A conserved regulatory module at the C-terminus of the papillomavirus E1 helicase domain controls E1 helicase assembly

Stephen Schuck and Arne Stenlund*

Cold Spring Harbor Laboratory
PO Box 100
Cold Spring Harbor NY, 11724

Running Title: Control of helicase assembly

* Corresponding author

Arne Stenlund
PO Box 100
1 Bungtown Road
Cold Spring Harbor Laboratory
Cold Spring Harbor NY 11724
Phone (516) 367-8407
FAX (516) 367-8454
e-mail stenlund@cshl.edu
Abstract

Viruses frequently combine multiple activities into one polypeptide to conserve coding capacity. This strategy creates regulatory challenges to ascertain that the combined activities are compatible and do not interfere with each other. The papillomavirus E1 protein, as many other helicases, has the intrinsic ability to form hexamers and double hexamers (DH) that serve as the replicative DNA helicase. However, E1 also has the more unusual ability to generate local melting by forming a double trimer (DT) complex that can untwist the double stranded ori DNA in preparation for DH formation. Here we describe a switching mechanism that allows the papillomavirus E1 protein to form these two different kinds of oligomers and to transition between them. We show that a conserved regulatory module attached to the E1 helicase domain blocks hexamer and DH formation and promotes DT formation. In the presence of the appropriate trigger, the inhibitory effect of the regulatory module is relieved and the transition to DH formation can occur.

Importance.

This study provides a mechanistic understanding into how a multi-functional viral polypeptide can provide different, seemingly incompatible activities. A conserved regulatory sequence module attached to the AAA+ helicase domain in the papillomavirus E1 protein allows the formation of different oligomers with different biochemical activities.
Introduction.

DNA helicases are important enzymes that catalyze the separation of the two complementary DNA strands in preparation for processes that require exposure of single stranded DNA, such as DNA replication, DNA repair and recombination (for reviews see (1, 2)). Hexameric DNA helicases are ring-shaped enzymes that function by encircling one of the DNA strands while displacing the other in an ATP dependent manner. These enzymes are highly processive and the key to the processivity is believed to be the encirclement of the DNA strand by the hexameric ring, which prevents disassociation of the enzyme from the substrate. Consistent with such a mode of action, a common feature of hexameric helicases is that the presence of ssDNA induces formation of hexamers or stabilizes existing hexamers (1).

The viral E1 protein from the papilloma virus family represents a subgroup of hexameric helicases. These DNA helicases, which belong to the SF3 helicase family (3), are not dedicated helicases but serve many different functions in the viral life cycle (for a review, see (4). The E1 proteins are termed initiator proteins because their primary function is to assist in the initiation of replication of the viral genome. The E1 protein takes part in initiation of DNA replication in at least four different ways. E1 binds cooperatively with the viral E2 transcription factor to the origin of DNA replication (ori) which contains binding sites for both proteins, forming a highly sequence specific E1-E2 complex (5-15). In this complex E1 uses its specific DNA binding domain (DBD) to bind to two of the four E1 binding sites present in ori as a head-to-head dimer(7, 8, 16, 17). The interaction between the E1 helicase domain and the E2 transactivation domain has two distinct
effects on DNA binding (10, 18-23). The cooperativity in the interaction contributes to the affinity of binding and the interaction also blocks a nonspecific DNA binding activity present in the E1 helicase domain, contributing to the specificity of DNA binding (12).

In the next step, through processes that are not well understood, E2 is displaced and the E1E2 complex is converted into an E1 double trimer (DT) complex on the ori (24-28). In the DT, E1 is bound to the ori as two head-to-head trimers, encircling dsDNA(24, 29). Formation of this complex depends on binding of the E1 DBD to the four E1 BS in the ori. However, a non-specific DNA binding activity present in the helicase domain, which makes non-sequence specific contacts with the DNA flanking the E1 BS is also critical for formation of the DT (24, 25, 27, 30). This complex is capable of local melting of the ori using an ATP hydrolysis dependent untwisting process (25).

The locally melted template, with the two head-to-head trimers constituting the DT now becomes the substrate for additional E1 molecules forming the E1 DH, which is composed of two head-to-head hexamers (31). Each hexamer encircles ssDNA and functions as a helicase with the ability to unwind the dsDNA bi-directionally in front of the replication forks (32-36). At this point the enzymes involved in the replication process such as topoisomerase I, DNA polymerase α, and RPA are recruited to the ori through interaction with E1 (21, 23, 37, 38). Although formation of the DH requires a DT precursor, hexamers of E1 can readily form directly if ssDNA is provided (31, 36). Both full length E1 and shorter C-terminal fragments containing the E1 helicase domain are able to form such ssDNA-induced hexamers utilizing the intrinsic ability of the
oligomerization domain and helicase domain to bind ssDNA and form hexamers. (36, 39, 40).

The need to utilize one protein to provide multiple functions presents a challenge not only in that multiple activities have to be incorporated into one polypeptide, but also because mechanisms that prevent the different activities from interfering with each other have to be built in. As described above some of the switches in activity that E1 undergoes are related to changes in oligomerization status. Because E1 has the intrinsic ability to form hexamers and double hexamers an obvious requirement is to prevent hexamer formation when complexes other than hexamers and DH (such as the E1E2 and the DT) are required. Conversely, a block to hexamer and DH formation has to be conditional and regulated such that the hexamer and DH can form when so required. Consequently, the transition from E1E2 to DT to DH must involve regulatory steps that control the ability of E1 to oligomerize.

Structural analyses provide an important part how we understand E1 oligomeriation. Multiple crystal structures containing the helicase and oligomerization domains of E1 have been solved providing excellent information about the overall structure of these helicases, how they form hexamers, how they bind nucleotide and how they associate with ssDNA (34, 41, 42). In two of these structures a 28 residue C-terminal peptide was deemed to be protease sensitive and therefore removed from the helicase domain to facilitate crystallization (34, 41). In the third structure, even though the peptide was
present, it was not ordered in the structure and therefore not visible, indicating that this peptide is not an integral part of the helicase domain (42).

Here we show that the short C-terminal peptide has critical functions in the biochemical activities of the E1 protein. The 28-residue peptide contains two separate elements, an acidic region and a C-terminal tail, which together regulate E1 oligomerization. The acidic region is a negative regulator of oligomerization and promotes DT formation. The C-tail is required for DH formation by counteracting the effect of the acidic region. Using these two elements E1 can undertake the orderly transition between the DT complex which melts the ori, and the DH which unwinds the viral DNA in front of the replication fork, providing a glimpse into how multifunctional initiator proteins transition from one activity to another.

Materials and Methods

Expression and purification of E1 and E2. Full length E1 and E1 fragments from BPV-1 were expressed in E. coli as N-terminal GST fusions and purified by glutathione agarose affinity chromatography. The GST portion was removed by digestion with thrombin and the material further purified by ion exchange chromatography. E2 was expressed without a tag and purified by mono S and mono Q ion exchange chromatography (43).
EMSA. Four percent acrylamide gels (39:1 acrylamide:bis) containing 0.5 × TBE, lacking EDTA, were used for all EMSA experiments. E1 was added to $^{32}$P labeled probe (~2 fmol) in 10 μl binding buffer, BB (20 mM HEPES [pH 7.5], 70 mM NaCl, 0.7 mg/ml BSA, 0.1% NP40, 5% glycerol, 5 mM DTT, 5 mM MgCl$_2$), and 2 mM ATP or ADP. After incubation at RT for 1 hr, the samples were loaded and run for 2 hr at 9V/cm (24).

