further E1E2-ori formation and acts as a sink for any template-bound E2 that dissociates. Products of the assembly reaction were removed at various times (T = 2.5, 5, 10, 17.5 and 25 min), cross-linked with glutaraldehyde, and analysed with or without specific antibodies on agarose gels. Chemical modification of E1 by glutaraldehyde instantly quenches further association of E1 with DNA, and stabilizes the protein–DNA complexes.

Figure 3B shows the time course of assembly of an E1-ori complex from the ^{HA}E1E2-ori substrate. Figure 3B, lanes 2–4 show E1-ori complex formation in the absence of E2 at low, intermediate and high E1 concentrations respectively. As shown in lane 8, in the presence of ^{HA}E1 and E2, the expected ^{HA}E1E2-ori complex formed, and this complex could be supershifted with excess anti-E1, anti-E2 and anti-HA epitope antibody 12CA5 (Figure 3B, lanes 5–7). Addition of the large excess of untagged E1 and nucleotide cofactor resulted in rapid formation of a complex that co-migrated with the slower migrating multimeric E1-ori complex (Figure 3B, lanes 9–13). Concomitant with formation of this complex, we observed a decrease in the amount of detectable E1E2-ori. Addition of anti-HA epitope antibody (Figure 3B, lanes 14–18) demonstrated that >85% of the slow migrating complex contained HA-tagged E1 at early time points (<5 min), whereas the proportion fell to 60–70% at later times

Fig. 3. Conversion of epitope-tagged $HA E1E2-ori$ complex to the multimeric $E1-ori$ complex. (A) Experimental design. The experiment is divided into three steps: (1) formation of epitope-tagged $HA E1E2-ori$ complex; (2) the assembly reaction itself; and (3) analysis of products by gel shift/supershift assay. Using a final ratio of tagged to untagged E1 of 1:100 (activity was determined to be approximately equivalent for the tagged and untagged proteins by titration, data not shown), the following predictions can be made. If the pre-formed $HA E1E2-ori$ complex is unable to act as a substrate for $E1-ori$ formation, the frequency at which epitope-tagged $E1-ori$ complexes appear would be determined by the input ratio of the two proteins and the order of the $E1-ori$ complex. For a tetrameric complex, ~4% of the products would be tagged and, likewise, for a hexameric complex, ~6% of the products would be tagged. If $HA E1E2-ori$ is able to act as a substrate for $E1-ori$ formation, with affinity similar to that of the free probe, the proportion of tagged $E1-ori$ complexes would be determined by the ratio of free probe to $HA E1E2-ori$ complex. If $HA E1E2-ori$ is a preferred substrate, the extent of $E1-ori$ complex formation would exceed that which occurs in the absence of a pre-formed $HA E1E2-ori$ complex. Also, the majority of complexes would be tagged. (B) Conversion of a pre-formed $HA E1E2-ori$ complex to the multimeric $E1-ori$ complex. Left: analysis of reaction products of the experiment outlined in (A). Lanes 1–8: products of the pre-incubation reaction in the absence of ATP/Mg²⁺ and control reactions. Lane 1, free probe; lanes 2–4, $E1-ori$ complex formation in the absence of E2; lanes 5–7, controls for antibody activity with $HA E1E2-ori$ substrate as indicated; lane 8, substrate $HA E1E2-ori$ complex at T = 0, after pre-incubation, before initiation of the assembly reaction. The following three groups of five lanes are time points T = 2.5, 5, 10, 17.5 and 25 min after initiation of the assembly reaction, analysed with or without specific antibodies, added after cross-linking. Right (lanes 24–35): control assembly reaction where E2 was omitted from the pre-incubation without ATP/Mg²⁺, but was included upon dilution and addition of untagged E1, ATP/Mg²⁺ and competitor DNAs. Lanes 24 and 25, markers for the $E1-ori$ complex and $E1E2-ori$ complex; lanes 26–30, time points as left, without antibody; lanes 31–35, analysis with anti-HA epitope antibody.

(Figure 3B, lane 18). Had the $E1-ori$ complex been formed independently of the tagged $HA E1E2-ori$ complex, only 4% of the complexes (assuming that the $E1-ori$ complex is a tetramer) would have incorporated tagged E1 and would be supershifted by this antibody. Therefore, the vast majority of the $E1-ori$ complexes that formed initially were formed from the $HA E1E2-ori$ complex, indicating that this is a good substrate for $E1-ori$ formation. The appearance of untagged complexes at later times could be the result of *de novo* $E1-ori$ formation from free probe;

alternatively, the tagged E1 molecules (from $HA E1E2-ori$ and $E1-ori$) could exchange with untagged E1 in solution, or become masked by further addition of E1. These possibilities could also account for the change with time in the relative proportions of the two anti-HA supershift species observed in Figure 3B, lanes 14–18. As expected, when ATP/Mg²⁺, ATP or Mg²⁺ alone were omitted from the complete assembly reaction, no $E1-ori$ complex formation was observed (not shown).

