The papillomavirus E2 proteins regulate the transcription of all papillomavirus genes and are necessary for viral DNA replication. Disruption of the E2 gene is commonly associated with malignancy in cervical carcinoma, indicating that E2 has a role in regulating tumor progression. Although the E2 proteins from all characterized papillomaviruses bind specifically to the same 12-base pair DNA sequence, the cancer-associated human papillomavirus E2 proteins display a unique ability to detect DNA flexibility and intrinsic curvature. To understand the structural basis for this phenomenon, we have determined the crystal structures of the human papillomavirus-18 E2 DNA-binding domain and its complexes with high and low affinity binding sites. The E2 protein is a dimeric β-barrel and the E2-DNA interaction is accompanied by a large deformation of the DNA as it conforms to the E2 surface. DNA conformation and E2-DNA contacts are similar in both high and low affinity complexes. The differences in affinity correlate with the flexibility of the DNA sequence. Preferences of E2 proteins from different papillomavirus strains for flexible or prebound DNA targets correlate with the distribution of positive charge on their DNA interaction surfaces, suggesting a role for electrostatic forces in the recognition of DNA deformability.

Networks of hydrogen bonds between amino acid side chains and the functional groups of DNA bases are a well documented mechanism of DNA sequence recognition by proteins. Not as well understood is the stereochemical and energetic basis by which protein-DNA binding affinity is modulated when all of these direct components of interaction are conserved. Biologically, such discriminatory abilities are essential when DNA-binding proteins have to select between multiple binding sites present among a vast excess of nonspecific DNA. This is the case in the papillomaviruses, where the primary transcriptional regulatory protein, E2, is confronted by numerous binding sites on the viral genome (Fig. 1a). These sites must be occupied in a defined order such that transcription and viral DNA replication proceed in a regulated fashion (1–7).

The papillomaviruses are a large family of double-stranded DNA viruses that vary in host species specificity, tissue tropism, and the clinical outcome of infection (8, 9). The E2 protein has the same functions in the over 100 characterized papillomavirus strains: to activate or repress transcription in a context-dependent fashion and to facilitate the initiation of viral DNA replication via interactions with the viral replication protein E1. Differences in the details of transcription and replication control have been documented between the viral strains that infect different host tissues and between the papillomavirus strains that are associated with different conditions in humans such as warts or cervical cancer (2–7). The E2 proteins from all viral strains have in common the fact that they bind a palindromic DNA sequence ACCgNNNNcGGT, referred to as the E2 binding site (E2BS;1 lowercase letters indicate preferred nucleotides, and the NNNN region is called the “spacer”). However, there exist virus strain-specific differences in the abilities of various E2 proteins to discriminate between binding sites. The E2 proteins from the human papillomavirus (HPV) strains that infect mucosa (including the cancer-associated strains HPV-18 and HPV-16) and the wart-causing strain HPV-11) bind with significantly greater affinity to E2BS with spacers rich in AT base pairs (1, 10–12). On the other hand, the E2 protein from bovine papillomavirus type 1 (BPV-1) displays no distinctive spacer sequence preference (12). The viral genomes reflect these trends; the mucosal HPV genomes have E2 binding sites with AT-rich spacers (Fig. 1a), while the genomes of the non-primate animal viruses (including BPV-1) have no such predominance of AT-rich spacer-containing binding sites (13). The crystal structures of the E2 DNA-binding domain from strains BPV-1 (14), HPV-16 (15), and HPV-31 (16) have been reported, as has the co-crystal structure of BPV-1 E2 bound to DNA (17). However, there is no structural information yet on a DNA complex of any HPV E2 protein. Here we report the crystal structures of the E2 DNA-binding domain (E2/D) from human papillomavirus type 18 and its complexes with high and low affinity binding sites.

The HPV-18 E2-DNA interaction uses an induced-fit mechanism, with a large deformation of DNA and only modest local rearrangements of the protein. Comparisons of the various E2/D structures reveal an unexpected similarity between the quaternary structures of the HPV-18 and BPV-1 E2/D proteins as compared with those of the more closely related (in sequence, evolutionary distance of viral strain, and pathology) HPV-16 and HPV-31 E2/D proteins. HPV-18 E2/D binds with very different affinities to the two E2BS sequences used in the crystallographic studies, yet no base sequence-specifying protein-DNA contacts differ in the two complexes. Like the related HPV-16 E2/D protein, but in contrast to BPV-1 E2/D, HPV-18

---

1 The abbreviations used are: E2BS, E2 binding site; HPV, human papillomavirus; BPV, bovine papillomavirus; DTT, dithiothreitol.
a) Structure of the HPV-18 genome. The 7857-base pair circular genome consists of two regions coding for early (E) and late (L) genes. An 800-base pair noncoding region is present upstream of the E6 gene and is called the long control region (LCR). The major early promoter (P_105) and a differentiation-specific promoter are indicated by arrows. The long control region contains both transcriptional regulatory elements and the origin of replication. A region of the long control region (of HPV strains 218, 216, 211, and 231) that contains the four E2 binding sites (E2BS1–4) is detailed below. In each case, the spacer sequence is shown in blue. The inset is the consensus DNA-binding site of the E2 proteins and the sequences of the two binding-sites referred to as E2BS(AATT) and E2BS(ACGT). The identity elements are in red, and the preferred nucleotides are in green.

b) Amino acid sequence alignments of the HPV-18, HPV-16, and BPV-1 E2 DNA-binding domains. Identical residues are in red, and similar residues are in green. The residue numbering for HPV-18 E2 and BPV-1 E2 are shown as is the secondary structure of HPV-18 E2/D. The arrows represent β-strands, and zigzag lines indicate helices.

