The recognition of local DNA conformation by the human papillomavirus type 6 E2 protein

Elizabeth Hooley, Victoria Fairweather, Anthony R. Clarke, Kevin Gaston and R. Leo Brady*

Department of Biochemistry, University of Bristol, Bristol BS8 1TD, UK

ABSTRACT

The E2 proteins are transcription/replication factors from papillomaviruses. Human papillomaviruses (HPVs) can be broadly divided in two groups; low-risk HPV subtypes cause benign warts while high-risk HPVs give rise to cervical cancer. Although a range of crystal structures of E2 DNA-binding domains (DBD) from both high- and low-risk HPV subtypes have been reported previously, structures of E2 DBD:DNA complexes have only been available for high-risk HPV18 and bovine papillomavirus (BPV1). In the present study we report the unliganded and DNA complex structures of the E2 DBD from the low-risk HPV6. As in the previous E2–DNA structures, complex formation results in considerable bending of the DNA, which is facilitated by sequences with A:T-rich spacers that adopt a pre-bent conformation. The low-risk HPV6 E2–DNA complex differs from the earlier structures in that minimal deformation of the protein accompanies complex formation. Stopped-flow kinetic studies confirm that both high- and low-risk E2 proteins adapt their structures on binding to DNA, although this is achieved more readily for HPV6 E2. It therefore appears that the higher selectivity of the HPV6 E2 protein may arise from its limited molecular adaptability, a property that might distinguish the behaviour of E2 proteins from high- and low-risk HPV subtypes.

INTRODUCTION

Human papillomaviruses (HPVs) are icosahedral DNA tumour viruses that infect epithelial and squamous tissue [reviewed in (1)]. Although there are >100 types of HPVs, the viruses can be divided into two groups: high-risk types such as HPV16 and HPV18 are associated with cancer and low-risk types such as HPV6 which are the causative agent of benign warts. The viral protein E2 is a transcription/replication factor that regulates production of the viral oncogene products E6 and E7 (2). High-risk HPVs often integrate into the host genome during tumourigenesis and this usually leads to disruption of the E2 gene and concomitant loss of E6 and E7 regulation. E2 also plays roles in viral genome replication and segregation by forming a complex with the helicase E1 (3) and by binding to host mitotic chromosomes (4).

The E2 protein is homodimeric, each monomer comprising ∼360 amino acids with a mass of 42 kDa. E2 consists of three domains—the N-terminal transcription activation/replication domain (TAD), a central hinge region and a C-terminal DNA-binding domain (DBD). Although a crystal structure of the whole E2 protein has not been reported, there are several structures from different papillomaviruses of both the N-terminal domain (5,6) and DBD (see below). The N-terminal domain is ∼200 residues in length, forms a dimer and is conserved throughout all HPV types. This domain interacts with the helicase E1 and recruits this protein to the viral origin of replication to facilitate viral genome replication (3). The proline rich central hinge region has ∼90 amino acids but is poorly conserved throughout the E2 proteins. However, there is evidence that the central hinge region plays more than just a structural role (2). The C-terminal DBD is basic (isoelectric point = 9.8 in the case of HPV6 E2) and forms a dimer of mass 20.4 kDa. We have reported previously the crystal structure of the low-risk HPV6 E2 DBD in the absence of DNA [PDB code 1r8h (7)] confirming that the structure is highly similar to other E2 DBDs [crystal structures of E2 DBDs have also been reported for the high-risk HPV16 (8), HPV31 (9), HPV18 (10) and the bovine virus BPV1 (11), including the latter two in complex with their cognate DNA sequences (10,11)]. HPV6 E2 DBD shares 56, 49, 50 and 33% sequence identity, respectively, with each of these proteins. Solution-based (NMR) structures have also been reported for BPV1 (12), HPV31 (13) and HPV16 E2 DBDs (14), including most recently the latter in the presence of DNA (15). The DBD dimer is formed around a central eight-stranded anti-parallel β-barrel structure, with four strands donated from each monomer. Each subunit contains a DNA-binding helix that binds two successive major grooves of the DNA target (10,11). Comparisons of the unliganded and complexed forms of HPV18 and BPV1 E2 proteins show that quaternary structural rearrangements within the E2 dimer accompany

*To whom correspondence should be addressed. Tel: +44 117 954 6852; Fax: +44 117 928 7436; Email: L.Brady@bristol.ac.uk

DNA binding (10,11). However, no structure has up till now been reported for an E2 DBD from a low-risk HPV in complex with DNA.

There are four E2-binding sites in the long control region (LCR) of the HPV genome. The consensus sequence is 5′-ACCgN$_4$cGGT-3′ (where upper case letters represent bases absolutely required, lower case letters represent bases that are preferred but not essential and N$_4$ is the variable central spacer region). The crystal structures of the E2–DNA complexes (10,11) confirm that, whereas direct contacts are made between E2 and the sequence-specific required bases, there are no direct contacts to the central spacer. The E2 DBD from the low-risk HPV6 preferentially binds a consensus sequence containing an A:T-rich central spacer region (e.g. AATT). Although this property is also shared by E2 DBD from the high-risk HPV16, the latter will also bind to sites with central spacers of sequence ACGT and CCGG, but with a 10-fold reduction in affinity compared to the 1000-fold reduction in affinity for the HPV6 protein (7,10). BPV1, on the other hand, appears to have no preference for particular central spacer sequences (16).

DNA is intrinsically flexible; however, some DNA sequences are more flexible than others. E2 DBD binding to its DNA target sequence has been shown to lead to bending of the DNA by up to 50° (10,11). This bending leads to narrowing of the minor groove of the central spacer region. HPV6 E2 DBD binds sequences containing A-tracts (4–6 consecutive A or T residues) which are intrinsically bent (17). This intrinsic bending becomes more pronounced on association with E2, leading to an even narrower minor groove. It therefore appears that E2 DBD–DNA binding is dependent not only on the conserved regions of the binding site forming contacts with the DNA-binding helices, but also on the conformation of the central spacer and inherent degree of flexibility of the DNA sequence. A recent study has shown that the binding of HPV16 E2 and BPV1 E2 to sites with different spacer regions is sensitive to the presence of cations that are presumed to stabilize bent DNA conformations by localizing within the narrowed minor groove (18). In both the BPV1 and HPV18 DNA complexes, deformation (especially in the placement of the DNA-binding helices) of the protein also accompanies binding, and partial ordering of the β2–β3 loop is observed in the BPV1:DNA complex. These observations are supported by the NMR studies in which the recognition helix is noted to be well defined but prone to rapid amide-exchange, and the β2–β3 consistently disordered in the absence of DNA ligand (12). The present consensus is therefore that association of E2 with its target DNA sequence is achieved through deformation of both the protein and DNA structures.

Here we present crystal structures of the low-risk HPV6 E2 DBD in the absence of DNA, and bound to 16 and 18 bp forms of its cognate DNA target (5′-CAACCG**AATT**CGGT-TG-3′ and 5′-GCAACCG**AATT**CGGTTGC-3′, respectively), spacer sequence shown in bold. These structures permit a comparison of the crystal conformation of HPV6 E2 DBD both in the presence and absence of its target DNA. Additionally, using stopped-flow fluorescence techniques we have monitored the binding of this protein–ligand pair in solution. Both of these methods confirm the HPV6 E2 DBD protein undergoes a small conformational rearrangement on DNA

binding, as has been noted previously for the BPV1, HPV16 and HPV18 E2 proteins. However, the crystal structures indicate that the extent of this structural change is limited for HPV6 E2 implying that the low-risk HPV6 E2 protein may have limited adaptability and therefore be more selective in its association with DNA.

MATERIALS AND METHODS

Protein purification

The HPV6 and HPV16 E2 DBD proteins were expressed and purified as described previously (7). In outline, HPV6 or HPV16 E2 DBD was expressed in *Escherichia coli* XL1-blue cells. An SP-sepharose column was used to purify HPV6 E2 DBD from the cell lysate. The eluted protein was then further fractionated on a MonoS 5/5 column and the protein was then dialysed into 20 mM Tris, pH 8.5, 100 mM NaCl and 10 mM DTT. For HPV6 E2 the fractions of the Mono S 5/5 column were dialysed into fresh column buffer and loaded onto a Heparin column. The protein was eluted off the column with a 0–100% 1.5 M NaCl gradient. The E2 protein fractions were dialysed into either the crystallization buffer (20 mM Tris, pH 8.5, 10 mM DTT and 100 mM NaCl) for structural studies or into 50 mM sodium phosphate, 150 mM NaCl and 2 mM DTT for stopped-flow studies.