In a parallel control reaction, E2 was omitted from the

Fig. 4. DNase I footprint analysis of E1-ori and E1E2-ori complexes. (A) Analytical gel shift of binding reactions treated with nuclease and examined in the left panel below (E1-ori). Reaction 1 (lane 1) contained probe alone and reaction 2 (lane 2) received E2 (2 nM final concentration). Reactions 3–5 (lanes 3–5) all contained E1 at a final concentration of 12.5 nM. E1-ori complex formation in reaction 5 (lane 5) was stimulated ~5-fold by the presence of E2 (2 nM). All reactions contained ATP/Mg²⁺ except reaction 3. In lanes 6–8, the products of reactions 3–5 were analysed after addition of excess E2 antiserum. (B) Nuclease footprints derived from the binding reactions (1–5) shown above. In order to obtain footprints at low site occupancy (only 12 and 14% of the probe was bound in reactions 3 and 4, respectively), reaction volumes were large (up to 1 ml). After DNase I treatment and cross-linking, products were concentrated and then resolved on preparative agarose gels. For E1-ori (left), lanes 1 and 10 are sequence ladders, and lanes 2 and 9 are the DNase I cleavage products of free probe, with and without ATP/Mg²⁺. Lanes 4 and 5 show footprints for the complex formed in the presence of ATP/Mg²⁺ without E2 (reaction 4), cross-linked before (lane 4) or after (lane 5) digestion with DNase I. Lane 6 is the footprint of the E1-ori complex whose formation was stimulated by E2 (reaction 5); products were cross-linked after DNase I digestion. Lanes 7 and 8 are footprints for the complex formed in the absence of ATP/Mg²⁺ (reaction 3), cross-linked after (lane 7) or before (lane 8) digestion with DNase I. Boundaries of the core and flanking protections are indicated, as are the positions of the A/T-rich region and E1/E2-binding sites. BPV-1 nucleotide co-ordinates of the minimal ori sequence are given on the far left. For the E1E2-ori footprint (centre), binding reactions contained a low concentration of E1 and E2 which give no detectable binding alone (lanes 13 and 14, respectively). Cleavage products for the total E1–E2 binding reaction are shown in lane 15, and the footprint of the isolated E1E2-ori complex in lane 16. Lanes 11 and 18 are sequence ladders, and lanes 12 and 17 are the DNase I cleavage products of free probe. Lane 21 shows E2 binding to the low affinity E2-binding site (lane 20, free probe; lane 19, G ladder). The E2 protection (hatched box) extends over 23 nucleotides from BPV nucleotide +8 to +31. At this high concentration of E2 (500 nM), required to observe protection of the E2 site, there is also evidence of under-cleavage of adjacent sequences, suggesting that more than one E2 dimer may be bound cooperatively to a proportion of the DNA templates.

pre-incubation of probe with tagged E1, but was added upon dilution and addition of untagged E1, ATP/Mg²⁺ and competitor DNAs (Figure 3B, lanes 26–35, right). Lanes 24 and 25 are markers for the E1-ori complex and the E1E2-ori complex respectively. The extent of E1 complex formation from free probe was clearly lower (~3-fold in this case; Figure 3B, lanes 26–30) in this reaction compared with that described above, where a high proportion of substrate was a pre-formed ^{HA}E1E2-ori complex. Also, the complex that formed could not be supershifted with anti-HA antibody (Figure 3B, lanes 31–35). Therefore, the large excess of untagged E1 protein (100-fold) results in an undetectable level of tagged E1-ori complex formation in these assays. Together, the results described above demonstrate that a pre-formed ^{HA}E1E2-ori complex is a preferred substrate for E1-ori complex formation compared with free probe. Further-

more, the E1 molecules of precursor ^{HA}E1E2-ori were incorporated into the resulting E1-ori complex.

To determine the fate of E2 in these reactions, we performed supershift assays using antibody directed against E2 (Figure 3B). At the earliest times (<5 min), up to 50% of the slow migrating E1-ori-like complex could be supershifted with anti-E2 antibody (Figure 3B, lane 19). By 25 min, however, little association of E2 with the E1-ori complexes could be detected (Figure 3B, lane 23). This suggests that E2 dissociates from the nascent E1-ori complex. Indeed, nuclease footprinting showed no evidence that E2 was bound to its site in these complexes (data not shown, but see Figure 4B). Like the E1-ori complex that formed at high E1 concentrations in the absence of E2, the E1-ori complex formed via E1E2-ori is stable in the presence of EtBr. Interestingly, the slow migrating E2-containing complexes apparent at early times

(Figure 3B, lanes 19–23) also do not dissociate in the presence of EtBr (data not shown). These are most likely intermediates in formation of the multimeric E1–*ori* complex, where E2 is cross-linked to E1 that it initially recruited to *ori*.

E1–ori complexes formed in the presence or absence of E2 generate identical nuclease footprints

To determine if E1 was similarly disposed in E1–*ori* complexes formed with or without E2, we footprinted complexes with DNase I (Figure 4). We compared E1–*ori* complexes formed at low E1 concentration in the presence and absence of ATP/Mg²⁺ with E1–*ori* complexes formed in the presence of E2. Also, we analysed complexes that were cross-linked before treatment with nuclease. To obtain solution footprints at low site occupancy, we isolated protein–DNA complexes by gel electrophoresis after digestion with DNase I. Cleavage was terminated by adding EDTA and glutaraldehyde. To sequester any E2-containing complexes from the multimeric E1–*ori* complex, excess anti-E2 antiserum was added prior to electrophoresis. Bound and free probe were then separated by agarose gel electrophoresis and the respective protein–DNA complexes were recovered after blotting to nitrocellulose. Figure 4A shows an analytical gel shift of the binding reactions treated with nuclease. E1 alone, in the absence and presence of ATP/Mg²⁺, gave rise to minimal complex formation (Figure 4A, lanes 3 and 4). Upon addition of E2, formation of the E1–*ori* complex was stimulated (Figure 4A, lane 5). In Figure 4A, lanes 6–8, anti-E2 antibody was added to a sample of reactions 3–5.