c) Ribbon diagram of the HPV-18 E2 DNA-binding domain. The two subunits (gold...
The HPV-18 E2 DNA-binding Domain and Its DNA Complexes

TABLE I

|                      | HPV-18E2/D | HPV-18E2D-E2BS(AATT) | HPV-18E2D-E2BS(ACGT) | BPV-1E2D/E2BS(AATT) |
|----------------------|-----------|----------------------|----------------------|----------------------|
| Resolution range (Å) | 99–1.9    | 99–2.3               | 99–3.0               | 99–2.25              |
| Unique reflections   | 25,003    | 19,322               | 6147                 | 23,731               |
| Rmerge(%)            | 7.8       | 7.9                  | 11.3                 | 8.9                  |
| Completeness (%)     | 97.5      | 88.0                 | 98.8                 | 95.                  |
| Completeness in last shell | 88.2 | 89.2                  | 99.8                 | 79.                  |
| Space group          | Pg2, 2_1  | C22_2                | C22_2                | Pg2, 2_1             |
| Unit cell            | a = 41.5  | a = 69.9             | a = 71.8             | a = 52.3             |
|                      | b = 46.8  | b = 82.2             | b = 82.6             | b = 54.8             |
|                      | c = 161   | c = 96.1             | c = 99.1             | c = 172.7            |
| Refinement           | Resolution range (Å) | 30–1.9 | 30–2.4 | 30–3.0 | 30–2.3 | 21,933 | 24,144 | 9277 | 5771 |
|                      | Reflections used for refinement | 24,144 | 9277 | 5771 | 21,933 | 24,144 | 9277 | 5771 | 21,933 |
|                      | R factor   | 23.7                 | 20.0                 | 23.7                 | 22.9                 |
|                      | Free R     | 29.4                 | 29.0                 | 30.2                 | 26.2                 |
|                      | Number of atoms | 2452 | 1847 | 1880 | 4081 |
|                      | Number of waters | 99 | 190 | 0 | 82 |
|                      | r.m.s.d. bond lengths (Å) | 0.028 | 0.008 | 0.007 | 0.006 |
|                      | r.m.s.d. angles (degrees) | 2.3 | 1.3 | 1.2 | 1.2 |

* Rmerge = ∑ (|I_k – I_A|)/ ∑ I_k, where |I_k – I_A| is the average of the absolute deviation of a reflection from the average I_k of its symmetry mates and Friedel equivalents.

** R-factor: ∑ |F_o| – |F_c| / ∑ |F_o|, where |F_o| |F_c| is the observed structure factor amplitude and |F_c| is the structure factor amplitude from the refined coordinates.

a Free R, R factor calculated on 10% of the data excluded from refinement.

b Root mean square deviation.

d Equilibrium binding constants for HPV-18 E2/D

The design of oligonucleotides used in these experiments is shown in Fig. 3. Hairpin oligonucleotides of the sequence shown in (Fig. 3a) were used to determine the affinity of HPV-18 E2/D for E2BS with the spacer sequences indicated in the first section. Dumb-bell shaped oligonucleotides (Fig. 3b) were used to introduce nicks in the phosphodiester backbone of the DNA probe.

| Binding site | K_eq | K_rel | ΔG(25% %) |
|--------------|------|-------|-----------|
| E2BS(ACGT)   | 1.8 × 10^-7 | 1.0 | -9.2 (0.4) |
| E2BS(AATT)   | 1.6 × 10^-9 | 112.5 | -12.0 (0.7) |
| E2BS(ACAT)   | 1.2 × 10^-8 | 15.0 | -10.8 (0.9) |
| E2BS(AAAT)   | 1.7 × 10^-8 | 10.6 | -10.6 (0.6) |
| E2BS(TAA)    | 3.3 × 10^-8 | 5.5 | -10.2 (0.6) |

* K_eq refers to K_{ACGT}/K_{ACAT}.

** ΔG < 65% refers to the confidence limits.

a|| refers to a break in the sugar-phosphate backbone.

d K_{rel} refers to K_{intact spacer}/K_{rel}.

** Footnotes:**

- E2/D recognizes and discriminates against conformational flexibility in DNA. There exists a correlation between the distribution of positive charge on the DNA interaction surfaces of these proteins and their preferences for prebent or flexible DNA targets.

** EXPERIMENTAL PROCEDURES**

** Protein and Oligonucleotide Preparation**

His-tagged HPV-18 E2/D (amino acids 286–365 and N-terminal residues GSHM that arose from the cloning procedure) was obtained by overexpression in Escherichia coli strain BL21(DE3)pLysS. The His-tagged protein was purified by nickel affinity chromatography. The His tag was removed by thrombin treatment, and the cleaved protein was further purified by ion exchange chromatography (Fast-S; Amersham Pharmacia Biotech), resulting in pure protein as determined by silver-stained SDS gels. BPV-1 E2/D was purified as described previously (17). All oligonucleotides were obtained from the Yale Keck Biotechnology Facility and purified by ion exchange chromatography (Mono-Q; Amersham Pharmacia Biotech) (17).

Fig. 2. Stereoview of a simulated annealing omit 2F_o-F_c map of the HPV-18 E2/D-E2BS(AATT) complex. A section of the recognition helix in the DNA major groove is shown. The map was calculated at 2.3-Å resolution and is contoured at 1.5 σ. The image was generated by the Swiss PDB viewer.

Fig. 3. Design of oligonucleotides used.