DNA purification

Oligonucleotides (10 μmol batches) were purchased from MWG-Biotech (Ebersberg, Germany). Samples of the 18mer were suspended in 25 mM Tris, pH 7.5, 100 mM NaCl and 0.02% sodium azide to a final concentration of 1 mM, and annealed by heating to 95°C for 2 min followed by slow cooling over 16 h to form duplex DNA.

The double-stranded DNA (dsDNA) was purified using a 1 ml ResourceQ column (Amersham Pharmacia) equilibrated in 25 mM Tris, pH 7.5 and 100 mM NaCl and eluted with a 0–700 mM NaCl gradient. Fractions containing the oligonucleotide were buffer exchanged and concentrated using a 3K Centricon concentrator (Millipore, Billerica, USA), the final DNA concentration being determined by A_{260} readings, assuming that dsDNA at a concentration of 50 μg/ml has an absorbance of $A_{260} = 1$ (19).

The 16mer oligonucleotide was purified using a similar procedure except that before annealing it was initially fractionated in its single-stranded form using a 1 ml Resource Q column (Amersham Pharmacia) equilibrated in 10 mM NaOH and eluted using a 10 mM to 1 M NaCl gradient.

Crystallization of E2 DBD/S

The alternative crystal form of HPV6 E2 DBD was fortuitously obtained in an attempt to co-crystallize E2 DBD with oligonucleotide. E2 DBD at a concentration of 2.5 mg/ml in 1× phosphate-buffered saline, 0.1% β-mercaptoethanol and 0.02% sodium azide was mixed at a 1:1.2 molar ratio with dsDNA in 25 mM Tris, pH7.4, 100 mM NaCl and 1 mM MgCl$_2$. After 20 min the complex was concentrated to 3 mg/ml in a 3 kDa Centricon concentrator (Millipore). Crystals—later found to only contain the protein—grew after 6 days at 18°C in 1–1.5 M ammonium sulphate,

Table 1. Summary of X-ray diffraction data collection and refinement statistics

	E2 DBD/S	E2 DBD–18mer	E2 DBD–16mer
Resolution range (Å)	17–2.3 (2.32–2.3)	50.0–3.1 (3.2–3.1)	50.0–3.2 (3.3–3.2)
Space group	C2	P6$_1$	P6$_1$
Unit cell	$a = 97.7$, $b = 106.9$, $c = 74.9$, $\alpha = \gamma = 90°$, $\beta = 121.7°$	$a = 73.4$, $b = 73.4$, $c = 109.2$, $\alpha = \beta = 90°$, $\gamma = 120°$	$a = 75.2$, $b = 75.2$, $c = 97.4$, $\alpha = \beta = 90°$, $\gamma = 120°$
Unique reflections	28323 (2402)	6087 (610)	5168 (570)
Redundancy	8.2 (7.0)	7.6 (7.1)	22.5 (22.5)
Completeness (%)	97.4 (82.7)	99.9 (99.8)	99.8 (100.0)
I/σ	37.3 (5.5)	17.9 (4.2)	19.8 (5.2)
R_{merge} (%)	5.3 (21.7)	11.8 (35.1)	13.8 (38.7)
Refined model			
R_{free}	25.4	28.8	29.7
R_{factor}	18.5	20.3	18.9
r.m.s.d. bond length (Å)	0.02	0.02	0.02
r.m.s.d. bond angle (°)	2.0	2.3	2.3

0.1 M Tris, pH 7.5–8.5 and 12% v/v glycerol. A solution of 1.1 M ammonium sulphate, 0.08 M Tris, pH 8.0 and 20% v/v glycerol was used as a cryoprotectant. X-ray diffraction data were collected at Daresbury SRS synchrotron on beamline PX 10.1 ($\lambda = 1.284$ Å) to a resolution of 2.3 Å. Data were processed using HKL2000 (20) in the space group C2 with a unit cell of $a = 97.7$, $b = 106.9$, $c = 74.9$, $\alpha = \gamma = 90°$, $\beta = 121.7°$. The structure was solved by molecular replacement using PHASER (21) and the coordinates of the previously solved structure of HPV6 E2 DBD (PDB code 1r8h) as a search model. All model building was performed in COOT (22) and refinement in Refmac5 (23). The X-ray data processing and model refinement statistics are shown in Table 1.

Co-Crystallization of E2 DBD and its DNA target

E2 DBD at a concentration of 7 mg/ml in 20 mM Tris, pH 8.5, 100 mM NaCl and 10 mM DTT was incubated in a 1:1.2 molar ratio with dsDNA of the required length for 20 min at room temperature. Crystals of the E2 DBD–18mer complex grew in 0.1 M sodium chloride, 0.1 M HEPES, pH 7.4–7.6, 1.4–1.8 M ammonium sulphate and 10 mM DTT after 4–5 days. A solution of 20% glycerol, 0.08 M sodium chloride, 0.08 M HEPES, 1.3 M ammonium sulphate and 8 mM DTT was used as a cryoprotectant. Crystals of E2 DBD–16mer complex grew in 1.4–1.6 M tri-Sodium citrate dehydrate pH 6.5 and 10 mM DTT after 4–5 days. A solution of 10% glycerol in 1.4 M sodium citrate and 9 mM DTT was used as a cryoprotectant. As DNA binding to E2 DBD requires a reducing environment the crystals were grown in 10 mM DTT and frozen in liquid nitrogen once they had reached a suitable size for data collection. X-ray diffraction data were collected at Daresbury SRS on beamline PX 14.1 ($\lambda = 1.488$ Å, 18mer co-crystal) and using a Bruker Proteum R ($\lambda = 1.548$ Å, 16mer co-crystal). Both co-crystal datasets were processed using HKL2000 (20) in space group P6$_1$. The structure was solved by molecular replacement using PHASER (21) and the coordinates of HPV18 E2 DBD bound to its DNA target as a search model [PDB code 1jj4, (10)]. Model building was performed in COOT (22) and QUANTA and refinement in Refmac5 (23). The X-ray data processing and model refinement statistics are shown in Table 1.

Solution-binding studies

In both the HPV6 E2 and HPV16 DBD E2 sequences there are two conserved tryptophans, at positions 317 and 319. These are located close to the core of the dimer and immediately below the DNA-binding site. Previous studies (24) on HPV16 E2 have indicated a change in fluorescence of these tryptophans can be detected on DNA complex formation. In this study we measured DNA–E2 DBD complex formation in solution for both HPV6 E2 DBD and HPV16 E2 DBD by monitoring tryptophan fluorescence emission at 340 nm following excitation at 285 nm under stopped-flow conditions using an Applied Photophysics SpectraKinetic Monochromator and Workstation. The protein dimer concentration was kept constant at 0.2, 0.05 or 0.5 μM in 50 mM sodium phosphate, 150 mM NaCl and 2 mM DTT. The DNA concentration was varied in integral steps from 0.5- to 15-fold molar excess of the protein concentration. The stopped-flow chamber was kept at a constant temperature of 25°C, and data collected at intervals from 500 to 50 ms with the filter set to 1% of the timescale and a voltage offset of 4 V. Approximately 25 sets of data were measured for each concentration of DNA and then averaged using Grafit 3 software (Erithacus Software, Staines, UK). The initial data were fitted to a double exponential curve with offset:

$$y = A_1 * \left(1 - e^{(-k_1 * t)}\right) + A_2 * \left(1 - e^{(-k_2 * t)}\right) + c,$$

where A_1 = Amplitude 1, A_2 = Amplitude 2, k_1 = fast rate, k_2 = slow rate and c = Offset.

The rates obtained were then plotted against the DNA concentration and fitted to a straight line from which the forward (gradient) and reverse (y-intercept) rates were derived. These data were combined with the published binding constants for HPV6 E2 and HPV16 E2 (7) to the oligonucleotide to derive an overall model for association.