The nuclease protections (bottom strand) derived from the binding reactions in Figure 4A are shown in Figure 4B, left panel (labelled E1–*ori*). The centre panel shows the footprint of the E1E2–*ori* complex for comparison (labelled E1E2–*ori*). The footprint of E2 bound to the low affinity E2-binding site is also shown on the right (labelled E2). Importantly, the footprints for the E1–*ori* complex formed in the presence of ATP/Mg²⁺, without and with E2 (Figure 4B, lanes 5 and 6), are indistinguishable. On this strand, E1 protects a 29 nucleotide ‘core’ sequence (BPV nucleotides 7934 to +15) over the E1 recognition site. In the flanking regions, extending 18 nucleotides upstream and 19 downstream (to nucleotide +34) of the core, some positions are completely protected, while others only partially. There is also evidence of hypersensitivity in these regions (labelled Hyp.). On this strand, the E1E2–*ori* protection extends from nucleotide 7935 to +31 (Figure 4B, lane 16). The downstream boundary of this protection is the same as when E2 alone is bound to the probe (Figure 4B, lane 21). Importantly, comparing the E1–*ori* complex protections over the E2 recognition sequence (and beyond) with the corresponding region in the E1E2–*ori* complex (Figure 4B, centre panel, lane 16, compared with left panel, lane 6) confirms that E2 does not occupy the E2-binding site in the majority of E1–*ori* complex whose formation it stimulates [Figure 4B, compare the three bracketed regions (1–3) in the left and centre panels]. Similar results were obtained for the other DNA strand (top, data not shown). On this strand, the core E1–*ori* protection encompasses 28 nucleotides

from nucleotide 7938 to +18. The flanking protections extend 19 nucleotides downstream and 21 nucleotides upstream over the A/T-rich region.

The E1–*ori* complex formed in the absence of ATP/Mg²⁺ (Figure 4B, lane 7) was virtually indistinguishable from the complex formed in the presence of ATP/Mg²⁺ (Figure 4B, lanes 5 and 6), suggesting that E1 is bound to *ori* in the same way in all complexes. The only significant change in the E1–*ori* footprints was observed when complexes were treated with nuclease after cross-linking (Figure 4B, lanes 4 and 8, left). The core protection is preserved; however, for the cross-linked complex formed in the absence of ATP/Mg²⁺, flanking protections appear to be lost. Similar results were also obtained for the other strand (data not shown).

These results indicate that the E1–*ori* complexes formed with or without E2 are identical, implying that E2 has a purely quantitative effect on E1–*ori* complex formation. Also, there is good evidence that E2 does not remain associated with its binding site in this complex. Together, these observations are consistent with a catalytic role for E2 in complex formation. In addition, differences between the nuclease footprints of complexes formed in the presence and absence of ATP/Mg²⁺ become significant only after cross-linking. This suggests that the interactions of E1 with itself and/or DNA are altered only in a subtle fashion when cofactor is present. However, these subtle changes in E1 that occur in the presence of nucleotide cofactor are necessary for E2-mediated E1–*ori* formation.

A transient and specific E2–DNA interaction stimulates formation of the multimeric E1–ori complex

The observations described above suggested that E2 is absent from the final E1–*ori* complex although the formation of this complex was stimulated by E2. A likely possibility is that E2 is bound to its site only transiently during complex assembly. To address this question, we used a chemical modification interference assay (Figure 5). As illustrated in Figure 5A, interference analysis would reveal sequences that are required for complex formation, even if they are only transiently bound by a protein. Preparative binding reactions were assembled on DNA modified at phosphate groups with *N*-ethylnitrosourea, and complexes were recovered after cross-linking and electrophoresis. The results are shown in Figure 5B, and the base positions that give rise to interference when modified are indicated in Figure 5C. For the E1E2–*ori* complex (Figure 5, lanes 3 and 4 bottom strand, 12 and 13 top strands), there is strong interference over the E1-binding site at four phosphates on the bottom (nucleotides 8–10 and 7946) and top (nucleotides 7941–7943 and 5) strands (indicated with filled triangles in Figure 5C). The interference over the E2-binding site is similar to that observed with E2 alone, as shown in Figure 5C. Several positions over the E1-binding site also result in minor interference when modified, and there are no significant cofactor-dependent differences. For the E1–*ori* complex that formed without E2 (Figure 5B, lanes 6 and 7, and 15 and 16), strong interference is seen at the same three contiguous positions as in the E1E2–*ori* complex, but the strong interference extends by an additional three phosphates in the 5′ direction on the bottom strand

Fig. 5. *Ori* complex ethylation interference analysis. (A) Principle of the interference assay. For a DNA sequence with two binding sites X and Y, where protein *x* stimulates binding of *y* via its site X, interference should be observed over both binding sites X and Y, regardless of whether *x* is maintained in stable complex with *y* on the DNA. (B) Interference for bottom (left) and top (right) strands. The complex from which the ethylated DNA was recovered and conditions under which it was formed are indicated above. E1+E2 indicates the reaction where E2 stimulated formation of the E1-*ori* complex in a similar way to that illustrated in Figure 4A. (C) Nucleotide positions that give rise to interference when phosphate groups are ethylated. Light triangles are for the E1-*ori* complex, filled triangles are for E1E2-*ori*, and grey triangles indicate the interference for E2 binding to the low affinity E2 site. The size of the triangles is a relative reflection of the degree of interference. The E1- and E2-binding sites are boxed.

(nucleotides 11–13; Figure 5B, compare lanes 3 and 4 with lane 7), and the 3' direction on the top strand (nucleotides 7944–7946; Figure 5B, compare lanes 12 and 13 with 16). These positions are indicated with open triangles. This pattern of extended interference over the E1-binding site in E1-*ori* compared with E1E2-*ori* is similar to a pattern of extending hydroxy radical protections previously described (Sedman and Stenlund, 1996). These data were interpreted as reflecting the binding of successive E1 molecules to nested sequences in the E1 recognition site. Only quantitative differences were observed between the interference patterns for complexes formed in the presence or absence of ATP/Mg²⁺ (Figure 5, lanes 6 and 7 bottom strand, 15 and 16 top strand).