** Crystallizations**

All of the crystals were grown by vapor diffusion from hanging drops at room temperature. HPV-18 E2/D—5 mg ml^-1 of HPV-18 E2/D in 25 mM Tris, pH 7.5, 100 mM NaCl, 10 mM DTT was mixed with an equal volume of a and lavender) associate to form an eight-stranded antiparallel β-barrel. Disordered regions are represented by dotted lines. The topology is shown beside them. The β-barrel is formed by curling the figure toward the viewer such that the β-strands at the edges can form hydrogen bonds with each other. The ribbon diagram was generated by RIBBONS (42).
reservoir solution containing 2.7 mM ammonium sulfate and 0.1 mM sodium acetate, pH 4.7. Crystals of dimensions 0.25 × 0.25 × 0.15 mm appeared in 3–5 days. Diffraction data were recorded on a RAXIS II image plate detector from a crystal maintained at −170 °C.

HPV-18 E2/D-E2BS(AATT)—An equimolar mixture of HPV-18 E2/D and E2BS(AATT) was made in 25 mM Tris-HCl, pH 7.5, 100 mM NaCl, 10 mM DTT. Drops were made by mixing 10 mg ml⁻¹ complex with an equal volume of a reservoir solution containing 47% 2-methyl-2,4-pentane diol, 100 mM sodium acetate, pH 4.7, 20 mM CaCl₂. Plate-like crystals of dimensions 0.4 × 0.25 × 0.05 mm appeared in 5–7 days. Diffraction data were recorded on a RAXIS II image plate detector from a crystal maintained at −170 °C.

HPV-18 E2/D-E2BS(ACGT)—An equimolar mixture of HPV-18 E2/D and E2BS(ACGT) was made in 25 mM Tris-HCl, pH 7.5, 100 mM NaCl, 10 mM DTT. Drops containing equal volumes of the protein-DNA complex and a well solution containing 48% 2-methyl-2,4-pentane diol, 100 mM sodium acetate, pH 4.7, 20 mM CaCl₂ were made. Very thin plate-like crystals (0.1 × 0.1 × 0.02 mm) appeared in 5–7 days. Diffraction data were recorded at Brookhaven beam X4A on a CCD detector from a crystal maintained at −170 °C.

BPV-1 E2/D-E2BS(AATT)—An equimolar mixture of BPV-1 E2/D and an annealed oligonucleotide of sequence CCAACCCAATTCCGGTTG was made in 25 mM Tris, pH 7.5, 100 mM NaCl, 10 mM DTT. Drops were made with 6.25 mg ml⁻¹ of the complex and an equal volume of a reservoir solution containing 30% PEG3550, 47 mM MES, pH 6.2, 2 mM CaCl₂. Rod-shaped crystals of dimensions 0.15 × 0.1 × 0.08 mm appeared in 5–7 days. Diffraction data were recorded on a RAXIS II image plate detector from a crystal maintained at −170 °C.

Structure Solution and Refinement

All images were indexed; the reflections were integrated, scaled, and postrefined with the HKL package (programs DENZO and SCALEPACK (18)); and the structures were determined by molecular replacement using the program AMORE (19). A model of BPV-1 E2/D-E2BS(ACGT) (17) was used to find a Molecular Replacement solution for the BPV-1 E2/D-E2BS(AATT) structure. Using data from 15–3.5 Å, the correct solution had a correlation coefficient of 0.25.

A polyalanine model constructed from the refined structure of the BPV-1 E2/D-E2BS(AATT) complex was used in Molecular Replacement to determine the structure of the HPV-18 E2/D-E2BS(AATT) complex. Using data from 15–4.8 Å, the correct solution had a correlation coefficient of 0.491.

A model consisting of one subunit of HPV-18 E2/D was constructed from the HPV-18 E2/D-E2BS(AATT) complex structure and used to obtain a Molecular Replacement solution for the HPV-18 E2/D structure. Using data between 15 and 4.5 Å, the correct solution had a correlation coefficient of 0.507.

Using the refined structure of the E2BS(AATT) complex as a model, rigid body refinement was conducted on the HPV-18 E2/D-E2BS(ACGT) complex data. Using difference maps, the correct oligonucleotide sequence was modeled.

All structures (except the HPV-18 E2/D-E2BS(ACGT) complex) were refined using a combination of simulated annealing, bulk solvent correction, positional refinement, and restrained individual B-factor refinement in CNS (20). Simulated annealing omit maps were systematically calculated and examined to minimize the effects of model bias. Only restrained grouped B-factors were refined for the HPV-18 E2/D-E2BS(ACGT) structure.

The final model of HPV-18 E2/D-E2BS(AATT) includes residues 287–323 and 329–364 (residues 309–312 had weak electron density, and some of these side chains were modeled as Ala). No unambiguous density corresponding to the C-terminal residue (Met365) was present, and it is not included in the refined model.

The final model of HPV-18 E2/D-E2BS(ACGT) includes residues 287–323 and 328–364 (residues 310–312 had weak electron density, and some of these side chains were modeled as Ala). No unambiguous density corresponding to the C-terminal residue (Met365) was present, and it is not included in the refined model.

410. There are two complexes in the asymmetric unit. In one complex, one 5'-overhanging C is flipped out packing against the protein surface, and no density corresponding to the other 5'-C is present. In the other complex, the C nucleotide at position 8 of one strand is flipped out and packed against the protein surface. The 5'-C then base pairs with G⁻¹ of the other strand. Since the DNA conformation at the 5'- and 3'-ends are clearly affected by crystal packing, all analyses of DNA conformation in the discussions are restricted to the central 14 base pairs that include the E2BS.