Probes. The 84 bp ori probe for EMSA was generated by PCR using one primer labeled with $^{32}$P at the 5' end using polynucleotide kinase and γ-$^{32}$P ATP with a cloned fragment containing the BPV-1 ori sequence as a template. The probe contains the sequences between nt 7906 and 48 in the BPV-1 genome. To generate the wt 84 bp probe used in Figure 6, four staggered oligonucleotides were generated which after phosphorylation and ligation generated the 84 bp probe with perfect complementarity. To generate the bubble probe in Figure 6, one of the four oligonucleotides was replaced with an oligonucleotide that after annealing generated a 6 bp single stranded region between nt 7928 and 7933 in the BPV-1 genome.

Glycerol gradient sedimentation. Glycerol gradients (5 ml) were generated in 20 mM HEPES [pH 7.5], 70 mM NaCl, and 5 mM DTT. In Figures 1D and E, 1μg of E1 protein was incubated with 200 ng of a 41 nt oligonucleotide labeled with $^{32}$P and then loaded onto a 15-30% glycerol gradient which was run for 16 hr at 49,000 rpm in an SW 55 rotor. In Figures 1F and 2D, 100 μg of E1$_{308-577}$ and E1$_{308-597}$ respectively, was loaded onto a 5-30% glycerol gradient, which was run for 23 hr at 49,000 rpm in an SW55 rotor. The gradients were fractionated into 150 μl fractions and the peaks were localized by
Bradford assays (for marker proteins and for E1\textsubscript{308-577} and E1\textsubscript{308-597}) or by counting in a scintillation counter to detect $^{32}$P in DNA containing samples. Marker proteins were purchased from Sigma-Aldrich and run in parallel gradients.

**DNA helicase assays.** Short helicase substrates were generated by annealing an end-labeled 41-mer oligonucleotide to a 59-mer oligonucleotide with partial complementarity, generating a molecule with 26 bp of ds DNA, 15 nt of 5’ ssDNA, and 33 nt of 3’ ssDNA. After gel purification the substrate (~2 fmol) was incubated with E1 in BB with 2 mM ATP, but lacking NaCl for 15 min at 37°C. After the incubation, SDS was added to 0.1% and the sample was loaded onto 10% acrylamide gels (29:1 acrylamide:bis).

The long helicase substrates were generated by annealing a primer to M13 mp18 and extending with Klenow DNA polymerase in the presence of 50 $\mu$M dATP, dGTP and dTTP, 10$\mu$M dCTP, 80$\mu$M ddATP and 40$\mu$Ci $\alpha$-$^{32}$P dCTP. The average length of the extension products is determined by the relative levels of ddATP. The helicase assays were carried out as described above except that 12 $\mu$g/ml of E. coli SSB was included in the reaction. The samples were treated with proteinase K (200$\mu$g/ml), and analyzed by agarose gel electrophoresis.

**Unwinding assays.** Unwinding assays were performed by incubating 2 fmol of probe with E1 at 32°C for 30 min under EMSA conditions (see above), but in the presence of 10 $\mu$g/ml E. coli SSB. Before loading samples on the EMSA gel, Sarkosyl was added to 0.1% to disrupt E1-DNA complexes, without affecting ssDNA-SSB complexes (see Figure 5D).
ATPase assays  ATPase assays were performed in a 20-μl reaction containing 30 mM HEPES (pH 7.5), 30 mM NaCl, 1 mM DTT, 7 mM MgCl₂, 100 μg of bovine serum albumin/ml, 100 μM ATP, and 40,000 cpm of [γ-³²P] ATP and E1. Reactions were incubated for 1 h at room temperature and stopped by the addition of EDTA to a final concentration of 10 mM. Then, 2-μl portions of the reactions were spotted onto a polyethyleneimine (PEI)-cellulose plate, and the plates were then developed in 1 M formic acid and 0.5 M LiCl₂ for 40 min. After drying, each plate was exposed to a Fuji imaging plate, and the level of free phosphate was determined by scanning the plate using a Fuji BAS imager.

Pull down assays. Pull down assays were performed as described. (44) E₁₆₀₈₋₅₇₇ and E₁₆₀₈₋₄₁₀ were sub cloned into the pET 11 C vector and in vitro transcribed and translated using a TNT coupled rabbit reticulocyte kit (Promega) according to the manufacturer’s instructions. Two hundred ng of GST fusion protein (GST, GST E₁₅₄₄₋₆₀₅, or GST E₁₅₄₄₋₆₀₅ ₃xR) was incubated with 1μl of in vitro translation mixture in 20 μl of a solution containing 20 mM HEPES pH 7.9, 50 mM KCl, 2.5 mM MgCl₂, 10 % glycerol, 1 mM DTT, 0.2% NP40, and 3 μl of fetal bovine serum. After 2hr incubation 5 ml of glutathione agarose beads were added and incubate for a further 30 min. The beads were washed repeatedly with Tris buffered saline (TBS) containing 0.1% NP40 and the samples were analyzed by SDS-PAGE.

In vitro DNA replication. In vitro DNA replication assays were performed essentially as described (11, 45). In vitro replication was performed in 25 μl reaction mixtures containing; 40 mM HEPES-KOH, pH 7.5, 8 mM MgCl₂, 0.5 mM DTT, 3 mM ATP, 0.2
mM each of GTP, UTP, CTP; 0.1 mM each of dATP, dGTP, dTTP, 10 μM of [32P]dCTP
(2 μCi; 3000 Ci/mmol), 40 mM creatine phosphate, 400 ng creatine kinase, 10 μl S100 extract, and 0.5 μl high salt nuclear extract from H293-cells. The concentration of template in the in vitro reactions was 2ng/μl. Reactions were incubated for 60 min at 37°C, stopped by addition of SDS to 1% and EDTA to 10 mM, and treated with proteinase K followed by phenol/chloroform extraction and precipitation with ethanol and ammonium acetate. The products were analyzed by electrophoresis on 1% agarose gels in TAE buffer.

Results

A 28 amino acid C-terminal peptide in the papillomavirus E1 protein controls the formation of hexamers on ssDNA.

An alignment of the 30-35 C-terminal amino acids from multiple E1 proteins shows the presence of a short ~10 aa stretch of largely acidic residues (acidic region, Figure 1A) followed by a 13-18 residue long sequence which we have termed the C-tail (Figure 1A). The presence of multiple aspartic and glutamic acid residues is well conserved in the acidic region although the exact sequence is not. The C-tail contains three highly conserved residues, G588, F594 and C596, but otherwise shows limited conservation. The acidic region and the C-tail clearly are not integral parts of the E1 helicase domain structure. These peptides were removed prior to crystallization in two structures of the E1 helicase domain(34, 41); in the third structure these peptides were present in the crystallized fragment but were not visible in the structure demonstrating that they were not sufficiently ordered to be resolved (42).
To address the function of the acidic region and the C-tail we expressed and purified E1\textsubscript{308-605}, which includes these two peptides and E1\textsubscript{308-577}, which lacks these peptides, from E. coli and tested the biochemical properties of these proteins. E1\textsubscript{308-605} formed a robust complex on ssDNA (Figure 1B, lanes 1-4). In contrast, E1\textsubscript{308-577} formed a complex that did not enter the gel indicating a very high molecular weight (Figure 1B, lanes 5-8).

