Distinctive Cognate Sequence Discrimination, Bound DNA Conformation, and Binding Modes in the E2 C-Terminal Domains from Prototype Human and Bovine Papillomaviruses†

Diego U. Ferreiro,‡ L. Mauricio T. R. Lima,§ Alejandro D. Nadra,‡ Leonardo G. Alonso,‡ Fernando A. Goldbaum,‡ and Gonzalo de Prat-Gay*,‡

Instituto de Investigaciones Bioquímicas Fundación Campomar and Facultad de Ciencias Exactas y Naturales, Universidad de Buenos Aires, Patricias Argentinas 435, (1405) Buenos Aires, Argentina, and Faculdade de Farmacia, Universidade Federal do Rio de Janeiro, Cidade Universitaria, Rio de Janeiro 21941-590, Brazil

Received July 20, 2000; Revised Manuscript Received September 6, 2000

ABSTRACT: The C-terminal DNA binding domain of the E2 protein is involved in transcriptional regulation and DNA replication in papillomaviruses. At low ionic strength, the domain has a tendency to form aggregates, a process readily reversible by the addition of salt. While fluorescence anisotropy measurements show a 1:1 stoichiometry at pH 5.5, we observed that a second HPV-16 E2 C-terminal dimer can bind per DNA site at pH 7.0. This was confirmed by displacement of bis-ANS binding, tryptophan fluorescence, native electrophoresis, and circular dichroism. The two binding events are nonequivalent, with a high-affinity binding involving one E2C dimer per DNA molecule with a K_D of 0.18 ± 0.02 nM and a lower affinity binding mode of 2.0 ± 0.2 nM. The bovine (BPV-1) E2 C-terminal domain binds to an HPV-16 E2 site with 350-fold lower affinity than the human cognate domain and binds 7-fold less tightly even to a bovine-derived DNA site. The ability to discriminate between cognate and noncognate sequences is 50-fold higher for the human domain, and the latter is 180-fold better than the bovine at discriminating specific from nonspecific DNA. A substantial conformational change in bound DNA is observed by near-UV circular dichroism. The bovine domain imposes a different DNA conformation than that caused by the human counterpart, which could be explained by a more pronounced bent. Structure−function differences and biochemical properties of the complexes depend on the protein domain rather than on the DNA, in line with crystallographic evidence. Despite the strong sequence homology and overall folding topology, the differences observed may explain the distinctive transcriptional regulation in bovine and human viruses.

Transcriptional regulation of gene expression is mediated by DNA binding proteins, which mostly share common features such as oligomeric structures and palindromic binding sites. However, there is a wide number of different topologies, each with distinctive sequence specificities (*1*). DNA replication in all organisms begins with the specific recognition of the DNA replication origin by proteins known as origin-binding proteins (OBPs)[1] (*2*). The E2 protein from papillomaviruses is a transcriptional regulator of all viral transcripts and plays an auxiliary role in viral DNA replication (*3*, *4*), and this dual function makes it a distinctive class within OBPs (*2*).

E2 proteins from bovine papillomaviruses can be either negative or positive regulators of transcription (*3*, *5*). The full-length protein acts as a transcriptional activator (E2-TA) and comprises two conserved domains, N-terminal transactivation and C-terminal DNA-binding and dimerization domain, linked by a flexible region without any sequence homology among the different strains (*6*, *7*). Two shorter forms found in the bovine virus (BPV-1) are the product of alternative splicing and act as transcriptional repressors, both having the C-terminal DNA-binding domain (E2C) intact. E2 proteins have also been involved in other key viral functions such as plasmid copy number control and viral genome segregation (*8*). Recently, HPV-16 E2-TA was shown to repress the expression E6/E7 oncogenes, with cancer cell growth inhibition as a result (*9*). All these functions are dependent on its DNA binding properties, hence our interest in a detailed structure−function analysis of this particular interaction.