For all structures, the final model had all non-Gly residues in allowed regions of the Ramachandran plot. The data and refinement statistics are summarized in Table I. A section of a composite simulated annealing omit map calculated on the HPV-18 E2/D-E2BS(AATT) co-crystal structure is shown in Fig. 2.

DNA Binding Studies

The oligonucleotides were engineered so that the termini of the DNA were either at one end or within the spacer region (Table II and Fig. 3) of the E2 binding site. This strategy allowed the introduction of breaks in the phosphodiester backbone of the E2 binding site. Oligonucleotides were purified by acrylamide gel electrophoresis, end-labeled, and annealed by heating the probe (2 nM concentration) to 90 °C for 10 min and transferring immediately to ice for 10 min. Quantitative gel mobility shift experiments were conducted following published protocols (21, 22). Briefly, serial dilutions of the proteins were added to binding buffer resulting in a final mixture containing 2 nM γ⁻³²P-labeled oligonucleotide, 150 µg/ml bovine serum albumin, 5 µg/ml sonicated salmon sperm DNA, 5.5 mM DTT, 22 mM HEPES, pH 7.9, 150 mM KCl, 5 mM MgCl₂, and 10% (v/v) glycerol. The reaction mixtures were incubated in a water bath at 25 °C for 30 min and then loaded onto prerun 12% polyacrylamide gels in 0.5× TBE, pH 8.0. The electrophoresis was run at 200 V for ~90 min. The gels were dried and the reaction products visualized by exposure to phosphor storage plates, which were scanned using a PhosphorImager (Molecular Dynamics, Inc., Sunnyvale, CA).

The density of the electrophoretic band representing the protein-DNA complex was quantitated using the ImageQuant software. Binding isotherms were obtained by monitoring the density of the electro-
The equilibrium binding constant, $K$, was determined by analysis of the titration curves against the coupled equations

$$\theta_1 = \frac{1}{1 + K[X]}$$

and

$$\theta_2 = \frac{K[X]}{1 + K[X]}$$

where $\theta_1$ is the DNA fraction unbound, $\theta_2$ is the DNA fraction complexed with protein, $K$ is the equilibrium association constant, and $[X]$ is the free active protein dimer concentration. The DNA binding activity of the HPV-18 E2/D protein preparations was determined from stoichiometric titrations (22); the values presented are corrected for this activity. Since the $^{32}$P-labeled oligonucleotide concentration is much lower than the equilibrium dissociation constants being measured, the approximation that total protein concentration is equal to free protein concentration is made. The standard state Gibbs free energy of binding was calculated from the equilibrium association constant by $\Delta G^0 = -RT \ln K$, where $R$ is the gas constant and $T$ is temperature. Each of the values of $\Delta G^0$ reported were determined by the global analysis of at least two independent titrations.

RESULTS

Overview of the Structures—The 1.9-Å crystal structure of HPV-18 E2/D and the 2.4-Å structure of its complex with the high affinity binding site E2BS(AATT) are described in detail below. In order to examine the stereochemistry at a low affinity HPV-18 E2/D-DNA interface, the structure of HPV-18 E2/D bound to E2BS(ACGT) has been determined. While the limited resolution (3 Å) of this crystal structure does not permit detailed analyses, it is of sufficient quality to allow comparisons of global features and direct protein-DNA interactions. To provide a direct comparison between the modes of DNA recognition of the HPV-18 and BPV-1 E2/D proteins, the 2.3-Å structure of BPV-1 E2/D bound to E2BS(AATT) has also been determined. This complex is similar in most respects to the previously reported crystal structure of BPV-1 E2/D bound to E2BS(ACGT) (17), and only salient features are presented here. In all of the discussions below, comparisons are drawn...
between the complexes of the HPV-18 or BPV-1 E2/D proteins with the same E2BS sequence, unless explicitly stated otherwise.

The HPV-18 E2 DNA-binding Domain—Each monomer of HPV-18 E2/D folds into an open-faced β-sandwich with a β-α-β-β-α-β topology (Fig. 1c). One α-helix in each subunit contains all of the amino acid residues involved in half-site recognition and is termed the “recognition helix.” Two E2/D monomers associate such that the β-strands at the edges (β4 and β5) form hydrogen bonds with their symmetry mates (Fig. 1c). This results in a continuous eight-stranded antiparallel β-sheet. Buried in the interface are numerous large side chains including Thr287, Ile289, His291, Lys293, Trp296, Trp320, Trp322, Ile332, Thr334, Thr336, Met363, and Met368 (Fig. 4a). The dimerization interface is extensive, occluding 1625 Å² of surface area. A herring bone type packing of four Trp residues in the dimer interface is common to all of the mucosal HPV E2 proteins (Fig. 4) and is likely to be a major contributor to the stability of the dimer. As a result of the intricate and closely packed dimer interface, any rearrangement of subunits upon DNA binding would entail a significant energetic penalty. There is a cavity at the center of the barrel, and solvent molecules have been modeled into the electron density clearly apparent in this region. Electron density corresponding to the loop between the recognition helix and β4 (residues 308–312) is weak, indicating local disorder in this region. The loop between strands β4 and β5 is disordered.

Comparison of the Structures of E2 Proteins from Different Viral Strains—The E2 proteins from different viral strains differ in quaternary structure (Fig. 5a). The HPV-18 and BPV-1 E2/D proteins are alike in the relative orientation of their subunits while differing from the HPV-16 and HPV-31 E2/D proteins. When one subunit of each protein is superimposed, the nonsuperimposed recognition helices of HPV-16 and HPV-18 E2/D are related by an average translation of 7 Å (Fig. 5a). Functionally, these features of the E2 proteins are critical, since they dictate the spatial arrangement of side chains presented to the major grooves for DNA sequence recognition. The large variation in subunit orientation among the HPV E2/D proteins suggests either that they undergo unique subunit rearrangements upon DNA binding or that DNA is bent very differently in each of the protein-DNA complexes.