To examine the nature of the complexes formed on ssDNA we analyzed them by glycerol gradient sedimentation (Figure 1C). E1\textsubscript{308-605} in the presence of radiolabeled ssDNA and ATP formed a discrete complex sedimenting slightly faster than β-amylase indicating an apparent molecular weight of approximately 200 kD, consistent with a hexamer. In contrast, under that same conditions E1\textsubscript{308-577} formed a complex that pelleted in the gradient, indicating a much greater molecular weight (Figure 1D).

To determine whether this large complex was formed because the E1\textsubscript{308-577} protein was aggregated we sedimented E1\textsubscript{308-577} in a glycerol gradient in the absence of ssDNA. E1\textsubscript{308-577} formed a discrete peak with an apparent molecular weight of ~29 kD indicating that it is monomeric and not aggregated (Figure 1E). This indicates that it is the presence of ssDNA that induces the formation of the high molecular weight complex observed for E1\textsubscript{308-577}. We next tested the E1\textsubscript{308-605} and E1\textsubscript{308-577} for DNA helicase activity using a short partially double stranded oligonucleotide. The two proteins had similar levels of helicase activity (Figure 1F, compare lanes 3-6 to 7-10), which was surprising since the E1\textsubscript{308-577} failed to form a hexamer with ssDNA.
To address the function of the C-terminal 28 residues we generated small deletions in this region (Figure 2A). We generated the proteins E1\textsubscript{308-600}, E1\textsubscript{308-597}, E1\textsubscript{308-592}, and E1\textsubscript{308-589}, which affect the C-terminal tail and the deletion E1\textsubscript{308-584}, which also affect the acidic region. In addition we generated alanine substitutions in the conserved residues G588, F594, and C596 in the C-terminal tail. We first tested these proteins for the ability to form the E1 hexamer on ssDNA (Figure 2B and C). As observed in Figure 1B, E1\textsubscript{308-577} formed a complex that did not enter the gel (Figure 2B, lanes 2-4), while E1\textsubscript{308-605} formed a robust hexamer (Figure 2B, lanes 20-22). E1\textsubscript{308-600} (Figure 2B, lanes 17-19) behaved like E1\textsubscript{308-605}, but the rest of the deletions, E1\textsubscript{308-584}, E1\textsubscript{308-589}, E1\textsubscript{308-592} and E1\textsubscript{308-597}, all showed severe defects for hexamer formation (Figure 2B, lanes 5-7, 8-10, 11-13, and 14-16, respectively). Similarly, the point mutations F594A and C596A in the C-terminal tail failed to form the E1 hexamer (Figure 2C, lanes 8-10 and 11-13) while G588A had no obvious effect on E1 hexamer formation (Figure 2C, lanes 5-7). These results demonstrate that the majority of the mutations in the C-terminal module affect E1 hexamer formation.

We were surprised that none of the deletions or point mutations, although defective for hexamer formation, exhibited the same phenotype as E1\textsubscript{308-577}, which forms large complexes in the well. After examining the structure of E1\textsubscript{308-577} we considered the possibility that the deletion to residue 577 could cause structural issues in the folding of the helicase domain, possibly because the presence of residues immediately C-terminal to residue 577 are important. To address this possibility we added either seven alanines or seven asparagines to the endpoint at residue 577 generating E1\textsubscript{308-577 \textsubscript{7xN}} and E1\textsubscript{308-577 \textsubscript{7xA}},

\[\text{Figure 2A}\]
respectively and tested these for hexamer formation in the presence of ssDNA. As shown in Figure 2C, lanes 14-19, these additions restored fully the ability to form the E1 hexamer. This result was surprising since it demonstrates that the acidic region and the C-tail are not per se required for hexamer formation by the E1 helicase domain, in contrast to expectation. We were also intrigued by the fact that E1\textsubscript{308-577} 7xN and E1\textsubscript{308-577} 7xA (Figure 2C, lanes 14-19) formed discrete hexamers at wt levels while E1\textsubscript{308-584} had a severe defect for hexamer formation (Figure 2B, lanes 5-7) although these proteins all have seven residues added to the C-terminus of E1\textsubscript{308-577}. This difference demonstrates that the 7 residues of the acidic region present in E1\textsubscript{308-584} have a direct negative effect on E1 hexamer formation and that the function of the acidic region therefore appears to be to inhibit hexamer formation. In the presence of the C-tail this block to hexamer formation on ssDNA can be bypassed.

To ascertain that deletions in the C-terminal module did not cause gross structural defects we performed sedimentation analysis with one of these mutants. As shown in Figure 2D, E1\textsubscript{308-592} sedimented as a discrete peak similar to the marker protein carbonic anhydrase (Mw~ 29kD) demonstrating that it remains a discrete monomer.

As we showed in Figure 1E, both E1\textsubscript{308-605} and E1\textsubscript{308-577} had helicase activity when tested using short partially double stranded oligonucleotides. For a more stringent helicase assay, more relevant to the role of E1 as a replicative DNA helicase, we generated substrates with much longer double stranded regions by extending a primer annealed to a single stranded M13 template in the presence of \textsuperscript{32}P labeled nucleotide and low levels of
dideoxynucleotide to generate a substrate with a double stranded region ranging between 1-5 kb in length. Such a substrate can provide an estimate of the processivity of the helicase (Figure 3 A-D). E1$_{308-577}$ failed to unwind this substrate, as did the deletions E1$_{308-584}$, E1$_{308-589}$, and E1$_{308-592}$ (Figure 3A, lanes 3-15). E1$_{308-597}$ showed very low level of unwinding (Figure 3B, lane 5), while E1$_{308-600}$ and E1$_{308-605}$ showed robust helicase activity (Figure 3B, lanes 7-12). Addition of 7xN or 7xA restored helicase activity to E1$_{308-577}$ (Figure 3C, lanes 7-12). The point mutations in the conserved residues in the C-terminal tail (F594 and C596, Figure 3D, lanes 10-15) showed a very low level of activity for DNA helicase activity while mutation of the conserved G588 showed no defect (Figure 3D, lanes 7-9). These results show that all mutants that could form the hexamer well were also able to unwind the long helicase substrate, while mutations defective for hexamer formation were unable to do so.

ATP hydrolysis by the E1 protein and other hexameric helicases is dependent on the ATP binding and hydrolysis site formed by two adjacent subunits in the hexamer (34, 42, 46). Because hexamer formation is greatly stimulated by the presence of ssDNA, consequently ATPase is greatly stimulated by ssDNA. We tested the ability of our C-terminal mutants to hydrolyze $\gamma$-$^{32}$P ATP (Figure 3E). In the absence of ssDNA E1$_{308-605}$ had a low level of ATPase activity (Figure 3E, lane 1), which was greatly stimulated (6-fold) in the presence of ssDNA (Figure 3E, lane 2). The remaining samples were tested in the presence of ssDNA. E1$_{308-577}$ showed a low level of ATPase activity, (Figure 3E, lane 4) while the deletion mutants E1$_{308-584}$, E1$_{308-589}$, and E1$_{308-592}$ showed only a trace of ATPase activity (Figure 3E, lanes 5-7). E1$_{308-597}$ had a significantly higher activity and
activity was restored completely for E1_{308-600} (Figure 3E, lanes 8-9). As observed with the helicase activity, E1_{308-577} 7xA and 7xN both showed a fully restored ATPase activity (Figure 3E, lanes 11-12). Two of the three point mutations in the C-terminal tail, G588A (Figure 3E, lane 13) and C596A (Figure 3E, lane 15) showed a modest reduction of activity while F594A (lane 14) lacked activity altogether. These results are consistent with the ability of the different mutants to form hexamers on ssDNA.