The principal reason for employing the interference assay was to determine whether an E2–DNA interaction was required for stimulation of E1-*ori* complex formation by E2. In reactions where E2 was used to stimulate the formation of E1-*ori*, the interference pattern is a combination of that observed for the E1-*ori* and E1E2-*ori* complexes described above (Figure 5, centre lanes 5 and 14, top and bottom strands, respectively). The complex was recovered after treatment with EtBr, so the E2 interference could not have arisen from contaminating E1E2-*ori* complex. Interference over the E2-binding site demonstrates that binding of E2 to its recognition sequence is required for formation of the E1-*ori* complex.

E2 stimulates formation of a complex active for origin distortion

Potassium permanganate (KMnO₄) has been used to probe DNA for open complex formation with transcription or replication proteins (see Sasse-Dwight and Gralla, 1991, and references therein). The reagent reacts primarily with T residues in single-stranded or sharply distorted DNA, resulting in modified bases where specific strand scission can occur. BPV-1 E1 has been shown previously to induce local ATP/Mg²⁺-dependent *ori* distortion using this method (Gillette *et al.*, 1994). To test if the E1-*ori* complex formed via E2 was active for local origin distortion, we formed E1-*ori* in the absence and presence of E2 and treated binding reactions directly with KMnO₄ (Figure 6). For each binding reaction, a small sample of the products was removed prior to addition of KMnO₄, cross-linked and processed for gel shift analysis (Figure 6A) in parallel with the modification reaction (Figure 6B). Anti-E2 antibody was added to all gel shift samples. This allowed a determination of the stimulatory effect of E2 on E1-*ori* formation, and confirmed the absence of E2 from the majority of the E1-*ori* complexes. The lane numbers for the KMnO₄ modification assay correspond to those for the gel shift shown above.

Both the degree of distortion and the fraction of probe bound increased with E1 protein concentration (Figure 6, lanes 1–4). Higher E1 protein concentrations than those shown resulted in little further increase in the degree of measurable distortion (data not shown). No hyperreactive thymidines were detected in the E1E2-*ori* complex assayed under similar conditions (data not shown). The degree of KMnO₄ sensitivity was stimulated by the addition of E2 (Figure 6, compare lane 3 with lanes 7–11 which all contained the same concentration of E1). For example, both E1-*ori* formation and the number of tem-

Fig. 6. E2 stimulates E1-dependent unwinding of the origin of replication. **(A)** Gel shift analysis of binding reactions. A small fraction of the binding reactions was removed, cross-linked and analysed after addition of excess anti-E2 antiserum. Reactions 1–4 (lanes 1–4), E1 alone, 50, 30, 17.5 and 10 nM final concentration, respectively. Reaction 5 (lane 5) is probe alone. In reactions 6–11 (lanes 6–11), E1 was at a concentration of 17.5 nM (corresponding to a 175-fold molar excess over probe) and E2 increasing from 0.25, 0.5, 1, 2, 3 to 4 nM. Lane 12, E2 alone (4 nM). SS indicates products supershifted with anti-E2 antibody. **(B)** KMnO₄ modification assay (top strand) for the reactions analysed above by gel shift. The remaining fraction of the binding reactions was treated directly with KMnO₄ in solution. Lanes 1–4: *ori* unwinding as a function of E1 protein concentration (reactions 1–4). The fraction of bound probe in each reaction was 100, 92, 30 and 9% respectively [from (A) above]. Lane 5, reactivity of free probe; lanes 6–11, stimulation of unwinding by E2 at low E1 concentration. For the reaction analysed in lane 9, 88% of the probe was bound, and <12% of the complex could be shifted by E2 antibody [from (A) above]. E2 concentrations in excess of those required to stimulate binding of E1 to all the probe also resulted in a modest increase in unwinding activity (lanes 10 and 11). E2 alone (or the E1E2-*ori* complex, not shown) induces no *ori* unwinding (lane 12). Nucleotide coordinates for major sites of modification are given on the left (BPV-1 genome). Positions of relevant *ori* sequences are shown on the right; the E2 site is indicated by the shaded box, and the white box indicates the E1-binding site.

plates modified by KMnO₄ were stimulated 3-fold by the addition of E2 in reaction 9 relative to reaction 3. These results strongly suggest that the stimulatory effect of E2 on *ori* distortion is due to stimulation of E1-*ori* complex formation, rather than a direct effect of E2 on the distortion process. However, distortion was not directly proportional to the fraction of probe bound. For example, in reaction 2, the E1-*ori* complex which was formed at high E1 concentration in the absence of E2 showed a 3-fold greater

Fig. 7. Proposed model for assembly of the BPV-1 initiator complex. E1E2-*ori* acts as a 'seed' for generation of an active initiator complex. Details are given in the text.

extent of KMnO_4 modification compared with reaction 9, where the extent of binding was similar. We observe similar results over a wide range of probe concentrations tested (0.005–0.4 nM), indicating that formation of a complex with high distortion activity is determined by the absolute E1 concentration. Together, these results indicate that qualitative differences may exist between complexes formed at different E1 concentrations. One possible explanation is suggested by the observation that the electrophoretic mobility of the E1-*ori* complex appears to decrease when the complex is formed at high E1 concentrations (Figure 6A, lane 1). More E1 molecules could be bound to *ori* when *ori* distortion is greatest as measured by the KMnO_4 modification assay.

Discussion

E1E2-*ori* 'seeds' formation of the active E1-*ori* complex

The BPV-1 transcription factor E2 acts transiently to stimulate formation of an E1 initiator complex lacking E2. In Figure 7, we present a model for assembly of the BPV-1 initiator complex at the minimal origin of replication. E2 confers specificity and affinity on E1 DNA binding in the form of the E1E2-*ori* complex and thus serves as a 'loading factor' for the initiator. Once formed, ATP-dependent conformational changes in E1 destabilize the E1E2-*ori* complex and facilitate transition to a multimeric E1-*ori* complex via an E2-containing intermediate. E1 from the precursor E1E2-*ori* complex is incorporated into E1-*ori* and E2 is displaced. In a sense, the action of E2 can be considered catalytic, having a quantitative but not qualitative effect on E1-*ori* complex formation. Initial tethering of E1 to *ori* by E2 overcomes a major rate-limiting step (stable and specific E1 binding) in initiation. Rapid and preferential conversion to E1-*ori* of the E1E2-*ori* complex compared with naked DNA is clearly observed at early times in our assembly experiment (Figure 3).