The first structure of the BPV-1 E2C bound to DNA uncovered a unique folding topology: the dimeric β-barrel (*10*). This unique fold consists of an 8-stranded β-barrel, with each monomer contributing half of the barrel. Two α-helices pack against outside of the barrel in each monomer, where the larger of these is involved in DNA recognition (Figure

* To whom correspondence should be adressed. Telephone: +54-11-48634011. Fax: +54-11-48652246. E-mail: pratgay@iib.uba.ar.

† This work was supported by a Wellcome Trust Collaborative Research Initiative Grant (OIA U41 RG27994). D.U.F. holds a Fellowship from CONICET, Argentina. G.P.G. is a John Simon Guggenheim Fellow and wishes to thank the support of Fundación Bunge & Born, and ANPCyT, Argentina.

§ Universidade Federal do Rio de Janeiro.

‡ Universidad de Buenos Aires.

[1] Abbreviations: OBP, origin binding proteins; HPV, human papillomavirus; BPV, bovine papillomavirus; hE2C, HPV-16 E2 DNA binding domain; bE2C, BPV-1 E2 DNA binding domain; Fl-site35, fluoresceinated site 35 DNA.; bis-ANS, bis-8-anilino-1-naphthalenesulfonate; Bis-Tris, 2-[bis(2-hydroxyethyl)amino]-2-(hydroxymethyl)-1,3-propanediol; CD, circular dichroism.

10.1021/bi001694r CCC: $19.00 © 2000 American Chemical Society
Published on Web 11/02/2000

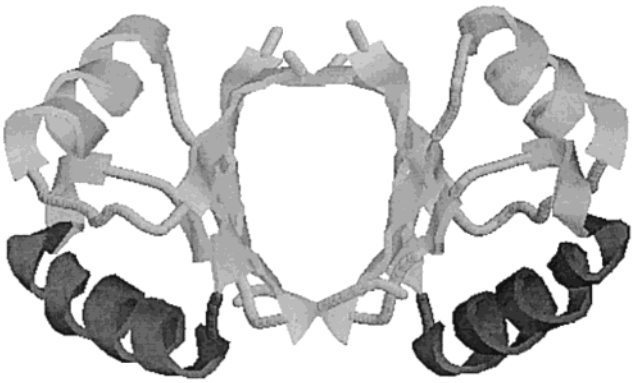

FIGURE 1: Schematic representation of the HPV-16 E2C dimer. The diagram was drawn from the crystallographic coordinates deposited. Protein Data Bank code 1by9 (*13*). DNA recognition helices are shown in dark gray.

1). This topology is shared only by EBNA1, another viral OBP from the Epstein−Barr virus, yet there is no sequence homology between them (*11*). Structures for the E2C domains from human papillomaviruses HPV-31 and HPV-16 are also available (*12*, *13*) and have the same folding topology as the BPV domain (*14*).

The consensus recognition sequence for the different E2C domains is $ACCGN_4CGGT$, and the binding affinities range from 10^{-9} to 10^{-11} M (*15*−*19*). To understand the binding mechanism with molecular detail, we aim at a spectroscopic study in solution in conditions of reversibility (*20*). We make use of fluorescein-modified oligonucleotides as well as intrinsic fluorescence of trp residues and circular dichroism (CD). Our work is centered on the E2C domain of the high risk strain HPV-16 E2C (hE2C), for which we have a preliminary characterization of its folding mechanism (*21*, *22*) and the effect of the specific DNA on its conformation and stability (*23*). We determined that the HPV-16 E2C domain binds much more tightly, has a strong tendency to reversible aggregation in the absence of salt, displays a higher sequence discrimination capacity, and distorts its cognate DNA sequence in a different way than the bovine domain. We discuss a number of distinctive properties in connection to structural evidence and the known different transcriptional regulation properties of both viruses.

EXPERIMENTAL PROCEDURES

Chemicals. All reagents and buffers were from maximum purity available. Urea, guanidine chloride, sodium chloride, DTT, and Bis-Tris buffer were from ICN (Costa Mesa, CA). Restriction endonucleases and T4 DNA ligase were from New England Biolabs (Beverly, MA).