The results shown in Figure 2 and 3 clearly demonstrate that for the activities associated with the C-terminal half of E1 (hexamer formation on ssDNA, DNA helicase activity, and DNA dependent ATPase activity) the acidic region and the C-tail are largely dispensable. However, when present, the acidic region clearly has a negative effect on all of these activities, likely because the acidic region inhibits E1 hexamer formation. In the presence of the C-tail, the acidic region is no longer inhibitory for hexamer formation.

The acidic region and the C-tail affect DT and DH formation by full length E1.

Because the acidic region and the C-tail appear to be dispensable for the activities found in the E1 oligomerization and helicase domains, we generated the same deletions and mutations that we had tested in E1_{308-605} in the context of full length E1 (Figure 4). In contrast to E1_{308-605}, full length E1 forms several discrete types of complexes with dsDNA. E1 can form an E1_{3}E2 complex together with the viral transcription factor E2. This complex serves to recognize the ori site specifically (5-12, 15, 47, 48). E1 also forms a DT complex on the ori; this complex generates local melting of the ori and serves as a precursor for the E1 DH, which is the viral replicative DNA helicase (24, 25, 27, 32,
We first tested the C-terminal deletions mutants for DT formation (Figure 4B and C). All the deletion mutants, with the exception of E1<sub>1-577</sub> formed the DT well (Figure 4B, lanes 5-16). Addition of 7xN or 7xA to E1<sub>1-577</sub> did not restore the ability to form the DT (Figure 4C, lanes 5-10), and the five alanine substitutions in the C-tail had no effect on DT formation (Figure 4C, lanes 11-25). These results establish two important facts. Firstly, there is a specific requirement for the acidic region for DT formation, since E1<sub>1-584</sub> forms the DT (Figure 4B, lanes 14-16) while E1<sub>1-577 7xN</sub> does not (Figure 4C, lanes 5-7). Secondly, the C-tail is clearly not required for DT formation.

We next examined the mutant proteins for their ability to form the E1 DH by EMSA (Figure 4D and E). When E1 is incubated with probe in the presence of ATP the DT is slowly converted into a DH, in a process that requires local melting of the dsDNA (24, 25, 27). Interestingly, only one of the deletions, E1<sub>1-600</sub> was still able to form the DH, demonstrating that the C-tail is critically important for DH formation (Figure 4D, lanes 20-22). Results consistent with this finding were observed with point mutations in the C-tail; G588A, M591A, and T593A formed the DH (Figure 4E, lanes 11-19). However, mutation of the conserved residues F594A (Figure 4E, lanes 20-22) and C596A (Figure 4E, lanes 23-25) resulted in failure to form the DH. Together these results show that the C-tail plays a specific role in the formation of the DH, but plays no role in DT formation, while the acidic region is specifically required for DT formation.

To ascertain that the acidic region plays a role for DT formation also in the context of the intact E1<sub>1-605</sub> we introduced arginine substitutions into the acidic region (Figure 4F).
Because the acidic region contains many acidic residues and the alignment with other E1 proteins indicate that considerable sequence variation is allowed as long as the acidic nature is maintained, we generated radical substitutions in multiple positions. We generated the mutants D586R, D579R/D583R, D579R/D583R/E586R, and D579R/D583R/E586R/D587R in the context of E1\textsubscript{1-605}. After expression and purification we tested these mutants for DT formation by EMSA. As expected, the single mutation (E586R) had little effect on DT formation (Figure 4F, lanes 8-11). The double mutant (D579R/D583R) could still form the DT but now also formed a larger complex (Figure 4F, lanes 5-7), something that we never observed with wt E1. The triple and quadruple mutations showed severe defects for DT formation generating complexes similar to those generated by E1\textsubscript{1-577} (Figure 4F, compare lanes 12-19 and 20-22). These results demonstrate that the acidic region in the context of full length E1 is necessary for formation of the DT.

To verify that the defects for DT formation were not caused by structural effects of the arginine substitutions we tested the mutants for formation of a different complex that does not rely on the ability of E1 to oligomerize (Figure 4G). E1 can together with the viral E2 protein form an E1\textsubscript{2}E2\textsubscript{2} complex on the ori. Formation of this complex depends on site-specific DNA binding by both proteins as well as interactions between the E1 and E2 DBDs and between the E1 helicase domain and the activation domain of E2 (6, 9, 10, 41, 49). Clearly, all the arginine substitutions were capable of forming the E1\textsubscript{2}E2\textsubscript{2} complex (Figure 4G, lanes 4-11) as was E1\textsubscript{1-577}, ruling out serious structural defects for these proteins.
The acidic region and the C-tail are required for unwinding and in vitro DNA replication.

The EMSA performed in Figure 4 showed that mutation of the acidic region results in a defect for DT and DH formation. To provide functional analysis of mutations affecting the acidic region and the C-tail we utilized an unwinding assay. The E1 DT formed on the ori can locally melt the DNA allowing formation of the E1 DH, which then unwinds the ori DNA and generates ssDNA (24, 25). We can detect the ssDNA generated in this process by the addition of E. coli SSB, which forms a complex with ssDNA that can be detected by EMSA. As expected, the deletions E1\textsubscript{1-584}, E1\textsubscript{1-589}, E1\textsubscript{1-592} (Figure 5A, lanes 3-11), which fail to form the DH lacked any detectable unwinding activity. E1\textsubscript{1-597} showed a trace of unwinding (lanes 12-14) while E1\textsubscript{1-600} and full length E1 (E1\textsubscript{1-605}) showed substantial activity (lanes 15-20). We next tested point mutations in the C-terminal tail and the acidic region for unwinding (Figure 5B and C). Similar to the results for DH formation, G588A and M591A showed wt levels of unwinding (Figure 5B, compare lanes 3-5 to lanes 6-11). However, the point mutations T593A, F594A, and C596A showed only a faint trace of activity (Figure 5B, lanes 12-20). E1\textsubscript{1-577}7xN (Figure 5B, lanes 21-23) lacked activity altogether. The arginine substitutions in the acidic region had more modest effects on unwinding. D586R showed a 2-4-fold reduction in unwinding (Figure 5C, lanes 7-10) while D579R/D583R/D586R showed a 4-8-fold reduction for unwinding (Figure 5C, compare lanes 3-6 to lanes 15-18). In Figure 5D we demonstrated that the sarkosyl treatment used to disrupt E1 complexes in panels A-C, does not affect SSB-ssDNA complexes (compare lanes 2 and 3, 5 and 6 and 7 and 8).
As a final step in measuring the activity of the mutations in the C-terminus of E1 we measured cell free DNA replication in human cell extracts. As shown in Figure 5E, E1$_{600}$ had close to wt activity (compare lanes 2 and 3) while all the remaining deletions, E1$_{597}$, E1$_{592}$, E1$_{589}$, E1$_{584}$, and E1$_{577}$ (Figure 5E, lanes 4-8) as well as the two mutants F594A and C596A (Figure 5E, lanes 12-13) lacked activity altogether. Of the remaining point mutations, T593A had a reduced activity (Figure 5E, lane 11) while the point mutations (G588A and M591A, Figure 5E, lanes 9-10) had close to wt activity.