Once released from the nascent E1-*ori* complex, E2 is presumably free to initiate a further cycle of complex assembly. This mechanism results in the formation of a complex with little inherent sequence specificity at a specific site (BPV *ori*). We hypothesize that this E1-*ori* complex is in turn a substrate for formation of a higher order E1 complex that is capable of processive DNA unwinding (helicase activity), as indicated schematically

in Figure 7. Recent studies have identified a hexameric form of E1 which binds DNA in a sequence-independent manner and has DNA helicase activity (Sedman and Stenlund, 1998). At this point, the relationship between the helicase and the complexes described here is unclear. However, a reasonable suggestion would be that formation of the helicase entails sequential addition of E1 molecules to an initial sequence-specific *ori* recognition complex. As more E1 molecules bind to nested recognition sites, specificity for *ori* sequences and stabilizing contacts with DNA decrease. A continuous ring of protein-protein contacts may determine the overall stability of the final complex.

Regulation of complex formation by ATP

The biochemical analysis of E1 complexes indicates that the E1E2-*ori* complex is the form of E1 best suited for *ori* recognition. This complex binds DNA with high sequence specificity and affinity. Binding of this form is destabilized by the presence of ATP, resulting in a complex with a shorter half-life. Hydrolysis of bound ATP is likely to be required, since we observe formation of E1E2-*ori* but not conversion to E1-*ori* in the presence of non-hydrolysable ATP analogues (C.M.Sanders and A.Stenlund, unpublished observation).

The mechanism of action of ATP in the conversion of the E1E2-*ori* complex to the E1-*ori* complex is not clear, but it is likely that ATP binding and hydrolysis by E1 are involved. ATP/Mg^{2+} may affect E1 binding directly, favouring formation of the E1-*ori* complex. However, the presence of ATP/Mg^{2+} affects the extent of E1 complex formation in the absence of E2 in only a very minor way (Figure 1B). In addition, the nuclease footprinting and interference analysis show that the E1-*ori* complexes formed in the absence or presence of ATP/Mg^{2+} are identical at the level of DNA binding, indicating that the E1 molecules are arranged similarly on the *ori* irrespective of the conditions under which they were formed (Figures 4 and 5). It is more likely that ATP binding by E1 affects the interaction between E1 and E2, allowing displacement of E2 once it has fulfilled its role as a loading factor. Consistent with this idea, it is only in the presence of E2 that ATP/Mg^{2+} has a substantial effect on the formation of the E1-*ori* complex (Figure 2). Furthermore, in the absence of ATP/Mg^{2+} , E2 is inhibitory for E1-*ori* complex formation, consistent with a role for ATP/Mg^{2+} in displace-

ment of E2. The energy of nucleotide hydrolysis may be required to displace E2, which otherwise inhibits further assembly of the initiator complex.

E2 as a loading factor

The properties of the process that we have described, i.e. the contribution to specificity by E2, the transient requirement for E2 and the ATP-dependent displacement of E2, show interesting similarities to other processes. Several components of the cellular DNA replication machinery from various organisms utilize factors that assist in assembly and targeting. The P-protein from phage λ functions to load the DnaB helicase at the λ *ori*, DnaC assists in loading of DnaB at *oriC* (Wahle *et al.*, 1989; Marians, 1992). The γ -complex clamp-loader in *E. coli*, and RFC in eukaryotes, function to load the β -clamp and proliferating cell nuclear antigen (PCNA), respectively, at the replication fork (Tsurimoto and Stillman, 1991; Kelman and O'Donnell, 1995). In these examples, the loading factors, just as E2, assist in targeting and assembly since the loaded factors lack sufficient specificity to determine their final destination. Another shared feature is the requirement for ATP binding and hydrolysis as part of the loading cycle. Release of the β -clamp by the γ -complex requires ATP hydrolysis, and release of DnaB from DnaC is likely to require ATP hydrolysis (Wahle *et al.*, 1989; Kelman and O'Donnell, 1995). Similarly, our data indicate that ATP is required for release of E2 from E1. An interesting difference, however, is that the cellular loading factors are responsible for ATP hydrolysis while, in the case of E2, ATP binding and hydrolysis are probably performed by the loaded factor E1. A third feature that the factors in these examples share is that they form complex structures with an unusual relationship to their substrates. Both the β -clamp and PCNA encircle DNA (Kong *et al.*, 1992; Krishna *et al.*, 1994). This is likely to be the case also for DnaB helicase, which belongs to the class of hexameric helicases that have been suggested to encircle their substrates (Egelman, 1996). E1 can also form a hexamer with helicase activity, which is likely to encircle DNA (Sedman and Stenlund, 1998).

Thus, key aspects of the loading process for these different factors are shared. A physical association between the two factors results in a complex that is positioned by virtue of the specificity of the loading factor. ATP-induced changes in one of the two factors results in abrogation of interaction and eventual removal of the loading factor. We believe that these similarities are consistent with a common theme for positioning of factors with limited capacity for independent localization. However, the targets differ: in the case of DnaB, the recognition is at the level of protein-protein interactions since the DnaB-DnaC complex and λ P recognize DnaA and λ O, bound to their respective *oris* (Marians, 1992). For loading of the β -clamp and PCNA, the determinant is DNA structure; primer-template junctions are recognized and bound by the γ -complex and RFC, while target recognition by E2 is at the level of DNA sequence (Tsurimoto and Stillman, 1991; Kelman and O'Donnell, 1995). Curiously, in spite of the importance and prevalence of proteins with sequence-specific DNA recognition, to our knowledge, loading factors that recognize specific DNA sequences have not been described previously.