E2 dimer architecture is governed by two features: first, the alignment of strands β2 and β4 against their symmetry mates (Fig. 5b) and, second, the packing of side chains in the barrel core (Fig. 4). While the alignment of the β2 strands is invariant, E2 proteins from different viral strains differ in the register of their dyad-related β4 strands (Fig. 5b). When one subunit each of the HPV-18 E2/D and HPV-16 E2/D proteins are superimposed, residues Gly361-Thr364 of HPV-18 E2/D align with residues Gly361-Ser364 of HPV-16 E2/D. The nonsuperimposed subunit is out of register by two residues; Tyr362-Met365 of HPV-18 E2/D lines up with Thr366-Met368 of HPV-16 E2/D. As a result, while the Phe362 residues of HPV-16 E2/D straddle a 2-fold axis, the corresponding Tyr362 residues in HPV-18 E2/D are not symmetrically disposed about a 2-fold axis. In this regard, HPV-18 E2/D resembles BPV-1 E2/D, while the HPV-16 and HPV-31 E2/D proteins are similar to each other (23). It has been proposed that strand register is specified by cross-strand side chain interactions (24), and this is clearly evident in the case of these E2/D proteins. The invariant glycine in the β4 strand of all E2 proteins (Gly363 in HPV-18 E2/D and HPV-16 E2/D and Gly403 in BPV-1 E2/D) has the potential to destabilize the β-sheet. In HPV-18 E2/D, the β-sheet is disrupted at Gly361, with the side chain of Thr364 participating in a cross-strand hydrogen bond while the β-strands diverge. In BPV-1 E2/D, the side chain of Phe410 participates in a cross-strand rescue of the Gly403-mediated β-sheet destabilization. Such pairing of glycine and aromatic residues across β-strands provides an increase in protein stability (25). The aromatic side chains of Tyr362 in HPV-18 and Phe362 in HPV-16 E2/D do not participate in analogous cross-strand stabilization. Thus, the surfaces formed by the β4 strands of different E2/D proteins are not alike. Interestingly, this surface of BPV-1 E2 has been implicated in an interaction with the E1 helicase that is essential for viral DNA replication (26). The differences in structure described above could contribute to the inability of heterolo-
Local DNA parameters

| Complex                  | Bend angle | Minor groove width |
|--------------------------|------------|--------------------|
|                           | degrees    | Å                  |
| HPV-18 E2-E2BS(AATT)     | 43         | 8.5                |
| HPV-18 E2-E2BS(ACGT)     | 49         | 8.5                |
| BPV-1 E2-E2BS(AATT)      | 47         | 9.8                |
| BPV-1 E2-E2BS(ACGT)      | 51         | 9.8                |

The HPV-18 E2 DNA-binding Domain and Its DNA Complexes

The HPV-18 E2-DNA Complex—DNA surrounds one face of the HPV-18 E2/D-β-barrel engulfing both recognition helices in successive major grooves (Fig. 6c). The global structure of the HPV-18 E2/D protein is similar in its free and DNA-bound states with a root mean square deviation of 1.1 Å for the peptide backbone of the dimer. There is a very small adjustment of recognition helix orientation that facilitates symmetric placement of the helices in contact with the identity elements of the E2BS. On average, the temperature factors of recognition helix backbone atoms are lowered by 11 Å² upon DNA interaction, indicating that the helices assume more rigid conformations as a network of protein-DNA interactions is formed.

HPV-18 E2/D-DNA binding is not accompanied by any large disorder-to-order transition. In particular, the disordered loop connecting β6 and β7 is not conformationally restricted upon DNA binding. In contrast, DNA binding significantly rigidifies the corresponding segments of the BPV-1 E2/D protein; six intersubunit hydrogen bonds and two tethering contacts with DNA phosphate oxygens contribute to this stabilization (Fig. 6c).

DNA Conformation in the E2-DNA Complexes—DNA wraps around the outside surface of the E2 β-barrel, tracing a smooth curve. The overall bend angle was measured between normals to the base pairs at positions +6 and −6 from the dyad axis, thus encompassing the consensus E2BS and eliminating any effects of end-to-end lattice packing (Table III). The bend angle is very similar for the HPV-18 E2/D and BPV-1 E2/D complexes with both E2BS sequences studied. The DNA is the B form with compressed major and minor grooves facing the protein. Three local parameters, roll, helix twist, and slide, that are the most sensitive to DNA deformation were analyzed and are listed in Table III. The roll angles in the contacted ACC/GGT half-sites are positive, thus compressing the major groove, and the roll angles in the central NNNN region are negative, compressing the minor groove. No kink at any single base step is observed. The DNA axis traces an out-of-plane bend. In all cases, the central four base pairs are overtwisted, while the ACCG/GGT half-sites are underwound.

The complexes formed by the HPV-18 and BPV-1 E2/D proteins differ in two related respects: the central minor groove width and the roll angle at the central dinucleotide step. The central minor groove is narrower in the HPV-18 E2/D complex (8.5 Å) than it is in the BPV-1 E2/D complexes (9.8 Å), regardless of spacer sequence. A large negative roll angle of −11.4° at the central dinucleotide step of the HPV-18 E2/D-E2BS(AATT) complex contributes to this narrowing of the minor groove. No classical spine of hydration is apparent in the minor groove of any E2-DNA complex. The complex between BPV-1 E2/D and E2BS(ACGT) has a more positive slide in the central minor groove, while the ACCG/GGT half-sites are underwound.