In Figure 5F, we tested the activities of the arginine substitutions in the acidic region for in vitro replication. The mutant E586R had partial activity (Figure 5F, lanes 8-10) while D579R/D583R (Figure 5F, lanes 5-7) and D579R/D583R/E586R (Figure 5F, lanes 11-13) lacked detectable replication activity. These results are generally consistent with the defects observed for complex formation and unwinding and show that the acidic region and the C-tail are essential for viral DNA replication.

Our ability to knock out the acidic region with multiple point mutations indicated to us that it might be possible to test whether the C-tail becomes dispensable when the acidic region is inactive. Such a result would strongly support the notion that the function of the C-tail is to counteract the activity of the acidic region. We therefore generated the constructs E1$_{308-605}$ 579R 583R 586R, which knocks out the acidic region, and E1$_{308-605}$ 579R 583R 586R C596A, which knocks out both the acidic region and the C-tail. Surprisingly, we failed to purify E1$_{308-605}$ 579R 583R 586R due to poor solubility, however E1$_{308-605}$ 579R 583R 586R C596A...
could readily be purified and we tested this protein for the formation of hexamer on ssDNA as well as for ATPase activity (Figure 5G). Strikingly, knocking out the acidic region in the context of E1_{308-605} C596A, which has a severe defect for hexamer formation (Figure 5G, lanes 5-7), restores hexamer formation to wt levels Figure 5G, compare lanes 2-4 and 8-10). Similarly, E1_{308-605} C596A, which has a significant defect for ATPase activity can be restored to wt activity by knocking out the acidic region (Figure 5H, compare lanes 3, 4 and 5). Although these results are not completely conclusive since we were unable to measure that activity of the acidic region mutant alone, they clearly suggest that the defects of the mutants in the C-tail are caused by the exposure of the acidic region.

The acidic region interacts with the oligomerization domain.

The results in Figure 2 and 3 demonstrate that the acidic region is a negative regulator of E1 hexamer formation. An obvious possibility is that the acidic peptide can interact with the oligomerization domain and prevent the interaction between the oligomerization domains that is required for hexamer formation. To determine whether we could detect such an interaction we performed pull down assays using a GST fusion containing a C-terminal fragment from E1 including the acidic region and the C-tail (GST{E1}_{544-605}). This protein interacted well with the in vitro translated oligomerization and helicase domain fragment E1_{308-577} (Figure 6A, lane 3). The interaction was entirely dependent on the acidic region as illustrated by the fact that the triple arginine substitution (D579R/D583R/E586R) in the acidic region lost the ability to bring down the fragment (GST 544-605 3xR, Figure 6A, lane 2). We observed a similar result using an in vitro
translated oligomerization domain fragment (E1\textsubscript{308-410}, Figure 6C, compare lanes 2-3) demonstrating that the oligomerization domain by itself is sufficient for the interaction with the acidic region.

**ssDNA is a trigger for DH formation.**

E1\textsubscript{308-605} forms hexamers on ssDNA in the presence of either ADP or ATP, demonstrating that nucleotide hydrolysis is not necessary for hexamer formation (36, 42). DH formation with E1\textsubscript{1-605} however, requires ATP likely because the generation of ssDNA, which is required for DH formation, is dependent on ATP hydrolysis by the DT (24, 25). We reasoned that it might be possible to bypass the ATP dependence if we could bypass the local melting step. We therefore generated an ori probe containing a 6 bp single stranded region at the position in the ori (nt 7928-7933) where the E1 helicase domain interacts with and melts the dsDNA (25-27). We then tested this “bubble probe” for complex formation in parallel with a completely double stranded probe. On the completely double stranded probe wt E1 forms a DT in the presence of ADP, but formation of DH requires ATP (Figure 6E, lanes 2-7). Strikingly, using the bubble probe, we could generate a DH complex also in the presence of ADP (Figure 6E, lanes 9-11). This result indicates that in the presence of ssDNA ATP hydrolysis is no longer necessary for DH formation, consistent with an ability of E1\textsubscript{308-605} to form hexamers on ssDNA without ATP hydrolysis.

Our data indicates that the acidic region is required for DT formation because it inhibits hexamer formation. This inhibition clearly is conditional since prolonged incubation of
the DT in the presence of ATP results in the transition to a DH. The timing of the first appearance of DH (~10-12 min) correlates with the time required to generate local melting, indicating that the trigger for DH formation likely is ssDNA (25). Consistent with this idea, the DH forms instantly (< 1 min) on the bubble probe where melting is not required (Figure 6E, lanes 9-11).

Our data shows that the C-tail in the presence of ssDNA can counteract the inhibitory effect that the acidic region exerts on oligomerization and a logical conclusion is therefore that it is the recognition of ssDNA by the C-tail that triggers DH formation. We therefore tested whether the C-tail mutants C594A and F596A could form the DH on the bubble probe in the presence of ADP (Figure 6E, lanes 12-17). Interestingly, both of these mutants failed to form the DH demonstrating that without the C-tail, ssDNA no longer triggers DH formation. Although this result does not prove that the C-tail recognizes ssDNA, it is clearly consistent with such a possibility. These data together provide strong evidence that ssDNA is the trigger for DH formation, likely through recognition by the C-tail.

Discussion

The E1 initiator protein combines multiple activities that are required for initiation of papillomavirus DNA replication. These activities are deployed sequentially and result in recognition of the ori by an E1E2 complex, local melting of the ds viral DNA by an E1 DT complex, formation of an E1 DH helicase that encircles ssDNA, and eventually
unwinding of the dsDNA. We have identified a short peptide in the C-terminus of E1 that functions as a regulatory module in this process. The peptide contains a negative regulator of oligomerization, the acidic region, which promotes DT formation. Another part of the peptide, the C-tail, can counteract the negative regulator and therefore promotes E1 oligomerization and DH formation.

We were initially interested in the function of the E1 C-terminal peptide because it is fairly well conserved, but not visible in the crystal structures indicating that it is not an integral part of the helicase domain. This prediction appears to be correct, because the acidic region and the C-tail can be functionally replaced by 7xN or 7xA the acidic region and the C-tail clearly represent add-ons that do not affect the structure of the helicase domain.

Our mutational analysis of the C-terminal peptide in the context of the oligomerization and helicase domain (E1308-605) resulted in several surprising observations. The intact E1308-605 fragment readily formed the expected hexamer with ssDNA, and had robust helicase and ATPase activities (Figures 2 and 3). The shorter E1308-577 fragment, which corresponds to the fragment used for structure determination of the E1 hexamer in complex with ssDNA (34), failed to form a discrete complex and had poor helicase activity on long substrates as well as reduced ATPase activity. Strikingly, although this defect is due to the absence of the C-terminal module, the defect is not specific since addition of 7xA or 7xN at residue 577 restored wt activity to the fragment. We believe that the deletion to residue 577 may compromise folding in some way. This folding
defect must be fairly subtle however, since E1\textsubscript{308-577} remains monomeric and does not aggregate in the absence of DNA (Figure 1). Also, under the crystallization conditions this protein clearly forms hexamers (34).