Transcription factors in loading

The results that we have presented apply specifically to the assembly of an E1 complex that can distort the BPV origin of replication. The use of a loading factor for E1 positioning may represent a special case since the biochemical activities of the final E1 complex may be incompatible with sequence-specific DNA binding. However, the requirements for generation of the E1-*ori* complex presumably are not substantially different from those for the assembly of other protein complexes on DNA. The process that we have described here could serve as a general mechanism to specify the positioning on DNA of proteins or protein complexes with little intrinsic specificity, just as RFC can position PCNA to a specific structure. We have been able to isolate the precursor E1E2-*ori* complex and observe its conversion to the E1-*ori* complex due to the dependence on ATP/Mg²⁺ for the conversion. However, because the requirement for E2 is transient, a factor acting like E2 in other systems could easily escape detection, and it is therefore plausible that similar examples may exist. In general, the concept of 'seeding' a complex is a reasonable model for formation of protein complexes on DNA, and some evidence exists that complexes such as ORC may serve as 'landing pads' for other factors required for initiation of DNA replication (Diffley *et al.*, 1994).

It is interesting that the viral E2 protein which appears to function as a loading factor in the E1 assembly process is a prototypic transcription factor (McBride *et al.*, 1991). Many different biochemical functions have been attributed to transcription factors for both transcription and DNA replication, but it seems likely that a major activity involves the recruitment of other factors to specific sites on DNA. An interesting proposition is that transcription factors in general can provide their interaction partners with the targeting activities we have described for E2. A role for transcription factors in binding of complexes such as ORC and TFIID to DNA similar to the one we have observed for E1 seems plausible. In addition to providing greater target specificity, such a mechanism would also provide regulatory opportunities for spatial and temporal control of both DNA replication and transcription through control of the transcription factor.

Materials and methods

Viral proteins

Expression and purification of E1 and E2 was, essentially, as described previously (Sedman *et al.*, 1997). To tag E1, two copies of the influenza virus HA epitope coding sequence were inserted upstream of the E1 ATG initiation codon (5'-CTAGCTATCCTTAATGACGTGCCTGAC-TATGCCAGCCTGGGAGGACCTT; annealing to the complementary oligonucleotide generates *Xba*I overhangs at each end).

Antibodies

Anti-HA epitope antibody (12CA5) and anti-E1 (104) are mouse monoclonals. Anti-E2 is polyclonal rabbit antiserum raised to the bacterially expressed protein. All were purified on protein A-Sepharose. For electrophoretic mobility shift supershift assays, antibodies were added to a predetermined excess.

Origin template

The minimal origin template (BPV-1 coordinates 7914-27) in pUC19 has been described previously (Sedman and Stenlund, 1995; Sedman *et al.*, 1997). The construct was modified by digestion with *Kpn*I, end-

polishing with T4 DNA polymerase and re-ligation. This removes a potential E2 half-site in the upstream polylinker sequence.

DNA binding assays

Probes were generated by PCR using ^{32}P end-labelled primers, normally USP (5'-GTAAACGACGGCCAGT) and RSP (5'-GGATAACAATTT-CACACAGG), flanking the pUC19 multiple cloning site. The specific product was purified by polyacrylamide gel electrophoresis and recovered by diffusion elution.

Binding reactions were performed in 20 mM sodium phosphate buffer pH 7.2, 100 mM NaCl, 1 mM EDTA, 0.1% NP-40, 1 mM dithiothreitol, 10% (v/v) glycerol and 0.1 mg/ml bovine serum albumin (BSA). Standard binding reactions contained high molecular weight competitor DNA poly(dA-dT)_n at 125 µg/µl. This low concentration of competitor reduces specific binding of E1 to DNA by ~2-fold. The probe concentration was 0.05 nM for gel shifts, but was increased for some footprinting procedures (see below). For the assembly experiment described in Figure 3, substrate ^{HA}E1E2-ori was formed at a probe concentration of 0.4 nM. Where ATP and magnesium chloride were included, both were at 5 mM. Incubations were at 20–22°C for 30 min unless otherwise stated. To stabilize complexes for gel shift analysis, glutaraldehyde was added to 0.08% (w/v) and cross-linking was terminated after 5 min by addition of Tris-HCl (pH 7.5) to 80 mM. Complexes were resolved on 1% agarose gels in Tris-acetate buffer. For supershift assays, antibodies were added after cross-linking. When complexes were analysed in the presence of EtBr, it was added to 25 µg/ml.

Competitor E2-binding site oligonucleotide was generated by annealing complementary synthetic oligonucleotides, followed by purification on acrylamide gels. The sequence used was high affinity E2 site BS9 as underlined (5'-ACAAAGTACCGTTGCCGTCGCGGGTC).

DNase I footprinting

DNase I footprinting was performed as described previously (Sedman and Stenlund, 1995), except that 2× cofactor salt solution contained 100 mM NaCl and 10% (v/v) glycerol in addition to Mg^{2+} (10 mM) and Ca^{2+} (5 mM) ions. For the experiment described in Figure 4, the concentration of probe was increased to 0.2 nM. To obtain footprints of native complexes at protein concentrations lower than those required to bind all the probe, the protein-DNA complexes were isolated after DNase I treatment. Digestion was terminated and complexes stabilized simultaneously by addition of EDTA to 20 mM and glutaraldehyde as above; free probe and protein-DNA complexes were then resolved by gel electrophoresis. Footprints of cross-linked E1-ori complexes were generated by DNase I digestion after treatment with glutaraldehyde. When site occupancy was low, reaction volumes were increased up to 1 ml. Following DNase I treatment and cross-linking, these reactions were concentrated (~10-fold) using Centricon 30 devices (Amicon, Beverly, MA). Reaction products were then resolved on agarose gels and protein-DNA complexes transferred to nitrocellulose by capillary blotting. After localizing bands by autoradiography, products were released by soaking membrane pieces in 20 mM Tris-HCl pH 8.0, 1 mM EDTA and 1% SDS. DNA was recovered by ethanol precipitation following phenol-chloroform extraction. As for ethylation interference and potassium permanganate modification assays (below), scission products were analysed on 8% urea-acrylamide gels. Sequence ladders (G and A>C) were generated by standard chemical sequencing techniques (Maxam and Gilbert, 1977).