RESULTS AND DISCUSSION

Three structures of the E2 DBD from the low-risk HPV type 6a (HPV6 E2 DBD) have been solved in this study. One is an unliganded (no DNA) structure that contains disulphide bonds between the DNA-binding helices (HPV6 E2 DBD-S). The remaining two are structures of HPV6 E2 DBD bound to
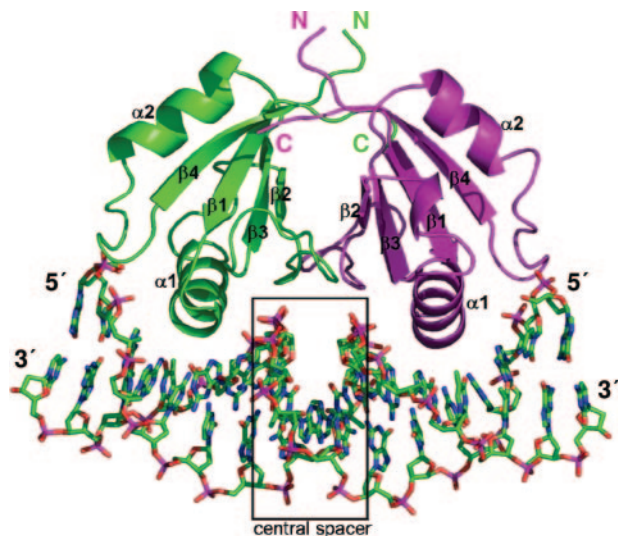
Figure 1. The E2 DBD of HPV6a bound to its DNA target (18 bp in length). The secondary structure elements are labelled and the AATT central spacer sequence is boxed. The N- and C-termini of the protein chains are labelled N and C, respectively, and 3′ and 5′ termini of the DNA labelled accordingly. Figures 1, 2 and 4 were prepared with PyMol [www.pymol.org (37)].

the dsDNA target sequences 5′-GCAACCG**AATT**-CGGTTGC-3′ (18mer) and 5′-CAACCG**AATT**CGGTTG-3′ (16mer). We have shown previously that HPV6 E2 DBD has an affinity of 2.0 ± 0.9 nM for a slightly extended (20mer) form of this DNA sequence (7).

The structure of HPV6 E2 DBD in the absence of DNA

We have reported previously the structure of unliganded HPV6 E2 DBD and shown that, like previous structures of E2 DBD proteins, this DBD of E2 forms a homodimer built from a central eight-stranded antiparallel β-barrel and four α-helices [as seen for HPV6 E2 in its complex with DNA in Figure 1]. Four strands of the β-barrel are donated from each monomer, and each monomer contains two α-helices, one of which inserts into the DNA major groove and will be referred to as the DNA recognition helix. The secondary structure elements occur in the order β1–α1–β2–β3–α2–β4. α1 is the DNA recognition helix and the dimer interface is between the β2 and β4 strands of each monomer.

A prominent feature of the novel crystal form of the unliganded E2 DBD crystal structure solved in this study (E2 DBD-S) is the presence of disulphide bonds between the DNA-binding helices of neighbouring molecules in the crystal lattice. These stabilizing disulphide bonds are made across the crystallographic 2-fold symmetry axis and form between Cys295 in one helix and Cys298 from the helix in the neighbouring molecule. Both of these cysteine residues are shown to form close associations with the DNA chain in the DNA complex structures (see below). Equivalent disulphide bonds are also present in the structure of unliganded BPV1 E2 DBD (11). Oxidation of these cysteine residues to form disulphide bonds would prevent the binding of DNA to the protein dimer. It is possible that oxidation of the cysteine residues acts as a mechanism to regulate the activity of the E2 protein, ensuring that the protein is only active in reducing

environments such as within the cell nucleus. Previous studies have shown that E2 proteins are sensitive to the redox environment both *in vivo* and *in vitro* (25). Redox regulation of many other RNA-binding and DNA-binding proteins has also been described in the literature.

E2 DBD-S crystallized in the space group C2 whereas the previous unliganded HPV6 E2 structure was determined in the space group P6_1 (7). Despite differences in the molecular packing between these crystal classes, and changes in the crystal contacts made by residues from the recognition helix including the formation of the intermolecular disulphide bond, there is no significant change in the quaternary structure of HPV6 E2 DBD between these structures (average r.m.s.d. ∼0.6 Å when dimers of apo DBD and apo DBD-S are compared based on 173 equivalent Cα atoms; the same value is obtained when comparing any three of the dimers from within each of the respective asymmetric units with one another). As the DNA-binding helices are central to E2:DNA complex formation, Hegde *et al.* (11) have compared E2 structures previously by overlaying the helix from one subunit, then using the relative displacement of the DNA-binding helix on the adjacent, non-superimposed subunit as a measure of conformational change. Using this criterion, the mean displacement of the non-superimposed recognition helix between the two HPV6 E2 unliganded structures is 0.9 ± 0.3 Å (based on 12 equivalent Cα atoms for the three copies of dimer in the asymmetric unit; Table 2). As this helix is also the site of the disulphide formation, this small displacement is not surprising and but of limited significance given the variation between each of the dimers in the respective asymmetric unit cells (∼0.6 Å) and the resolution of these structures (1.9 and 2.3 Å, respectively). These data suggest that the crystal contacts made by the helix (in this case including a covalent disulphide bond) do not overtly influence its placement. A similar finding has been reported previously for the unliganded BPV1E2 DBD, where the structure in a range of different crystal lattices has been found to be essentially identical [0.4 Å r.m.s.d. (11)]. Although the possibility always needs to be considered that crystal lattices may trap proteins in non-minimum energy conformations, this does not appear to be the case for E2 DBD for which highly similar structures have been observed in many different crystal lattices.

In contrast to all other reported structures of unliganded E2 DBDs, in both of these unliganded structures for HPV6 E2 DBD the β2–β3 loop is ordered. All other published E2 DBD structures have disordered β2–β3 loops in the DNA free state (8,10–14). This loop in HPV6 E2 DBD-S is loosely ordered (residues 323–328 are modelled at 50% occupancy in the electron density), makes only weak, water-mediated contacts in both crystal lattices and yet adopts a similar conformation in both unliganded structures, and in the HPV E2:DNA complex (Figure 2, described below). As has been noted previously (7), the HPV 6 β2/β3 loop conformation appears to be maintained primarily through the inclusion of prolines at positions 322 and 325, and it seems likely that the burial of the large hydrophobic tryptophan side chains at positions 317 and 319 at the base of the loop may also stabilize its conformation. Although the equivalent loop is not ordered in previous E2 DBD:DNA complexes, a recent mutagenesis study of HPV16 E2 DBD suggested residues Lys-325

Table 2. The global bend angles and mean displacements of DNA-binding helices of various E2 DBD–DNA complexes

Protein	DNA sequence	PDB Code	Global bend[a] (°)	Minor groove width (Å)[b]	pI of DNA binding helices
HPV6 E2 DBD (16mer)	5′-CAACCG**AATT**CGGTTG-3′	2AYB	24	2.8	9.3
HPV6 E2 DBD (18mer)	5′-GCAACCG**AATT**CGGTTGC-3′	2AYG	28	2.8	9.3
HPV18 E2 DBD	5′-CAACCG**AATT**CGGTTG-3′	1JJ4	43	2.7	11.0
HPV18 E2 DBD	5′-CAACCG**ACGT**CGGTTG-3′	Not deposited. Ref. (30)	49	2.7	11.0
BPV1 E2 DBD	5′-CCGACCG**AATT**CGGTCG-3′	R. S. Hegde, personal communication	47	4.0	10.5
BPV1 E2 DBD	5′-CCGACCG**ACGT**CGGTCG-3′	2BOP	51	4.0	10.5
None	5′-ACCG**AATT**CGGT-3′	1ILC	8.1	3.4	—
None	5′-ACCG**ACGT**CGGT-3′	423D	0.4	6.1	—

The minimal minor groove widths [calculated as described in (17)] of the DNA duplex, the pI of the relevant DNA-binding helices and their displacements between unliganded and DNA-complexed forms are shown. The global bend angles for HPV6 E2 were calculated in Madbend (33) and the remainder are as reported in (10).
[a]Global bend angles for HPV18 and BPV1complexes are from (10) or (30). Those for HPV6 complexes were calculated using Madbend (33), which gave similar values for HPV18 and BPV1 to those previously reported.
[b]Calculated as closest distance between equivalent phosphate groups from opposing base pairs, less two times the atomic radius of a phosphorous atom.
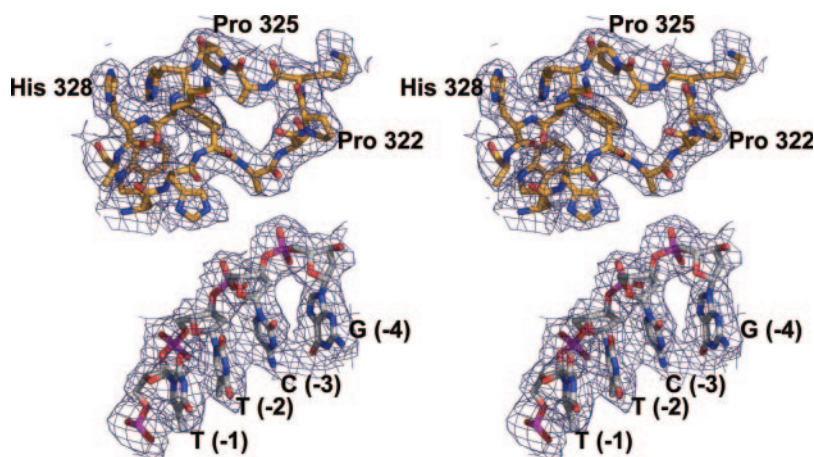
Figure 2. Electron density for the β2–β3 loop and surrounding regions, including the bound DNA. The electron density map shown is calculated from $2F_o–F_c$ coefficients, and is contoured at ∼2σ. The residues in the β2–β3 loop (323–328) are seen to the upper right projecting away from the DNA, a segment of one strand of which is shown below and numbered as in the schematic representation in Figure 3.

and Lys-327 in this loop participate in coulombic interactions with the backbone phosphates of the DNA (26).