Because the C-terminal module is not visible in the E1 hexamer structure Whelan et al. (47) performed small angle X-ray scattering to determine the position of the C-terminal module in the hexamer. The authors found that it occupied a cleft between adjacent subunits in the ring and based on electrostatic potential calculations the authors suggested that the C-terminal module may bridge adjacent subunits. Comparison of E1\textsubscript{399-605} and E1\textsubscript{292-579} showed that a 26 aa deletion that removes the C-terminal module resulted in a protein that was less capable of forming hexamers with ssDNA, had reduced ATPase activity, and especially on longer substrates had poor helicase activity compared to the intact protein. Based on these data the authors suggested that the C-terminal module functions a brace that holds the hexamer together during unwinding of ds DNA. In the hexameric T7 gene 4 helicase a 31 aa N-terminal tail crosses over from one subunit and interacts with the adjacent subunit (50, 51).

The data that we present here clearly contradicts a model where the C-terminal module functions as a brace. The restoration of ATPase, DNA binding and DNA helicase activities that can be accomplished by replacing the 28 residue C-terminal module with short oligopeptides (7xN or 7xA) clearly shows that neither the sequence nor the length of the peptide is critical for restoration of activity and it is hard to imagine a brace without sequence dependence. Our interpretation, as stated above, is that the deletion of
the whole C-terminal module results in folding problems that can be rescued by
increasing the length of the polypeptide slightly.

Overall, the deletion mutant (E1<sub>299-579</sub>) used by Whelan et al. showed similar defects that
we observe in E1<sub>308-577</sub> including a reduced ability to form hexamers, reduced
processivity of the helicase and reduced ATPase activity. The slight differences that do
exist, e.g. E1<sub>299-579</sub> forms a hexamer poorly, while E1<sub>308-577</sub> forms a complex in the well is
likely due to subtle differences in the constructs and/or expression and purification
procedures. In our hands a mutant with a slightly smaller deletion than E1<sub>308-577</sub> (E1<sub>308-581</sub>) produces a mixture of hexamers and complexes that stick in the well (Data not
shown).

We observe excellent correlation between hexamer formation by E1<sub>308-605</sub> on ssDNA and
DH formation on dsDNA by E1<sub>1-605</sub> indicating that these processes are closely related and
that E1 oligomerizes in the same way in the hexamer and the DH (Figure 7A). The
critical component for forming the hexamer is the oligomerization domain, which is
positioned immediately N-terminal to the helicase domain and forms a tight structure that
holds the hexamer together (34, 42). In the DH the oligomerization domain likely adopts
the same structure. We have less information about the structure of the DT complex, but
based on biochemical analysis the DT is critically dependent on template length while the
formation of the DH is not. This is consistent with formation of planar rings in the
hexamer and DH. In contrast, in the DT, E1 is likely to bind in a helical arrangement (24,
30, 34, 42). Because the oligomerization arrangement is unlikely to be shared between
the hexamer and the DT it is plausible that blocking of the oligomerization domain to prevent formation of hexamers and DHs would still allow the DT to form.

The function of the acidic region and C-tail in E1\textsubscript{308-605}.

We were surprised that small deletions from the C-terminus (E1\textsubscript{308-584}, E1\textsubscript{308-589}, E1\textsubscript{308-592}, and E1\textsubscript{308-597}) resulted in proteins that behaved differently than either E1\textsubscript{308-577}, E1\textsubscript{308-577}\textsuperscript{7xN} or E1\textsubscript{308-605}. The small deletions resulted in severe defects for hexamer formation, ATPase and helicase activity (Figures 2 and 3). However, the defects were not due to formation of larger complexes (as in E1\textsubscript{308-577}) but were caused by a failure to form any complex, indicative of a defect in ssDNA binding and/or oligomerization. Comparison of the two proteins E1\textsubscript{308-584} and E1\textsubscript{308-577}\textsuperscript{7xN}, which both have seven residues added after residue 577 shows that E1\textsubscript{308-577}\textsuperscript{7xN} in all respects behaved as E1\textsubscript{308-605}, while E1\textsubscript{308-584} was defective for all the activities that reside in the helicase/oligomerization domain fragment (which are all dependent on hexamer formation) (Figures 2 and 3). This result demonstrates first of all that the acidic region and the C-tail are not required for hexamer formation in the context of E1\textsubscript{308-605}. Secondly, when present alone the acidic region has a strong negative effect on hexamer formation. When both the acidic region and the C-terminal tail are present (as in E1\textsubscript{308-605}) this negative effect of the acidic region is neutralized indicating that the C-tail is a modulator of the acidic region and that the acidic region and the C-tail likely function together. This notion is supported by the results in Figure 5F and G, which indicate that mutation of the acidic region restores the activity to the C-tail mutant C596A, consistent with a role for the C-tail in blocking the activity of the acidic region.
The function of the acidic region and C-tail in E1\textsubscript{1-605}

Because the acidic region and the C-tail do not serve a specific function in the context of E1\textsubscript{308-605} we examined mutations in these two regions in full length E1\textsubscript{1-605}. In the presence of ds ori DNA, which is the substrate for E1\textsubscript{1-605}, we observed striking effects of removal of the acidic region and the C-tail. E1\textsubscript{1-577} \textsubscript{7xN} could still bind DNA, but no longer formed discrete complexes (Figure 4). Restoration of the acidic region e.g. E1\textsubscript{1-584}, E1\textsubscript{1-589}, E1\textsubscript{1-592} and E1\textsubscript{1-597} resulted in formation of the discrete DT complex on dsDNA. This result establishes that the acidic region has to be present to form the DT. Based on the strong inhibitory effect of the acidic region on hexamer formation in the context of E1\textsubscript{308-605} it is likely that the way the acidic region promotes DT formation is by inhibiting hexamer formation. We believe that the acidic region inhibits oligomerization by directly interacting with the oligomerization domain (Figure 6A and B). Because the monomer of the E1 helicase domain in solution may look quite different than the monomer extracted from the hexamer structure we cannot predict with any accuracy how such an interaction comes about.

Inactivation of the C-tail also has a very clear phenotype in E1\textsubscript{1-605}. Point mutations or deletions in the C-tail result in failure to form the DH complex, although DT formation is unaffected. In analogy with the function of the C-tail in E1\textsubscript{308-605} where the C-tail clearly counteracts the inhibitory effect of the acidic region and promotes hexamer formation, the C-tail in full length E1 likely promotes DH formation by counteracting the acidic
region. We can clearly rule out any direct effect of the C-tail in hexamer formation since the whole C-terminal module can be removed without affecting hexamer formation.

We can combine the data obtained from E1_{308-605} and from E1_{1-605} and describe a comprehensive model for the function of the C-terminal module in viral DNA replication (Figure 7B). Our data shows that two short peptides, the acidic region and the C-tail, together form a regulatory module, which is attached to the C-terminus of the helicase domain. This module is not required for hexamer formation on ssDNA. In the absence of this module E1_{308-577} can use the intrinsic ability of the oligomerization and helicase domains to bind ssDNA and to form hexamers, an ability that is common to many helicas (Figure 7B I). In full length E1 however, the C-terminal module plays an essential role. The ability of E1 to perform local melting of dsDNA is critically dependent on the DT, which forms only in the presence of the acidic region. In the absence of the acidic region ssDNA cannot be generated, and therefore the DH cannot form either.