Ethylation interference

Detailed procedures for ethylation interference can be found in Wissmann and Hillen (1991). Briefly, 0.5–1 µg of end-labelled probe was dissolved in 200 µl of cacodylate buffer (50 mM NaCac, 1 mM EDTA, pH 7.0), and an equal volume of a saturated solution of *N*-ethylnitrosourea (Sigma) in ethanol was added. The reaction was incubated at 50°C for 60 min, chilled on ice and the DNA precipitated after addition of 20 µl of 3 M sodium acetate pH 7.0 and 2 vols of ethanol. The DNA was then dissolved and re-precipitated from 300 mM sodium acetate, the pellet rinsed in ethanol, dried and dissolved in 10 mM sodium phosphate (pH 7.0), 50 mM NaCl, 0.1 mM EDTA. Protein-DNA complexes were formed with modified probe at a final concentration of 0.1 nM and cross-linked with glutaraldehyde. As with the DNase I analysis described above, large reaction volumes were concentrated prior to resolution on agarose gels and recovery of DNA from specific complexes following blotting to nitrocellulose. To obtain the interference for E2 binding, protein-DNA complexes were resolved on acrylamide gels (Li *et al.*, 1991) and the modified DNA recovered by electroelution.

Recovered DNA was dissolved in 30 µl of 10 mM sodium phosphate

(pH 7.0), 1 mM EDTA, and 5 µl of 1 M NaOH was added. Cleavage at modified phosphates was achieved by incubation at 90°C for 30 min. The solution was neutralized with hydrochloric acid and the DNA recovered by precipitation.

Potassium permanganate modification assay

Binding reactions with end-labelled probe at a final concentration of 0.1 nM were assembled and incubated as described above. For the modification reaction, KMnO_4 was added to a final concentration of 6 mM and reactions incubated for a further 2 min. Modification was terminated by adding β-mercaptoethanol to 80 mM, SDS to 0.3% and EDTA to 10 mM. Reactions were then digested with proteinase K (20 µg/ml) for 60 min at 37°C and DNAs recovered by phenol-chloroform extraction and ethanol precipitation. Cleavage at modified bases was achieved with piperidine (30 min at 90°C).

Acknowledgements

We thank Nouria Hernadez, Winship Herr, Bill Henry and Grace Chen for critical reading of the manuscript. We would also like to thank Masafumi Tanaka for his enthusiastic comments. This work was supported by a grant NIH CA 13106 to A.S. C.M.S. was supported by a Wellcome International Prize Travelling Fellowship.

References

- Bergsma,D.J., Olive,D.M., Hartzell,S.W. and Subramanian,K.N. (1982) Territorial limits and functional anatomy of the simian virus 40 replication origin. *Proc. Natl Acad. Sci. USA*, **79**, 381–385.
- Chen,G. and Stenlund,A. (1998) Characterisation of the DNA-binding domain of the bovine papillomavirus replication initiator E1. *J. Virol.*, **72**, 2567–2576.
- DePamphilis,M.L. (1993) Eukaryotic DNA replication: anatomy of an origin. *Annu. Rev. Biochem.*, **62**, 29–63.
- DePamphilis,M.L. (1996) Origins of DNA replication. In DePamphilis,M.L. (ed.), *DNA Replication in Eucaryotic Cells*. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, pp. 45–86.
- de Villiers,J., Schaffner,W., Tyndall,C., Lupton,S. and Kamen,R. (1984) Polyoma virus DNA replication requires an enhancer. *Nature*, **312**, 242–246.
- Diffley,J.F.X., Cocker,J.H., Dowell,S.J. and Rowley,A. (1994) Two steps in the assembly of complexes at yeast replication origins *in vivo*. *Cell*, **78**, 303–316.
- Egelman,E.H. (1996) Homomorphous hexameric helicases: tales from the ring cycle. *Structure*, **4**, 759–762.
- Gillette,T.G., Lusky,M. and Borowiec,J.A. (1994) Induction of structural changes in the bovine papillomavirus type 1 origin of replication by the viral E1 and E2 proteins. *Proc. Natl Acad. Sci. USA*, **91**, 8846–8850.
- Guo,Z.-S., Gutierrez,C., Heine,U., Sogo,J.M. and DePamphilis,M.L. (1989) Origin auxiliary sequences can facilitate initiation of simian virus 40 replication *in vitro* as they do *in vivo*. *Mol. Cell. Biol.*, **9**, 3593–3602.
- Jacob,F., Brenner,S. and Cuzin,F. (1963) On the regulation of DNA replication in bacteria. *Cold Spring Harbor Symp. Quant. Biol.*, **28**, 329–348.
- Kelman,Z. and O'Donnell,M. (1995) DNA polymerase III holoenzyme: structure and function of a chromosomal replicating machine. *Annu. Rev. Biochem.*, **64**, 171–200.
- Kong,X.P., Onrust,R., O'Donnell,M. and Kuriyan,J. (1992) Three-dimensional structure of the β subunit of *E.coli* DNA polymerase III holoenzyme: a sliding DNA clamp. *Cell*, **69**, 425–437.
- Krishna,T.S., Kong,X.P., Gary,S., Burgers,P.M. and Kuriyan,J. (1994) Crystal structure of the eukaryotic DNA polymerase processivity factor PCNA. *Cell*, **79**, 1233–1243.
- Lai,J.-S. and Herr,W. (1992) Ethidium bromide provides a simple tool for identifying genuine DNA-independent protein associations. *Proc. Natl Acad. Sci. USA*, **89**, 6958–6962.
- Li,R., Knight,J., Bream,G., Stenlund,A. and Botchan,M.R. (1991) Specific recognition nucleotides and their context determine the affinity of E2 protein for 17 binding sites in the BPV-1 genome. *Genes Dev.*, **3**, 510–526.
- Lusky,M., Hurwitz,J. and Seo,Y.-S. (1993) Cooperative assembly of the bovine papilloma virus E1 and E2 proteins on the replication origin requires an intact E2 binding site. *J. Biol. Chem.*, **268**, 15795–15803.