In summary, both E2 proteins induce a similar global deformation of DNA characterized by compressed major and minor grooves at the interaction surface. The trends in roll angle and groove width are characteristic of the particular E2 protein present in the complex. Since both HPV-18 E2/D and BPV-1 E2/D must induce the same overall deformation in DNA, any differences in their preferences for intrinsic DNA conformational propensities must reflect differences in their innate abilities to stabilize the appropriate distortion.
Fig. 7. a and c, view of the recognition helices of HPV-18 E2/D (a) and BPV-1 E2/D (c) as positioned in the major groove of DNA. Side chains of amino acids involved in direct contacts with the bases are shown. Hydrogen bonds are indicated by dashed lines. b and d, schematic drawings of the protein-DNA contacts observed in the HPV-18 E2/D complex with E2BS(AATT) (b) and BPV-1 E2/D complex with E2BS(AATT) (d). Base pairs in the absolutely conserved identity elements are depicted by filled rectangles, and preferred base pairs are shown as gray rectangles. Phosphates contacted by the proteins are shown as gray-filled circles. Circles labeled W represent solvent molecules. The hatched line represents a hydrophobic van der Waals contact. All of the interactions are dyad symmetric; however, in these images base contacts are shown in the bottom half and phosphate contacts in the top half.
DNA sequence. Additionally, Arg305 contacts the adenine base at position -7 outside the minimal E2BS. This interaction may contribute to the 16-fold greater affinity with which HPV-18 E2/D binds to sites that have an AT base pair at this location (e.g., BS2 versus BS1 in the HPV-18 genome (1) (Fig. 1a)). Mutagenesis data show that an Arg305 to Ala mutation abrogates this specificity (28). HPV-18 E2/D also contacts the phosphodiester backbone of the DNA in both the major and the minor grooves. In total, 14 direct and 20 water-mediated phosphate contacts have been identified. Phosphate contacts made by Thr319, Arg303, and Ser298 at positions 3, 4, and 7 of the DNA have exact counterparts in the BPV-1 E2-DNA complex via residues Thr359, Arg342, and Gln337. Thus, all major groove phosphate contacts are conserved between these two complexes. The minor groove phosphate contact made by Arg370 is unique to the HPV-18 E2/D-DNA complex. Notably absent from the HPV-18 E2/D-DNA complexes is any contact between phosphates and amino acids in the β/β loop, such that mediated by the Arg370 side chain of BPV-1 E2/D.

All of the direct protein-DNA contacts described above are present in the low affinity complex between HPV-18 E2/D and E2BS(ACGT). Although solvent molecules were not modeled, clear electron density is present at positions in the interface where solvent-mediated protein-DNA contacts are seen in the high affinity E2BS(AATT) complex.

**Noncontacted Spacer Sequence Preferences of HPV-18 E2/D**—The affinity of HPV-18 E2/D toward a representative set of E2BS that differed in spacer sequence was measured. These included the spacer sequences found on the HPV-18 genome and those used in the crystallographic analyses reported here. The results are presented in Table II. HPV-18 E2/D binds with greater affinity to E2BS containing AT-rich spacer sequences. The relationship between binding affinity and spacer sequence can be summarized as AATT > AAAT > AAAA > TTAA > AATT > AATT. Similar preferences for AT-rich spacer sequences were reported for HPV-16 E2/D and HPV-11 E2/D (10, 12). The co-crystal structures reported here eliminate the possibility that direct contact with the bases is responsible for the ability of HPV-18 E2/D to discriminate between E2BS based upon their spacer sequence.

A tract-containing binding sites would be predisposed to adopt the narrowed central minor groove conformation found in the E2-DNA complex (29), thus lowering the enthalpic cost of DNA deformation upon E2 binding. Conversely, E2BS that are not conformationally complementary to the protein surface would bind to E2 with lower affinity, reflecting the distortion free energy of the DNA as it assumes the conformation found in the complexed state. This is clearly evident in the nearly 2 kcal/mol greater affinity of HPV-18 E2/D for E2BS(AATT) than for E2BS(TTAA). While both of these binding sites have AT-rich spacer sequences, the TpA step in E2BS(TTAA) rolls into the major groove, resulting in central minor groove widening (30), a conformation not complementary to the E2 DNA-binding surface. A similar widening of the central minor groove was seen in the crystal structure of free E2BS(ACGT) (31), toward which HPV-18 E2/D also displays low affinity. The AT-rich spacer sequence preference of the HPV E2 proteins thus correlates with the conformational complementarity of such intrinsically curved DNA to the E2 surface.

Although BPV-1 E2/D also distorts DNA in a globally analogous fashion, it does not show any significant preference for AT-rich spacer-containing E2BS (12). Two differences between the HPV-18 and BPV-1 E2/D-DNA complexes may contribute to this effect. First, the central minor groove is not as narrow in BPV-1 E2/D-DNA complexes as it is in the HPV-18 E2/D-DNA complexes (Table III). Thus, a predisposition toward narrow minor grooves may be less important for the BPV-1 E2/D-DNA interaction. Second, Arg370 in the β/β loop of BPV-1 E2 makes a critical phosphate interaction in the central minor groove of the DNA. This contact could significantly attenuate repulsion between phosphate groups on the two strands caused by minor groove narrowing. The spacer sequence preferences of HPV-18 E2/D may reflect its inability to compensate for this electrostatic repulsion between the strands of a narrowed minor groove.