HPV6 E2 DBD bound to its DNA target sequence

In this study the crystal structures of HPV6 E2 DBD bound to 18mer and 16mer dsDNA have been solved. Both of these structures are identical, except for the area where neighbouring DNA helices are in contact with each other to form an extended helix through the crystal. In the 18mer structure symmetry related DNA helices are slightly staggered, whereas in the 16mer structure the DNA helix is continuous throughout the crystal. This latter complex was prepared after solution and inspection of the 18mer complex, which suggested that a shorter segment of DNA could be more readily incorporated within the crystal lattice leading to a more ordered crystal.

As has been described previously for the BPV1 and HPV18 E2:DNA complexes, the DNA recognition helices of the HPV6 E2 DBD dimer bind to two successive major grooves of the target DNA sequence (Figure 1). Protein–DNA contacts are limited to the conserved regions of the target sequence (Figures 1 and 3), and no direct contacts are made with the central spacer of the DNA sequence. In the E2–DNA complex, the DNA is seen to wrap around the protein with a global bend angle of ∼24–28° (Table 3). The length of the DNA may also influence the bend angle as the 18mer bound to HPV6 E2 DBD has a slightly greater bend angle (28°) than the 16mer (24°), although this is unlikely to be significant given the resolution of the structures.

Direct DNA contacts are made between eight amino acid side chains in each recognition helix and to backbone phosphate or base atoms of the DNA. Most of these amino acids are highly conserved throughout the papillomavirus family. Owing to the limited resolution of the diffraction data (∼3 Å) water-mediated protein–DNA interactions cannot be reliably assessed. Although solvent molecules are believed to play an important role in E2 DBD–DNA binding [e.g. see (15)], water molecules have only been conservatively included in these crystal structures (12 in the 16mer complex, 19 in the 18mer complex). For this description, the DNA base numbering of the 18mer structure will be used (with numbers in brackets referring to their position from the centre of the 5′–3′oligonucleotide) and is shown in Figure 3.

Table 3. Displacement of recognition helices on DNA binding

| E2	Resolution of structures (complex/apo) (Å)	Mean displacement of recognition helix (Å)
HPV18 (AATT)	2.3/1.9	3.5 ± 0.2
BPV1 (AATT)	2.3/2.5	2.6 ± 0.2
BPV1 (AGCT)	1.7/2.5	2.8 ± 0.5
HPV16	(NMR)	6.2 ± 0.7
HPV6 (AATT, 16mer)	3.2/1.9	1.9 ± 0.1
HPV6 (AATT, 16mer versus 18mer)	3.2/3.1	1.3 ± 0.2
HPV6 apo versus HPV6-S apo	1.9/2.3	1.2 ± 0.1

Displacements were calculated by overlaying the DNA-binding helix Cα's of one subunit (residues 293–304 in HPV18 and HPV6, 325–346 in BPV, and 11–22 in HPV16) of the unliganded and DNA-complexed crystal structures, then measuring the mean displacement of the equivalent 12 Cα atoms of the same helix from the other subunit. For each pair the mean displacements of the helices used for the overlay were <0.5 Å. The errors were estimated as the standard deviations of the mean.

Starting from the 5′ end of the oligonucleotide (Figure 3) the Arg302 guanidinium group forms charged hydrogen bonds with the backbone phosphate of Cyt2 (+8) and a bifurcated interaction with both the base and phosphate group of Ade3 (+7). The phosphate group of Ade3 (+7) also interacts with the side chain of Thr353. Further along the oligonucleotide chain, the base of Cyt5 (+5) is hydrogen bonded to the side chain of Asn294. Consistent with the HPV18 E2 structure, there are then no direct contacts made between the protein and DNA in the stretch from Cyt6 (+4) through to Thy10 (−1) inclusive (a region that incorporates the AATT cognate sequence).

Contacts resume at Cyt12 (−3), where the phosphate backbone makes a series of charged bonds with both the side chain and main chain of Thr316, and Arg300 side chain. This latter side chain also contacts the phosphate group of Gua13 (−4), which in turn is also bonded to the Lys297 side chain. The base of Gua14 is in van der Waals contact (3.2 Å) with the Tyr301 side chain, from which the hydroxyl group also forms a hydrogen bond with the Lys297 side chain amine group. The remainder of the nucleotides in the chain is projected away from the protein surface. As the structure is ~2-fold symmetric, most of these contacts are replicated on the reverse (3′–5′) strand. No contacts are made between the DNA and residues in the ordered β2–β3 loop.

HPV E2 DBD has two cysteine residues: Cys298 and Cys295. These residues form disulphide bonds in the unliganded E2 DBD-S structure, but have been suggested previously to be important for DNA binding (11,25). Neither of these amino acids make direct contact with the DNA in the HPV6 E2 DBD:DNA complex. However, Cys298 lies close (~4 Å) to the bases from Ade3 (+7), Ade4 (+6), Cyt5 (+5) and Gua14 (3′–5′ strand) and Cys295 is a similar distance from the phosphate group of Ade3 (+8). In all cases water-mediated contacts may be feasible.

A comparison of the unliganded HPV6 E2 DBD structures with those incorporated within the DNA complexes shows that only three of the residues in the DNA-binding helices significantly change conformation on complex formation: Lys297, Tyr301 and Arg302. These are surface-exposed amino acids that exchange crystal contacts in the apo structures for contacts with DNA in the complexes. Both
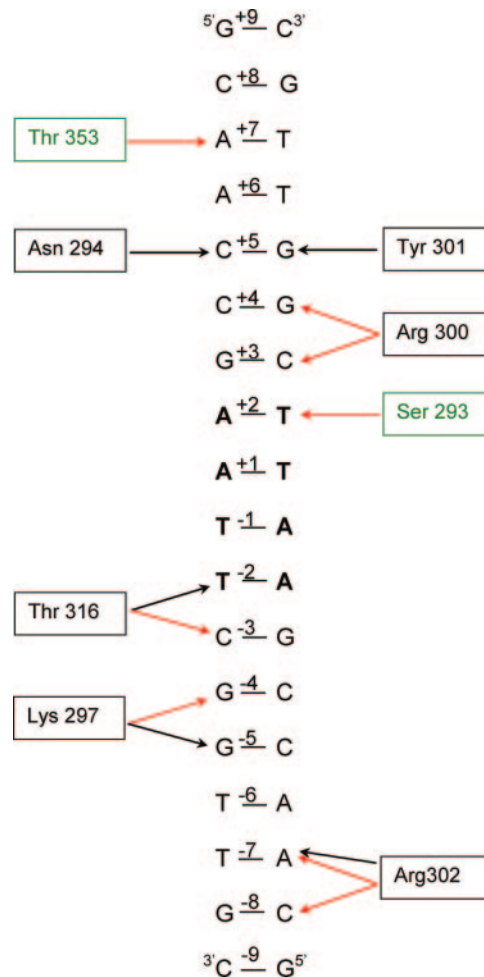
Figure 3. Schematic representation showing the contacts between the DNA-binding helices of HPV6 E2 DBD and the 18 bp DNA target. The central spacer is highlighted in bold. Each of the bases of the conserved region either side of the central spacer have contacts with the DNA-binding helices. Red arrows represent bonding to backbone phosphate atoms, black arrows represent bonding to the bases. Thr353, Ser293 and Glu292 are highlighted in green as these contacts are unique to HPV6 E2 DBD. For simplicity, each amino is shown only once.