If the C-tail is removed or mutated in E1_{308-605}, where the acidic region is present, we observe a strong inhibitory effect of the acidic region on hexamer formation with ssDNA (Figure 7B II). Because we can detect a physical interaction between the acidic region and the oligomerization domain and because the oligomerization domain is required for hexamer formation, we believe that the acidic region inhibits oligomerization directly. If the C-tail is removed or mutated in E1_{1-605}, where the acidic region is present, E1 is fully competent to form the DT complex. However, the DT cannot progress to DH since the C-
When both the acidic region and the C-tail are present in the context of E1\textsubscript{308-605}, the inhibitory effect of the acidic region is no longer observed and in the presence of ssDNA, the hexamer is formed (Figure 7B III). We believe that ssDNA triggers the C-tail to counteract the inhibitory effect that the acidic region has on oligomerization. Full length E1 first forms a DT, demonstrating that the acidic region, which promotes DT formation, is active. As local melting proceeds and ssDNA becomes available the C-tail recognizes the ssDNA and triggers DH formation.

In conclusion, our data clearly shows that the C-terminal module, which is present in all papillomavirus E1 proteins, is an essential part of the machinery required for viral DNA replication. This module is an add-on to the helicase domain that modifies the activity of the E1 protein. The net result is that the acidic region and the C-tail extend the functionality of E1 since in addition to forming the hexamer and DH helicases which are intrinsic to the helicase and oligomerization domains, the presence of the acidic region allows the formation of the DT complex, which can melt dsDNA locally. The transition to the DH then requires the C-tail. The critical importance of these two short peptides in the C-terminal module raises the possibility that these peptides could represent targets for antiviral compounds.
Acknowledgments. This work was supported by a grant RO1 AI072345 to A.S.

References

Figure legends

Figure 1. A C-terminal 28 aa peptide is important for complex formation by E1<sub>308</sub>. A. The C-terminal regions from 6 different papillomavirus E1 proteins and SV40 T-ag were aligned. Highly conserved residues are highlighted and the acidic region is bracketed.
B. A 41-mer ssDNA probe was incubated with three levels (8, 16, and 32 ng) of E1$_{308-605}$ (lanes 2-4) or E1$_{308-577}$ (lanes 6-8) in the presence of ADP and subjected to EMSA. In lanes 1 and 5, no protein was added. The position of the E1 hexamer complex is indicated.

C. Two fmol of helicase substrate was incubated with four quantities (8, 16, 32 and 64 ng) of full length E1$_{1-605}$ (lanes 3-6), E1$_{308-605}$ (lanes 7-10) and E1$_{308-577}$ (lanes 11-14) in the presence of ATP. Lane 1 contained probe alone, while lane 2 contained boiled probe. The migration of dsDNA and ssDNA are indicated.

D. Binding reactions for EMSA with E1$_{308-605}$, as shown in Figure 1B, were scaled up and sedimented in a 15-30% glycerol gradient. The radioactivity in the fractions was quantitated, and plotted. The peak fractions were loaded onto an EMSA gel to verify that the gradient peak corresponded to the complex that we observe by EMSA (see inset). Arrows indicate the sedimentation of the marker proteins bovine serum albumin (BSA), alcohol dehydrogenase (ADH), and β-amylase.

E. Binding reactions for EMSA with E1$_{308-577}$, as shown in Figure 1B, were scaled up and sedimented in a 15-30% glycerol gradient. The radioactivity in the fractions was quantitated, and plotted. The last fraction plotted corresponds to the bottom of the tube, which was cut off and counted.
One hundred μg of E1\textsubscript{308-577} was sedimented in a 5-30\% glycerol gradient. The protein was detected by Bradford assays and compared to the sedimentation of the marker proteins carbonic anhydrase and bovine serum albumin (BSA).

Figure 2. The C-terminal 28 residues regulate E1 hexamer formation

A. A cartoon showing the sequence of the C-terminal module with deletions and point mutations in the context of E\textsubscript{308-605}.

B. The deletion mutants in the C-terminal module were tested for hexamer formation in the context of E\textsubscript{308-605} using an ssDNA probe. Two, four and eight ng each of the proteins E\textsubscript{308-577}, E\textsubscript{308-584}, E\textsubscript{308-589}, E\textsubscript{308-592}, E\textsubscript{308-597}, E\textsubscript{308-600} and E\textsubscript{308-605}, were incubated with a 41-mer ssDNA probe (~2 fmol) in the presence of ADP and tested for hexamer formation by EMSA. The migration of the E1 hexamer is indicated.

C. Two, four and eight ng each of the point mutations G588A, F594A and C596A, and the deletion mutants E\textsubscript{308-577} 7xA, and E\textsubscript{308-577} 7xN were incubated with a 41-mer ssDNA probe in the presence of ADP and tested for hexamer formation by EMSA.

D. One hundred μg of E1\textsubscript{308-592} was sedimented in a 5-30\% glycerol gradient. The protein was detected by Bradford assays and compared to the sedimentation of the marker proteins carbonic anhydrase and bovine serum albumin (BSA).
Figure 3. Mutations in the C-terminal module affect the helicase and ATPase activity of E1. A helicase substrate was generated by annealing a primer to M13 ssDNA followed by extension with a mixture of dNTPs and ddNTPs and Klenow DNA polymerase to generate partially double stranded templates with long (0.2 - 5 kb) radio-labeled strands. In each set, two fmol of the substrate was incubated with 3 levels (4, 8, and 16 ng) of the proteins. In panel A, E1308-577, E1308-584, E1308-589 and E1308-592 were tested. In panel B, E1308-597, E1308-600, E1308-605 were tested. In panel C, E1308-577, E1308-577 7xA, E1308-577 7xN, E1308-605 were tested and in panel D, E1308-605, E1308-605 G588A, E1308-605 F594A, E1308-605 C596A were tested. In each panel a size marker (M), a boiled probe (B), and a no E1 (-) sample were included. E. Point mutants and deletions in the C-terminal module in the context of E1308-605 were tested for ATPase activity. Samples were incubated with γ-32P ATP in the presence of ssDNA except where otherwise indicated, followed by separation of free phosphate from ATP by thin layer chromatography. The level of hydrolysis is shown below each lane.

Figure 4. The acidic region is required for DT formation and the C-terminal tail is required for DH formation by E11-605.
A. A cartoon showing the deletions and point mutations in the C-terminal module in the context of E11-605.
B. The 84 bp ori probe was incubated with three levels (2.5, 5, and 10 ng) of the deletion mutants E11-597, E11-592, E11-589, E11-584, and E11-577 in the presence of ADP and tested for DT formation by EMSA. The mobility of the DT complex is indicated.
C. The 84 bp ori probe was incubated with three levels (2.5, 5, and 10 ng) of the mutants
E1<sub>1-577</sub> 7xN, E1<sub>1-577</sub> E1<sub>1-605</sub> G588A, E1<sub>1-605</sub> M591A, E1<sub>1-605</sub> T593A, E1<sub>1-605</sub> F594A, and
E1<sub>1-605</sub> C596A in the presence of ADP and tested for DT formation. The mobility of the
DT complex is indicated.

D. The 84 bp ori probe was incubated with three levels (2.5, 5, and 10 ng) of the mutants
E1<sub>1-577</sub> 7xN, E1<sub>1-577</sub>, E1<sub>1-584</sub>, E1<sub>1-589</sub>, E1<sub>1-592</sub>, E1<sub>1-597</sub>, and E1<sub>1-600</sub> in the presence of ATP
and tested for DH formation. The mobility of the DH complex is indicated.