In order to explore the relationship between DNA flexibility and HPV-18 E2/D binding, we measured the affinity of HPV-18 E2/D toward representative nicked E2BS sequences (Table II). HPV-18 E2/D binds with significantly lower affinity to a nicked E2BS(AATT) target than it does to the corresponding intact oligonucleotide. In contrast, upon introducing a nick in a DNA target that is not conformationally complementary to the E2 DNA-binding surface (E2BS(TTAA)), the affinity of the interaction is slightly enhanced. It appears that HPV-18 E2/D discriminates against flexibility in the E2BS. In a more exhaustive analysis, it was shown that HPV-16 E2/D-DNA binding affinity was also drastically attenuated when the flexibility of the E2BS was increased via nicks in the phosphodiester backbone (12). Further support for the conclusion that these HPV E2 proteins discriminate against flexibility in their DNA targets comes from an analysis of DNA deformability (32), which suggests that A runs (of 3 or more A nucleotides) are more deformable than AA dimers. This correlates with the greater affinity exhibited by HPV-18 E2/D for E2BS(AATT) than for E2BS(AAAT) or E2BS(ACGT). DNA flexibility can reduce the energy expended in deforming DNA upon ligand or protein binding, but it also increases the entropic cost in going from a flexible, free DNA to a relatively rigid complexed state. The fact that the HPV E2/D-DNA interaction is adversely affected by flexibility in DNA indicates that the entropic cost is not balanced by favorable protein-DNA interactions. Interestingly, the BPV-1 E2 protein does not disfavor flexibility in the E2BS (12), suggesting that the DNA-bending mechanisms of these E2 proteins may be dissimilar. Furthermore, in the case of the E. coli phage 434 repressor-operator interactions (33, 34), where a preference for AT-rich spacer sequences was also observed, the repressor-DNA interaction was favored by DNA flexibility.

Thus, no general relationship between DNA flexibility and protein binding affinity has yet been derived.

There is a substantial difference in the distribution of positive potential on the DNA contact surfaces of HPV-18 E2/D and

**Fig. 8. The potential energy surfaces presented to DNA by HPV-18 E2/D and BPV-1 E2/D.** The electrostatic surface potential of the HPV-18 E2 DNA-binding domain (a) and the BPV-1 E2 DNA-binding domain (b) was calculated by GRASP (44). Negative potentials are red; positive potentials are blue. Water probe radius was 1.4 Å; potential is displayed on a −15 kBT to +15 kBT scale. The view is that shown in c. The disordered β/β loop is indicated by a dotted line, and the recognition helices are in blue.
BPV-1 E2/D (Fig. 8). A large accumulation of positive charge is present in the central region of HPV-18 E2/D, which is juxtaposed against the minor groove of the E2BS (Fig. 8a). This is primarily due to the basic residues Arg296, Arg303, and Arg307. According to finite difference Poisson-Boltzmann analyses, the DNA minor groove potentials are more negative for ApT sequences than they are for GpC sequences. This is because the electrostatic potentials on T(O2) and A(N3) atoms are more negative than the corresponding values for (O2) and (N3) (35). E2BS with AT-rich spacer sequences thus have minor grooves that are electrostatically complementary to the HPV-18 E2/D surface. Interestingly, HPV-16 E2/D does not have a similar accumulation of positive charge at the center of its DNA-binding surface (15), yet it too prefers AT-rich spacer-containing binding sites (11, 12). Thus the spacer sequence discrimination ability of HPV-16 E2/D may rely less on simple electrostatic complementarity than in the case of HPV-18 E2/D. In contrast to both of these HPV E2/D proteins, positive charge on BPV-1 E2/D is concentrated on surfaces flanking the E2BS (Fig. 8b), correlating with its inability to favor binding sites with AT-rich spacer sequences. Thermodynamic and structural studies are in progress to critically assess the relationships between electrostatics, DNA flexibility, and the mechanism of E2/D-DNA binding.

DISCUSSION

Viral DNA replication and gene expression patterns reflect the relative occupancy of E2 binding sites as the concentration of the E2 protein varies (1, 36, 37). The mucosal HPVAs have a conserved arrangement of four E2BS on their genomes (Fig. 1a). E2 binding at BS4 can specifically up-regulate viral early gene expression, including the expression of oncoproteins E6 and E7. E2 binding at the promoter-proximal sites BS1 and BS2 leads to repression of transcription of the early genes, including the oncoproteins E6 and E7. E2BS3 is important for viral DNA replication.

Papillomaviruses replicate in terminally differentiated cells that have exited the cell cycle. Since viral DNA replication is dependent on host cell factors, they need to activate the transcription of the E6/E7 genes. The expressed E6 and E7 proteins are uniquely tailored for optimal regulation of gene expression, including the expression of oncogenes E6 and E7. E2/D interacts with an average of two proteins (1, 36, 37). Modulation of this intrinsic affinity by cooperative binding has been suggested as one of the mechanisms by which HPV-18 E2/D-DNA binding.

Acknowledgments—We thank Dr. Zippora Shakked (Weizmann Institute, Israel) and Dr. Michael Brenowitz (Albert Einstein College of Medicine, NY) for helpful discussions and for critical reading of the manuscript.