Lys297 and Tyr301 are required to move in order to prevent steric clashes with the DNA, and both adopt alternative conformations that enable the formation of effective hydrogen bonds with the DNA. Arg302 also alters its conformation so that the side chain is closer to the DNA and improved contacts can therefore be made. Both Tyr301 and Arg302 are in the vicinity of the ±7 bases, noted previously to form an extension to the E2 consensus-binding site (7,27). Binding data have indicated previously that A:T is preferred at the −7:+7 sites, respectively. In this crystal structure the Arg302 makes direct contacts with both the base and backbone phosphate group in the +7 position, explaining the base preference noted in this extended binding site region. A similar observation has been noted in the crystal structure of the HPV18 E2:DNA complex (Arg 305 in HPV18 E2) (10).

The majority of the DNA-binding residues are conserved between all E2 DBD domains; however, there are some differences. Asp295, Arg296 and Ser355 of HPV18 E2

DBD are replaced by Glu292, Ser293 and Thr353, respectively, in HPV6 E2 DBD. The latter two interact directly with the DNA backbone; and Glu292 does not interact with the crystal structure but is sufficiently close to possibly establish a water-mediated contact. Arg307 of HPV18 E2 DBD directly contacts the DNA backbone; however, this contact is lost in HPV6 E2 DBD as this residue is Asn304 which is too short to directly reach the DNA although it may do so via a water molecule. Phe343 of BPV1 E2 DBD is replaced by Tyr301 in HPV6 E2 DBD; both form van der Waals contacts with the DNA. The interactions between DNA and residues Glu292, Ser293 and Thr353 in HPV6 E2 DBD are not present in the BPV1 E2:DNA complex as these positions are occupied by Thr334, Ala335 and Gly398, respectively.

The solvent-accessible surface area [calculated with Areaimol (28)] buried within the 16mer and 18mer complexes is 1260 and 1374 $\mathring{A}^2$, respectively. This is very similar to the buried surface in the HPV18E2:AATT complex (1223 $\mathring{A}^2$) but less than in the BPV1:AATT complex (1515 $\mathring{A}^2$). The similarity to the HPV18 E2 complex is consistent with the comparable binding affinities reported for E2 proteins from both high- and low-risk viruses.

The effect of DNA binding on the HPV6 E2 DBD conformation

It has been noted previously that formation of the E2 DBD:DNA complex results in deformation of both the protein and DNA structures (10,11). Hegde *et al.* (11) noted that changes in the BPV1 E2 protein structure equate to small rearrangements of the secondary structure elements rather than overall gross structural changes. As the DNA recognition helices are critical to the interaction of E2 with its target DNA, changes in the relative orientation of these helices have been used as a measure of the deformation the protein undergoes on complex formation. Hence, for BPV1 E2, monomers in the free and complexed forms can be readily superimposed (r.m.s.d. of 0.55 Å for backbone atoms). However, if one of the DNA recognition helices (α1) from the free and complexed subunits is overlaid, the equivalent helix from the non-superimposed subunits is seen to be displaced by up to 3.5 Å (11). In Figure 4, we show similar overlays for the BPV1, HPV18, HPV16 and HPV6 E2 DBDs. In each of these cases, structural data are available both for the unliganded form of the E2 DBD protein, and for the same protein in complex with its cognate DNA sequence. For these pairs of structures, we have overlaid one subunit using the positions of 12 Cα atoms from the α1 helix (Figure 4) and then measured the mean displacement for the same Cα positions from the equivalent helix from the non-superimposed subunit. The results are summarized in Table 2.

Both from the figure and these data, it is evident that HPV18 E2 undergoes a substantial rearrangement on complex formation. The extent of this distortion is reduced for BPV1 E2, and smallest for HPV6 E2. In each case the changes are most prominent in the surface features such as the α1 and α2 helices, there being minimal movement in the residues forming the dimer interface [as reported for BPV1 (11)]. Given the varying resolutions of the crystal structures used for this analysis, some caution is warranted
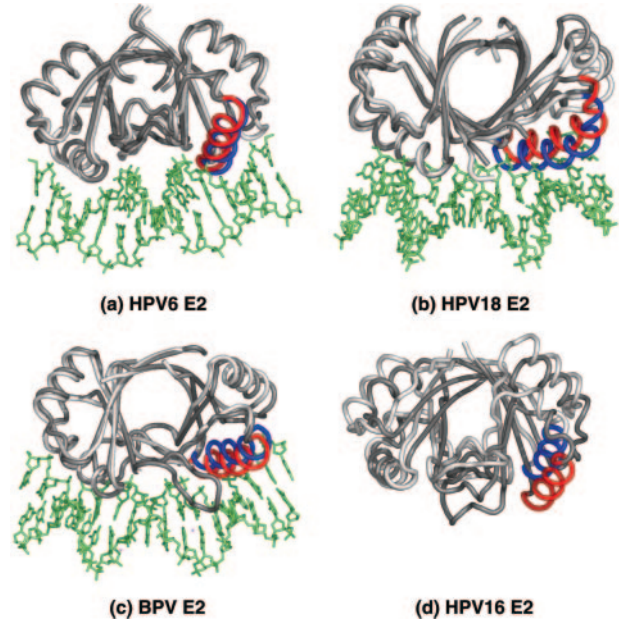
Figure 4. Rearrangement of the DNA-binding helices on DNA complex formation with (**a**) HPV6 E2 DBD, (**b**) HPV18 E2 DBD, (**c**) BPV1 E2 and (**d**) HPV16 E2 DBD. In each case E2 DBD is shown as a ribbon trace with the unliganded structure in dark grey overlaid on the structure of the DNA complex (light grey) to maximize the fit of the α1 helix from the left subunit. The resulting relative positions of the α1 helix from the right subunit are shown in red for the unliganded structure, and blue for the DNA complex structure. For HPV6, HPV18 and BPV1 the DNA is shown in stick representation in green. Note that for HPV16 both structures were determined by NMR spectroscopy but there was no experimental determination of the bound DNA conformation (hence no DNA is included in the model for the complexed form). The α1 helix amino acids used for the overlays were as listed in Table 3.

in these interpretations. Nonetheless, the α1 recognition helix is a prominent segment of secondary structure in the E2 DBD, and is consistently well defined in all of the crystal and NMR structures. Such elements are usually reliably placed even in medium resolution crystal structures. The multiple copies of the HPV6 E2 dimer within the two apo crystal structures provide an internal comparison whereby variations in displacement between the various dimers are consistently ∼0.6 ± 0.1 Å. This small variation comprises both inherent flexibility of the helix relative to the domain fold and errors within the crystal structures. Similarly, the availability of both a 16mer and 18mer structure for the HPV6 E2 DBD:DNA complex, both of restricted resolution (∼3 Å), allows a comparison of the helix displacement between these structures (1.3 ± 0.2 Å) reflecting both the small change in DNA bending angle between these structures (∼4°) and the increased error associated with the diminished resolution. The displacements measured between all pairs of liganded/unliganded E2 DBDs exceed these values.

The crystal structure of the unliganded E2 DBD from another high-risk papillomavirus, HPV16, is also available (8) as is its solution structure (14), and very recently its solution structure in the presence of cognate DNA has also been reported (15). The deposited coordinates have been included in our analysis (Table 2) and suggest there is a considerable shift in the relative placement of the recognition helices associated with DNA binding for HPV16 E2 (Figure 4d).

Nonetheless, this needs to be treated with caution, as NMR spectroscopy—based on local interactions—has limited accuracy for the analysis of long-range changes of this kind. However, a comparison of the unliganded HPV6 and HPV16 E2 DBD crystal structures showed a displacement of the DNA recognition helices by ∼7 Å (7), whereas in this same study circular permutation assays indicated that the DNA bend in the complexes that each of these proteins forms with their cognate DNA sequences is highly similar. HPV16 E2 DBD must therefore change to a similar conformation to that reported here for HPV6 E2 DBD in order to induce a similar bend in the DNA.

The effect of complex formation on the DNA structure

The DNA used for this study has a central AATT sequence, sometimes referred to as an A-tract (4–6 consecutive adenine-thymine residues). Solution studies and gel migration data have indicated previously that A-tract sequences are characteristically bent towards the minor groove (29). A crystallographic structure has been reported (17) for a dsDNA dodecamer corresponding to the central sequence of the 16mer and 18mer used in this study. All copies of this unliganded form of the E2 consensus sequence in the crystal lattice displayed global curvatures of 8–10°, with an unusually narrow minor groove (<4 Å) at the A-tract core. In contrast, dsDNA of sequence 5′-ACCG**A**CGTCGGT-3′ or 5′-ACCG**GTA**CCGGT-3′ has been shown by X-ray crystallography to have either a straight or slightly bent conformation with no preference for the direction of the bend (30). These degrees of curvature reported within crystal lattices approximate to those measured using a lower-resolution but solution-based method (31) and a predictive cyclization kinetics method (32).