E. The 84 bp ori probe was incubated with three levels (2.5, 5, and 10 ng) of the mutant
proteins E1<sub>1-577</sub> 7xA, E1<sub>1-577</sub>, E1<sub>1-605</sub> G588A, E1<sub>1-605</sub> M591A, E1<sub>1-605</sub> T593A, E1<sub>1-605</sub>
F594A, and E1<sub>1-605</sub> C596A and tested for DH formation in the presence of ATP by
EMSA.

F. The 84 bp ori probe was incubated with three levels (2.5, 5, and 10 ng) of wt E1 or the
mutants E1<sub>1-605</sub> 579R/583R, and E1<sub>1-577</sub> and tested for DT formation in the presence of
ADP by EMSA. For the mutants E1<sub>1-605</sub> 586R, E1<sub>1-605</sub> 579R/583R/586R, and E1<sub>1-605</sub>
579R/583R/586R/587R, four levels of E1 (1.2, 2.5, 5 and 10 ng) were used.

G. To determine whether the arginine substitutions in the acidic region caused structural
defects these mutants were tested for the ability to form the E1<sub>2</sub>E2<sub>2</sub> complex. 1.2 and 2.5
ng of wt E1 or the mutants E1<sub>1-605</sub> 586R, E1<sub>1-605</sub> 579R/583R, E1<sub>1-605</sub> 579R/583R/586R, or
E1-605 579R/583R/586R/587R were incubated with the 84 bp ori probe and 2 ng of E2 and tested for E1-E2 complex formation by EMSA. In lane 12, 1.2 ng of E1,577 was tested for E1-E2 complex formation.

Figure 5. The C-terminal module is required for unwinding and DNA replication in vitro. A. The 84 bp ori probe was incubated with three levels (2.5, 5, and 10 ng) of E1,584, E1,589, E1,592, E1,597, E1,600 and E1,605 in the presence of ATP and E. coli SSB and analyzed by EMSA. ssDNA + SSB complexes are indicated.

B. The 84 bp ori probe was incubated with three levels (2.5, 5, and 10 ng) of E1,605, E1,605 G588A, E1,605 M591A, E1,605 T593A, E1,605 F594A, E1,605 C596A, and E1,605 7xN in the presence of ATP and E. coli SSB and analyzed by EMSA.

C. The 84 bp ori probe was incubated with four levels (2.5, 5, 10 and 20 ng) of E1,605, E1,605 586R, E1,605 579R/583R, E1,605 579R/583R/586R and E1,577 in the presence of ATP and E. coli SSB and analyzed by EMSA.

D. Low levels of sarkosyl disrupt E1 complexes with ssDNA and dsDNA, but do not affect complexes between E. coli SSB and ssDNA. The double stranded 84 bp ori probe was incubated with E1 in the presence of ADP (lanes 2 and 3). Immediately before loading onto an EMSA gel one sample (lane 3) was treated with 0.1% sarkosyl. The ssDNA probe was incubated with E1 in the presence of ADP (lanes 5 and 6). Immediately before loading onto an EMSA gel one sample (lane 6) was treated with...
0.1% sarkosyl. The ssDNA probe was incubated E. coli SSB (lanes 7 and 8). Immediately prior to loading onto the EMSA gel one sample (lane 8) was treated with 0.1% sarkosyl.

E. Deletions and point mutations in the acidic region and C-tail were expressed and purified from E. coli and tested for DNA replication in vitro as described in Materials and Methods. 200 ng of each mutant protein was incubated in replication extract in the presence of $\alpha^{32}$P-dCTP and analyzed by agarose gel electrophoresis.

F. The arginine substitutions in the acidic region were tested for DNA replication in vitro. Three quantities (100, 200 and 400 ng) of wt E1 and the mutants 579R/583R, 586R, 579R/583R/586R and E1,577 were used in cell free replication reactions in the presence of $\alpha^{32}$P-dCTP and analyzed by agarose gel electrophoresis.

G. Inactivation of the acidic region in the context of the E1$_{308-605}$ C$_{596}$A restores E1 hexamer formation and ATPase activity to wt levels. Mutants in the C-terminal module were tested for hexamer formation in the context of E1$_{308-605}$ using an ssDNA probe. Two, four and eight ng each of the proteins E1$_{308-605}$, E1$_{308-605}$ C$_{596}$A, and E1$_{308-605}$ C$_{596}$A D$_{579}$/D$_{573}$/R586R were incubated with a 41-mer ssDNA probe (~2 fmol) in the presence of ADP and tested for hexamer formation by EMSA. The migration of the E1 hexamer is indicated. The altered mobility of the complexes formed with E1$_{308-605}$ C$_{596}$A D$_{579}$/D$_{573}$/R586R is caused by the E586R substitution.
H. Mutants in the C-terminal module in the context of E1_{308-605} were tested for ATPase activity. The proteins E1_{308-605}, E1_{308-605} C596A, and E1_{308-605} C596A D579R/D573R/E586R were incubated with γ-32P ATP and ssDNA, unless otherwise indicated, and tested for ATPase activity as described in the Materials and Methods. The migration of ATP and free phosphate is indicated and the level of hydrolyzed ATP is shown below the lanes.

Figure 6. The C-terminal module interacts with the E1 oligomerization domain.

A. E1_{308-577} and luciferase were translated in vitro in a reticulocyte lysate in the presence of 35S methionine and used in pull down experiments with GST, GST_{544-577} 579R/583R/586R, or GST_{544-605} and analyzed by SDS-PAGE followed by autoradiography. In the input lane, 1% of input was loaded. B. The same gel as in A was stained with Comassie brilliant blue (CBB) to show the presence of the GST fusion proteins.

C. E1_{308-410} was translated in vitro in a reticulocyte lysate in the presence of 35S methionine and used in pull down experiments with GST, GST_{544-577} 579R/583R/586R or GST_{544-577} and analyzed by SDS-PAGE followed by autoradiography. D. The same gel as in C was stained with CBB to show the presence of the GST fusion proteins.

E. The C-tail is required to sense the presence of ssDNA. Three quantities (2.5, 5 or 10 ng) of wt E1 or the mutants F594A or C596A were used in EMSA using two different probes. Wt E1 was incubated with the wt 84 bp ori probe in the presence of ADP or ATP. Wt E1 or the mutants F594A or C596A were incubated with a probe that differed from
the wt probe by a 6 bp mis-matched region (bubble probe) in the presence of ADP.

Migration of the DT and DH complexes is indicated.

Figure 7. Effects of mutations in the C-terminal module on E1 complex formation, and a model for how the C-terminal module controls oligomerization of E1.

A. A summary of how point mutations and deletions in the C-terminal module affect hexamer (H), double trimer (DT) and double hexamer (DH) formation in the context of E1 \textsuperscript{308-605} and E1 \textsuperscript{1-605}.

B. A model for how the acidic region and C-tail controls oligomerization of E1. See text for details.


Figure 1
Figure 2

A Helicase domain

C-terminal module

Acidic region

B

C

D

Conserved residues

577+7xA

Carbonic anhydrase (29 kD)

BSA (66 kD)

E16

A595

Figure 2
Figure 3
Figure 5
Figure 6
Forms neither DT nor DH. The exposed oligomerization domain prevents DT formation. Because DT cannot form, DH cannot form.

Forms DT but not DH. Blocked oligomerization domain allows DT formation. In the absence of the C-tail the oligomerization domain can not be unblocked to allow DH formation.

Forms DT and DH. Blockage of the oligomerization domain by the acidic region allows DT formation. In the presence of ssDNA, the oligomerization domain can be unblocked by the C-tail, which allows DH formation.