REFERENCES

1. Steger, G., and Corbach, S. (1997) J. Virol. 71, 50–58
2. Garrido-Guerrero, E., Carrillo, E., Guido, M., Zamorano, R., Garcia-Carranca, A., and Gariglio, P. (1996) Arch. Med. Res. 27, 389–394
3. Cheng, S., Schmidt-Grimminger, D.-C., Murant, T., Broker, T., and Chow, T. L. (1995) Genes Dev. 9, 2335–2349
4. Rapp, B., Paweleck, A., Kraetzter, F., Schaefer, M., May, C., Purdie, K., Grassmann, K., and Ifner, T. (1997) J. Virol. 71, 6956–6966
5. Grassmann, K., Rapp, B., Maschek, H., Petry, K. U., and Ifner, T. (1996) J. Virol. 70, 2339–2349
6. Di Lorenzo, T. P., Chen, D., Zhang, P., and Steinberg, B. M. (1998) Virology 243, 130–139
7. Di Lorenzo, T. P., and Steinberg, B. M. (1998) J. Virol. 69, 6865–6872
8. zur Hausen, H., and de Villiers, E. M. (1994) Annu. Rev. Microbiol. 48, 427–447
9. zur Hausen, H. (1996) Biochem. Biophys. Acta 1288, F55–F78
10. Alexander, K. A., and Phelps, W. C. (1996) Biochemistry 35, 9864–9872
11. Bednarski, C. L., and Bastia, D. (1990) Virology 174, 575–577
12. Hines, C. S., Meghoo, C., Shetty, S., Bubimar, M., Brenowitz, M., and Hegde, R. S. (1998) J. Mol. Biol. 276, 809–818
13. Li, R., Knight, J., Bream, G., Stenlund, A., and Botchan, M. (1989) Dev. Genes 11, 510–526
14. Hegde, R. S., Wang, A.-F., Kim, S.-S., and Schapira, M. (1998) J. Mol. Biol. 276, 797–808
15. Hegde, R. S., and Androphy, E. (J. 1988) J. Mol. Biol. 284, 1479–1489
16. Busiero, D., Kong, X., Edan, D. A., Walter, K., Holzman, T. F., Lindh, F., Robins, T., and Giranda, V. L. (1998) Acta Crystallogr. D Biol. Crystallogr. 54, 1367–1376
17. Hegde, R. S., Grossman, S. R., Laimins, L. A., and Sigler, P. B. (1992) Nature 359, 505–512
18. Otwinowski, Z. (1983) in Data Collection and Processing (Sawyer, L., Isaacs, N., and Bailey, S., eds) pp. 56–62, SERC Daresbury Laboratory, Warrington, UK
19. Navaza, J. (1994) Acta Crystallogr. A 50, 157–163
20. Brunger, A. (1992) CNS version 0.9, Yale University Press, New Haven, CT
21. Senear, D. F., and Brenowitz, M. (1991) J. Biol. Chem. 266, 13661–13671
22. Senear, D. F., Dalma-Weishaus, D., and Brenowitz, M. (1993) Electrophoresis 14, 704–712
23. Liang, H., Petrov, A. M., Meadows, R. P., Yoon, H. S., Edan, D. A., Walter, K., Holzman, T. F., Robins, T., and Freis, B. S. (1996) Biochemistry 35, 2095–2105
24. Merkel, J. S., Surtetvart, J. M., and Regan, L. (1999) Struct. Fold Des. 17, 1333–1343
25. Merkel, J. S., and Regan, L. (1998) Fold Des. 3, 449–455
26. Chen, G., and Stenlund, A. (2000) J. Virol. 74, 1506–1512
27. Berg, M., and Stenlund, A. (1997) J. Virol. 71, 3853–3863
28. Thain, A., Webster, K., Emery, D., Clarke, A. R., and Gost, K. (1997) J. Biol. Chem. 272, 8236–8242
29. Brothers, D. M., Hanen, T. E., and Nadeau, J. G. (1990) J. Biol. Chem. 265, 7093–7096
30. Bostock-Smith, C. E., Laughton, C. A., and Searle, M. S. (1999) Biochem. J. 342, 125–132
31. Rozenberg, H., Rabinovich, D., Frow, F., Hegde, R. S., and Shakked, Z. (1998) Proc. Natl. Acad. Sci. U. S. A. 95, 15194–15199
32. Olson, W. K., Gorin, A. A., La, X. J., Heck, L. M., and Zhurkin, V. B. (1998) Proc. Natl. Acad. Sci. U. S. A. 95, 11163–11168
33. Koudelka, G. B., Harbury, P., Harrison, S. C., and Patschne, M. (1998) Proc. Natl. Acad. Sci. U. S. A. 95, 4633–4637
34. Koudelka, G. B., and Carlson, P. (1992) Nature 355, 89–91
35. Dungo, T. H., and Zabrowska, K. (1998) J. Mol. Biol. 280, 31–39
36. Stubenrauch, F., and Pfister, H. (1994) J. Virol. 68, 6959–6966
37. Stubenrauch, F., Colbert, A. M., and Laimins, L. A. (1998) J. Virol. 72, 8115–23
38. Vossen, K. (1993) FASEB J. 7, 872–879
39. Schneider-Gadicke, A., and Schwartz, E. (1986) EMBO J. 5, 2285–2292
40. Schneider-Gadicke, A., and Schwartz, E. (1987) Hamatol. Bluttransfus. 31, 389–381
41. Baker, C. C., Phelps, W. C., Lindgren, V., Braun, M. J., Gonda, M. A., and Howley, P. M. (1987) J. Virol. 61, 962–971
42. Carson, M. (1987) J. Mol. Graphics 5, 103–106
43. Abasyan, R. A., Toet, M. M., and Kuznetsov, D. N. (1994) J. Comp. Chem. 15, 488–506
44. Nicholls, A., Smith, A., and Horne, B. (1991) Proteins Struct. Funct. Genet. 11, 281–296
45. Dicks, R. E. (1998) Nucleic Acids Res. 26, 1906–1926