Assuming these conformations are representative of the structures encountered within the cell, it is evident that formation of a complex with the E2 DBD requires the DNA to undergo significant further bending. Table 3 shows the extent of this bending as observed in the three E2 DBD:DNA crystal complexes. On association with HPV6 E2, the target sequence is required to bend an additional ∼20°, and the minor groove narrows further. This distortion, however, is less than that observed for BPV1 E2 and HPV18 E2 (each with net bends of ∼30°). The minor groove in the latter is similarly closely spaced to the DNA in the current study, implying a limit has been reached for narrowing of the groove.

The predisposition of A-tracts to adopt a pre-bent structure has been suggested to have entropic advantages promoting formation of the E2 DBD:DNA complex (7,11). The associated decrease in the width of the minor groove may also create a more complementary docking surface for the E2 DNA recognition helices: the dodecamer 5′-ACCG**ACGT**-CGGT-3′, with a non-optimal spacer sequence, has an increased minor groove width of ∼6 Å (Table 2). Both HPV18 and (especially) HPV6 E2 proteins have significantly reduced affinity for this sequence.

Surface electrostatic charges have been correlated frequently with the capacity of proteins to bind to DNA, and proposed to contribute to the ability of E2 DBD to promote bending of DNA (2). HPV6 E2 DBD is a basic protein

with an overall pI value of 9.8. Calculating a pI for the DNA recognition helices alone shows that HPV6 E2 DBD is less positive (pI = 9.3) than HPV18 E2 DBD (pI = 11.0) or BPV1 E2 DBD (pI = 10.5). This trend correlates with an analysis of surface charge distribution of these proteins (7) which indicated that the DNA-binding surfaces of HPV18 and BPV1 E2 DBDs carry a higher density of positive charge than either HPV6 or HPV16. An increase in positive charge would be expected to enhance the electrostatic interaction between the DNA-binding helices and the DNA backbone, and may therefore facilitate fitting of the DNA to the protein recognition helices. The program MadBend (33) was used to calculate the global bend angles of various DNA sequences. The centre of the central spacer was taken as a reference plane. The twist, roll and tilt angles were calculated in FREEHELIX (34) and incorporated into the MadBend angle calculation. Table 2 shows the results of these calculations, which are broadly in agreement with measurements of the bend angle observed in the crystal structures. From these data it can be seen that an increase in pI of the recognition helices correlates with an increased global bend angle.

Structural rearrangements of E2 DBDs on DNA binding: solution studies

Comparisons of the available crystal structures of unliganded and DNA-complexed E2 proteins indicate that structural rearrangements of E2 DBD protein accompany DNA binding (10,11). This has been further validated in solution studies of HPV16 E2 DBD (15) and by monitoring fluorescence of protein tryptophan residues and an incorporated label on the DNA (26). This latter study indicated the interaction of the high-risk HPV16 E2 DBD with cognate DNA could be separated into two phases: (i) a faster ($k = 8$ s^{-1}) non-specific association that is diffusion-controlled and appears to correspond to encounter complex formation, and (ii) a slower ($k = 0.04$ s^{-1}) phase that does not depend on protein concentration, is characterized by a decrease in fluorescence, and has been interpreted as resulting from a substantial change in the conformation of the protein after formation of the DNA complex. This latter phase is also accompanied by a slow enthalpic solvent exclusion step, to form the tightly associated protein–DNA complex. These data are consistent with the expectation, drawn from the crystallographic studies of HPV18 E2 DBD, that HPV16 E2 DBD must also undergo substantial rearrangement on complex formation. Because of the close structural similarity of the HPV6 E2 DBD unliganded and complexed forms, we applied a similar solution-based fluorescence analysis (although limited to intrinsic tryptophan fluorescence only) to both HPV6 and HPV16 E2 DBDs to explore whether a distinctive two-phase association was also observed for the low-risk viral protein.

Representative data and derived rate constants are shown in Figure 5 and Table 4, respectively. Under the conditions studied, both HPV6 and HPV16 E2 DBDs showed a typical increase in overall amplitude of fluorescence with an increase in DNA concentration, as has been noted previously for HPV16 E2 (24). The association curves for DNA binding could readily be dissected into fast and slow components by fitting to a double exponential rate equation, and the
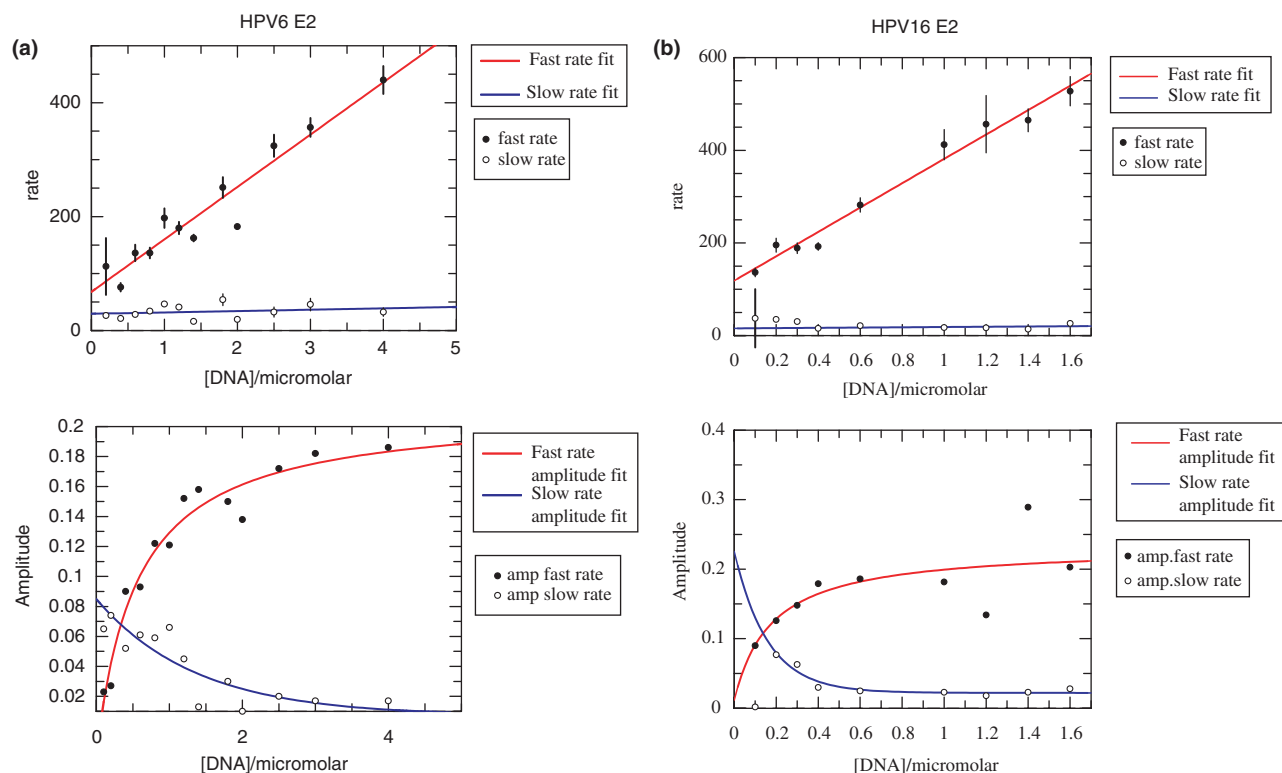
Figure 5. Stopped-flow fluorescence data for the binding of (**a**) HPV6 E2 DBD and (**b**) HPV16 E2 DBD to the DNA sequence. Changes in the amplitude of fluorescence measured at 340 nm were measured on the addition of aliquots of DNA to produce association curves for DNA binding, which were then fitted to a double exponential rate equation to dissect the fast and slow components (lower graphs). The rates obtained were then plotted against the DNA concentration (upper graphs).