- Lusky, M., Hurwitz, J. and Seo, Y.-S. (1994) The bovine papillomavirus E2 protein modulates the assembly of but is not stably maintained in a replication-competent multimeric E1–replication origin complex. *Proc. Natl Acad. Sci. USA*, **91**, 8895–8899.
- Marians, K.J. (1992) Prokaryotic DNA replication. *Annu. Rev. Biochem.*, **61**, 673–719.
- Maxam, A.M. and Gilbert, W. (1977) A new method for sequencing DNA. *Proc. Natl Acad. Sci. USA*, **74**, 560–564.
- McBride, A.A., Romanczuk, H. and Howley, P.M. (1991) The papillomavirus E2 regulatory proteins. *J. Biol. Chem.*, **266**, 18411–18414.
- Mohr, I.J., Clark, R., Sun, S., Androphy, E.J., MacPherson, P. and Botchan, M.R. (1990) Targeting the E1 replication protein to the papillomavirus origin of replication by complex formation with the E2 transactivator. *Science*, **250**, 1694–1699.
- Sasse-Dwight, S. and Gralla, J.D. (1991) Footprinting protein–DNA complexes *in vivo*. *Methods Enzymol.*, **208**, 147–169.
- Schroter, H., Maiet, G., Ponsting, H. and Nordheim, A. (1985) DNA intercalators induce specific release of HMG 14, HMG 17 and other DNA-binding proteins from chicken erythrocyte chromatin. *EMBO J.*, **4**, 3867–3872.
- Sedman, J. and Stenlund, A. (1995) Co-operative interactions between the initiator E1 and the transcriptional activator E2 is required for replicator specific DNA replication of bovine papillomavirus *in vivo* and *in vitro*. *EMBO J.*, **14**, 6218–6228.
- Sedman, J. and Stenlund, A. (1996) The initiator protein E1 binds to the bovine papillomavirus origin of replication as a trimeric ring-like structure. *EMBO J.*, **15**, 5085–5092.
- Sedman, J. and Stenlund, A. (1998) The papillomavirus E1 protein forms a DNA-dependent hexameric complex with ATPase and DNA helicase activities. *J. Virol.*, **72**, 6893–6897.
- Sedman, T., Sedman, J. and Stenlund, A. (1997) Binding of the E1 and E2 proteins to the origin of replication of bovine papillomavirus. *J. Virol.*, **71**, 2887–2896.
- Seo, Y.-S., Müller, F., Lusky, M., Gibbs, E., Kim, H., Phillips, B. and Hurwitz, J. (1993a) Bovine papilloma virus (BPV)-encoded E2 protein enhances binding of E1 protein to the BPV replication origin. *Proc. Natl Acad. Sci. USA*, **90**, 2865–2869.
- Seo, Y.-S., Müller, F., Lusky, M. and Hurwitz, J. (1993b) Bovine papilloma virus (BPV)-encoded E1 protein contains multiple activities required for BPV DNA replication. *Proc. Natl Acad. Sci. USA*, **90**, 702–706.
- Tsurimoto, T. and Stillman, B. (1991) Replication factors required for SV40 DNA replication *in vitro*. I. DNA structure-specific recognition of a primer–template junction by eukaryotic DNA polymerases and their accessory proteins. *J. Biol. Chem.*, **266**, 1950–1990.
- Ustav, M. and Stenlund, A. (1991) Transient replication of BPV-1 requires two viral polypeptides encoded by the E1 and E2 open reading frames. *EMBO J.*, **10**, 449–457.
- Ustav, M., Ustav, E., Szymanski, P. and Stenlund, A. (1991) Identification of the origin of replication of bovine papillomavirus and characterization of the viral origin recognition factor E1. *EMBO J.*, **10**, 4321–4329.
- Ustav, E., Ustav, M., Szymanski, P. and Stenlund, A. (1993) The bovine papillomavirus origin of replication requires a binding site for the E2 transcriptional activator. *Proc. Natl Acad. Sci. USA*, **90**, 898–902.
- Wahle, E., Laskens, R.S. and Kornberg, A. (1989) The DnaB–DnaC replication protein complex of *Escherichia coli*. II. Role of the complex in mobilizing DnaB functions. *J. Biol. Chem.*, **264**, 2469–2475.
- Wissman, A. and Hillen, W. (1991) DNA contacts probed by modification protection and interference studies. *Methods Enzymol.*, **208**, 365–379.
- Yang, L., Li, R., Mohr, I.J., Clark, R. and Botchan, M.R. (1991) Activation of BPV-1 replication *in vitro* by the transcription factor E2. *Nature*, **353**, 628–632.
- Yang, L., Mohr, I., Fouts, E., Lim, D.A., Nohaile, M. and Botchan, M. (1993) The E1 protein of bovine papilloma virus 1 is an ATP-dependent DNA helicase. *Proc. Natl Acad. Sci. USA*, **90**, 5086–5090.

Received June 19, 1998; revised September 1, 1998;
accepted September 10, 1998