Table 4. Kinetic data for E2:DNA complex formation obtained from stopped-flow fluorescence experiments

|	Fast rate k_{1on} (M^{-1} s^{-1})	k_{1off} (s^{-1})	Slow rate k_{2on} (s^{-1})
HPV6 E2	$(9.2 \pm 0.8) \times 10^7$	68 ± 16	29 ± 6
HPV16 E2	$(2.63 \pm 1.4) \times 10^8$	119 ± 14	16 ± 5

rates obtained were plotted against the DNA concentration (Figure 5). The slow component is characterized by a decrease in tryptophan fluorescence, consistent with a molecular rearrangement event. For both proteins, the fast rate is seen to increase linearly with increasing concentrations of DNA, whereas the slow rate remains approximately constant. We have interpreted these data according to a two-step model for E2:DNA association in which there is (i) a fast step representing the formation of the initial E2:DNA complex and (ii) a subsequent, slower step correlating with a molecular rearrangement of the protein to form a tight-binding complex. This model is summarized in Figure 6. For both HPV6 and HPV16 E2, the initial protein–DNA complex is loosely associated with a dissociation constant of ~0.5 μM. We note that the on-rate is about twice as fast for HPV16 E2; however, the off-rate is similarly elevated. These differences might be explained by the differing electrostatics of the presenting surfaces of the respective proteins, as discussed above. Both protein–DNA initial complexes then undergo a slower rearrangement step, likely

to include both the molecular adaptation of the DNA-binding helices as illustrated in the crystal structures, and expulsion of solvent from the interface as proposed previously (24). The final tight-binding complex achieved is of similar affinity (~2 nM) for both HPV6 and HPV16 E2 (7). However, the molecular rearrangement step proceeds at different rates. The forward rate is approximately twice as fast (30 s^{-1}) for HPV6 E2 than for HPV16 E2 (16 s^{-1}). This is consistent with a lower energetic barrier (ΔG) for molecular rearrangement in the HPV6 E2:DNA complex, correlating with the closer structural similarity of the unliganded and complexed crystal structures of this protein.

The rate constants derived for HPV16 E2 DBD in this study differ from those reported previously (26), although this is not surprising as the solution conditions and methods differ between these two studies. This previous study also supported a two-step model for E2:DNA association. However, their more detailed analysis obtained through the stronger signal from the label incorporated within the DNA permitted further dissection of the slow step.

HPV6 E2 DBD exhibits reduced deformability relative to high-risk E2 DBDs

In combination, these data confirm that the association of HPV6 E2 DBD with its cognate DNA—like HPV18, HPV16 and BPV1 E2 DBDs—is characterized by a multi-step procedure encompassing complex formation, DNA-bending and conformational rearrangement of the
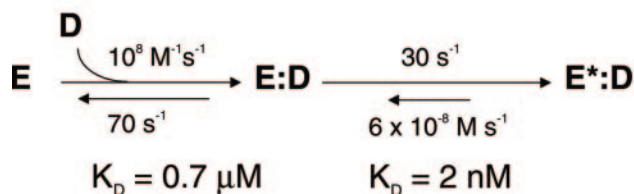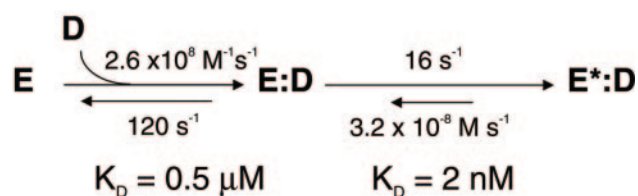
Figure 6. Proposed model for DNA complex formation for E2 proteins. In each case the E2 protein is abbreviated as E and the DNA as D. E:D represents the initial weakly associated complex, and E*:D the tight complex formed after molecular rearrangement of the protein.

DNA-binding helices of the protein. However, the crystallographic studies in particular support a model for HPV6 E2 DBD binding in which the protein has reduced deformability relative to the equivalent domains from BPV1, HPV18 and—most likely—HPV16. Firstly, the crystal structures indicate that domain rearrangements for HPV6 E2 are minimal (Figure 4 and Table 3) on complex formation. In contrast, much larger changes in the protein conformation are evident for both HPV18 and for HPV16—both from high-risk HPV. BPV1 exhibits more segmental flexibility than HPV6 E2, but less than seen in the HPV18 and HPV16 E2 structures. These changes are mirrored in the degree of DNA bending evident in the available crystal structures, where HPV6 E2 gives rise to the smallest structural change within the DNA. Overall, the interaction of HPV6 E2 with its DNA target is closer to the traditional 'lock and key' mode of molecular recognition than has been indicated for other E2 DBDs.

This model correlates with functional differences that have been reported previously to distinguish HPV6 E2 from its high-risk counterparts. Although E2 DBDs from HPV6 and HPV16 bind their cognate DNA sequences with similar affinity, HPV6 E2 is exquisitely sensitive to changes in the spacer sequence. Altering the spacer from AATT to ACGT or CCGG reduced binding of HPV16 E2 by about an order of magnitude, but by greater than three orders of magnitude for HPV6 E2 (7). As these latter sequences are known not to induce the preferred 'pre-bent' DNA conformation with narrowed minor groove approaching the conformation observed within the E2:DNA complexes, it appears that binding will be increasingly compromised for HPV6 E2 where the protein is less able to alter its own conformation to meet the restrictions imposed by these altered targets.

We also note that the reduced adaptability of the HPV6 E2 DBD correlates with a stabilized β2–β3 loop in this protein. In all other unliganded structures of HPV E2 proteins, this loop is disordered, and it remains so in the HPV18 E2:DNA complex. The BPV1 E2:DNA complex shows an ordered loop that interacts indirectly with the central spacer of the DNA sequence. This interaction is believed to stabilize the β2–β3 loop and reduce the electrostatic repulsion within the compressed minor groove. Mutagenesis data for HPV16 E2 also suggest loop residues from this protein make contacts to the DNA (26). However, in all structures of HPV6 E2 DBD, this loop is at least partially ordered, and no direct contacts to the DNA are observed. In HPV6 E2, the β2–β3 loop is shorter than in HPV18 E2, and is likely to be stabilized by the presence of two proline residues (Pro322 and Pro325) and additionally possibly through the formation of a salt bridge between Lys323 in the loop and Asp311 from the other molecule in the dimer. In the HPV6 E2:DNA complex, the loop remains in the same conformation, positioned such that no contact is made between loop residues and the DNA. Although it seems unlikely that this loop would have an effect on the stable central β-barrel structure, it is possible that its presence across the centre of the dimer might impose a restriction on the degree of movement accessible to the nearby α1 recognition helix.

Previous solution-based (NMR) studies of E2 DBDs have indicated that, although the α1 recognition helix is consistently well defined with average backbone r.m.s.d. values in all structures determined, most amide groups from this region exchange relatively quickly with solvent (12,13,15). This is believed to indicate that the helix is comparatively flexible within the E2 DBD. However, the high correlation in its placement when the six independent copies of the dimer from the two different apo HPV6 E2 structures are compared implies that the observed position reflects a genuinely preferred minimum energy conformation. Recent studies indicate that at least 90% of the binding energy between E2 DBD and its DNA arises from additive direct and water-mediated interactions between the protein and DNA (26). The adaptation of the helix positioning from a preferred location as observed in the apo structures to its bound conformation is consistent with this model. Although the helix may indeed need to sample a range of conformations in order to effect binding of its ligand, the closer proximity of the favoured bound conformation for HPV6 E2 implies this position could be more readily reached, as reflected in the kinetic data.

Biological implications of reduced flexibility in HPV6 E2

HPV E2 proteins are known to differ substantially in their ability to activate transcription, with E2 proteins from 'high-risk' papillomaviruses generally proving more potent activators (35,36). These differences might be attributable to differences in the E2 *trans*-activation domains, or could also reflect changes in the association of E2 subtypes with the viral DNA. There are four E2-binding sites in the HPV LCR, each at conserved positions relative to the transcription start point. It has been noted previously that there is a clear hierarchy in the occupation of these sites that distinguishes the high-risk HPV16 E2 protein from low-risk HPV6 E2 (7). Specifically, greater variation was observed in the binding of HPV6 E2 to the four sites than was found for

HPV16 E2 protein. These sites are most readily distinguished through the differing composition of their A-tract spacer sequences, which in turn implies differing propensities for forming pre-bent structures. The limited deformability of the HPV6 E2 DBD noted in the current study provides a molecular explanation for the reduced tolerance for spacer variation observed for this low-risk HPV protein. This higher level of selectivity in binding the viral genome may correlate with diminished ability to activate transcription.

Atomic coordinates

The coordinates and structure factors for E2 DBD–18mer, E2 DBD–16mer and E2 DBD-S have been deposited in the Protein Data Bank under the accession codes 2AYG, 2AYB and 2AYE, respectively.

REFERENCES

1. Dell,G. and Gaston,K. (2001) Human papillomaviruses and their role in cervical cancer. *Cell. Mol. Life Sci.*, **58**, 1923–1942.
2. Hegde,R. (2002) The papillomavirus E2 proteins: strucuture, function and biology. *Annu. Rev. Biophys. Biomol. Struct.*, **31**, 343–360.
3. Abbate,E.A., Berger,J.M. and Botchan,M.R. (2004) The X-ray structure of the papillomavirus helicase in complex with its molecular matchmaker E2. *Genes Dev.*, **18**, 1981–1996.
4. You,J., Croyle,J.L., Nishimura,A., Ozato,K. and Howley,P.M. (2004) Interaction of the bovine papillomavirus E2 protein with Brd4 tethers the viral DNA to host mitotic chromosomes. *Cell*, **117**, 349–360.
5. Harris,S.F. and Botchan,M.R. (1999) Crystal structure of the human papillomavirus type 18 E2 activation domain. *Science*, **284**, 1673–1677.
6. Antson,A.A., Burns,J.E., Moroz,O.V., Scott,D.J., Sanders,C.M., Bronstein,I.B., Dodson,G.G., Wilson,K.S. and Maitland,N.J. (2000) Structure of the intact transactivation domain of the human papillomavirus E2 protein. *Nature*, **403**, 805–809.
7. Dell,G., Wilkinson,K.W., Tranter,R., Parish,J., Leo Brady,R. and Gaston,K. (2003) Comparison of the structure and DNA-binding properties of the E2 proteins from an oncogenic and a non-oncogenic human papillomavirus. *J. Mol. Biol.*, **334**, 979–991.
8. Hegde,R.S. and Androphy,E.J. (1998) Crystal structure of the E2 DNA-binding domain from human papillomavirus type 16: implications for its DNA binding-site selection mechanism. *J. Mol. Biol.*, **284**, 1479–1489.
9. Bussiere,D.E., Kong,X., Egan,D.A., Walter,K., Holzman,T.F., Lindh,F., Robins,T. and Giranda,V.L. (1998) Structure of the E2 DNA-binding domain from human papillomavirus serotype 31 at 2.4 A. *Acta Crystallogr. D Biol. Crystallogr.*, **54**, 1367–1376.
10. Kim,S.S., Tam,J.K., Wang,A.F. and Hegde,R.S. (2000) The structural basis of DNA target discrimination by papillomavirus E2 proteins. *J. Biol. Chem.*, **275**, 31245–31254.
11. Hegde,R.S., Wang,A.F., Kim,S.S. and Schapira,M. (1998) Subunit rearrangement accompanies sequence-specific DNA binding by the bovine papillomavirus-1 E2 protein. *J. Mol. Biol.*, **276**, 797–808.
12. Veeraraghavan,S., Mello,C.C., Androphy,E.J. and Baleja,J.D. (1999) Structural correlates for enhanced stability in the E2 DNA-binding domain from bovine papillomavirus. *Biochemistry*, **38**, 16115–16124.
13. Liang,H., Petros,A.M., Meadows,R.P., Yoon,H.S., Egan,D.A., Walter,K., Holzman,T.F., Robins,T. and Fesik,S.W. (1996) Solution structure of the DNA-binding domain of a human papillomavirus E2 protein: evidence for flexible DNA-binding regions. *Biochemistry*, **35**, 2095–2103.
14. Nadra,A.D., Eliseo,T., Mok,Y.K., Almeida,C.L., Bycroft,M., Paci,M., de Prat-Gay,G. and Cicero,D.O. (2004) Solution structure of the HPV-16 E2 DNA binding domain, a transcriptional regulator with a dimeric beta-barrel fold. *J. Biomol. NMR*, **30**, 211–214.
15. Cicero,D.O., Nadra,A.D., Eliseo,T., Dellarole,M., Paci,M. and de Prat-Gay,G. (2006) Structural and thermodynamic basis for the enhanced transcriptional control by the human papillomavirus strain-16 E2 protein. *Biochemistry*, **45**, 6551–6560.
16. Hines,C.S., Meghoo,C., Shetty,S., Biburger,M., Brenowitz,M. and Hegde,R.S. (1998) DNA structure and flexibility in the sequence-specific binding of papillomavirus E2 proteins. *J. Mol. Biol.*, **276**, 809–818.
17. Hizver,J., Rozenberg,H., Frolow,F., Rabinovich,D. and Shakked,Z. (2001) DNA bending by an adenine–thymine tract and its role in gene regulation. *Proc. Natl Acad. Sci. USA*, **98**, 8490–8495.
18. Blakaj,D.M., Kattamuri,C., Khrapunov,S., Hegde,R.S. and Brenowitz,M. (2006) Indirect readout of DNA sequence by papillomavirus E2 Droteins depends upon net cation uptake. *J. Mol. Biol.*, **358**, 224.
19. Sambrook,J., Fritsch,E.F. and Maniatis,T. (1987) *Molecular Cloning: A Laboratory Manual, 2nd edn.* Cold Spring Harbor Press, Woodbury, New York, USA.
20. Otwinowski,Z.M.W. (1997) Processing of X-ray diffraction data collected in oscillation mode. *Methods Enzymol.*, **276**, 307–326.
21. Storoni,L.C., McCoy,A.J. and Read,R.J. (2004) Likelihood-enhanced fast rotation functions. *Acta Crystallogr. D Biol. Crystallogr.*, **60**, 432–438.
22. Emsley,P. and Cowtan,K. (2004) Coot: model-building tools for molecular graphics. *Acta Crystallogr. D Biol. Crystallogr.*, **60**, 2126–2132.
23. Winn,M.D., Isupov,M.N. and Murshudov,G.N. (2001) Use of TLS parameters to model anisotropic displacements in macromolecular refinement. *Acta Crystallogr. D Biol. Crystallogr.*, **57**, 122–133.
24. Ferreiro,D.U. and de Prat-Gay,G. (2003) A protein–DNA binding mechanism proceeds through multi-state or two-state parallel pathways. *J. Mol. Biol.*, **331**, 89–99.
25. McBride,A.A., Klausner,R.D. and Howley,P.M. (1992) Conserved cysteine residue in the DNA-binding domain of the bovine papillomavirus type 1 E2 protein confers redox regulation of the DNA-binding activity *in vitro*. *Proc. Natl Acad. Sci. USA*, **89**, 7531–7535.
26. Ferreiro,D.U., Dellarole,M., Nadra,A.D. and de Prat-Gay,G. (2005) Free energy contributions to direct readout of a DNA sequence. *J. Biol. Chem.*, **280**, 32480–32484.
27. Thain,A., Webster,K., Emery,D., Clarke,A.R. and Gaston,K. (1997) DNA binding and bending by the human papillomavirus type 16 E2 protein. Recognition of an extended binding site. *J. Biol. Chem.*, **272**, 8236–8242.
28. Collaborative Computational Project, Number 4. (1994) The CCP4 suite: programs for protein crystallography. *Acta Crystallogr. D Biol. Crystallogr.*, **50**, 760–763.
29. Crothers,D.M., Haran,T.E. and Nadeau,J.G. (1990) Intrinsically bent DNA. *J. Biol. Chem.*, **265**, 7093–7096.
30. Rozenberg,H., Rabinovich,D., Frolow,F., Hegde,R.S. and Shakked,Z. (1998) Structural code for DNA recognition revealed in crystal structures of papillomavirus E2–DNA targets. *Proc. Natl Acad. Sci. USA*, **95**, 15194–15199.
31. Zimmerman,J.M. and Maher,L.J. III (2003) Solution measurement of DNA curvature in papillomavirus E2 binding sites. *Nucleic Acids Res.*, **31**, 5134–5139.
32. Shinoda,T., Arai,K., Shigematsu-Iida,M., Ishikura,Y., Tanaka,S., Yamada,T., Kimber,M.S., Pai,E.F., Fushinobu,S. and Taguchi,H.

(2005) Distinct conformation-mediated functions of an active site loop in the catalytic reactions of NAD-dependent D-Lactate dehydrogenase and formate dehydrogenase. *J. Biol. Chem.*, **280**, 17068–17075.

33. Strahs,D. and Schlick,T. (2000) A-Tract bending: insights into experimental structures by computational models. *J. Mol. Biol.*, **301**, 643–663.

34. Dickerson,R.E. (1998) DNA bending: the prevalence of kinkiness and the virtues of normality. *Nucleic Acids Res.*, **26**, 1906–1926.

35. Kovelman,R., Bilter,G.K., Glezer,E., Tsou,A.Y. and Barbosa,M.S. (1996) Enhanced transcriptional activation by E2 proteins from the oncogenic human papillomaviruses. *J. Virol.*, **70**, 7549–7560.

36. Hou,S.Y., Wu,S.Y. and Chiang,C.M. (2002) Transcriptional activity among high and low risk human papillomavirus E2 proteins correlates with E2 DNA binding. *J. Biol. Chem.*, **277**, 45619–45629.

37. DeLano,W.L. (2002) The PyMOL Molecular Graphics System.