DNA Synthesis. Double-stranded 18 bp oligonucleotides containing one E2 recognition sequence (site 35 in the HPV-16 genome) were prepared as follows: single-stranded oligonucleotides were purchased, HPLC purified, from Integrated DNA Technologies (Coralville, IA). Site 35A: 5′ GTA *ACCG* AAAT *CGGT* TGA 3′ (recognition sequence italicized) Fl-site35B is the complementary strand with a fluorescein molecule attached to the 5′ end via a 6 carbon linker. Single-stranded oligonucleotide concentration was calculated using the molar extinction coefficient obtained from the nucleotide composition. Annealing was performed by mixing equal amounts of the oligos in 10 mM Bis-Tris-HCl buffer, pH 7.0, and 20 mM NaCl, incubating 5 min at 95 °C, and slowly cooling to 25 °C for 16 h. This yielded a double-stranded oligonucleotide termed Fl-site35, and no detectable single-stranded oligonucleotide was present as judged by PAGE (not shown). A similar procedure was followed to anneal the BPV-derived oligonucleotide: 5′ CCG *ACCG* ACGT *CGGT* CGG 3′ (site BPV). This sequence corresponds to the same oligonucleotide found in the cocrystal structure of the BPV-1 E2C−DNA (*10*). The double-stranded EBNA1 oligonucleotide used as nonspecific DNA was 5′ GGG TAG CAT ATG CTA CCC 3′.

Protein Expression and Purification. The recombinant HPV-16 E2C (hE2C) was expressed in *Escherichia coli* BL21(DE3) and purified as described previously (*21*). Inverse PCR mutagenesis (*24*) was used to produce the HPV-16 E2-C W339F mutant, which was expressed and purified as the wild-type protein. DNA encoding the carboxy-terminal 86 amino acids from BPV-1 E2 (bE2C) plus an initiator methionine was obtained by PCR, using plasmid pTZE2kz as template (a generous gift from Alison McBride). The PCR fragment obtained was cloned into plasmid pTZ18u, and recombinant bE2C was expressed in *E. coli* JM109 as follows: cells were grown in 600 mL of 2YT medium (16 g of peptone, 10 g of yeast extract, and 5 g of NaCl/L) with 100 μg/mL ampicilin at 37 °C until $OD_{600} = 0.5$. IPTG was added to 0.4 mM, and expression was achieved infecting with phage M13/*T*7 (Invitrogen, San Diego, CA) at a multiplicity of infection of 5 and grown overnight. Cells were harvested, resuspended in 0.1 vol of lysis buffer (100 mM Tris-HCl, pH 6.8, 0.6 M NaCl, 1 mM EDTA, and 10 mM 2-mercaptoethanol), and sonicated at 0 °C twice. Inclusion bodies were collected by centrifugation at 20000*g* for 20 min, washed with 8 M urea, and dissolved in 500 mM Tris-HCl (pH 6.8), 10 mM DTT, and 6 M guanidine chloride. Refolding was accomplished by dialysis against 100 vol of 50 mM Tris-HCl (pH 6.8), 0.15 M NaCl, and 2 mM 2-mercaptoethanol. The refolded sample was loaded onto a Heparin HyperD affinity column (BioSepra, Villeneuve la Garenne, France) equilibrated with 50 mM Tris-HCl (pH 8.0), 0.15 M NaCl, and 2 mM 2-mercaptoethanol, and washed with 5 column vol of the same buffer. Bound protein was eluted with a 0.15−1.0 M NaCl linear gradient. Fractions that were >90% pure were pooled, dialyzed against sodium acetate buffer (10 mM sodium acetate, pH 5.6, 0.6 M NaCl, and 1 mM DTT), concentrated using Centriprep-3 (Amicon, Bedford, MA), and loaded onto a Superdex75 gel filtration column (Pharmacia Biotech, Uppsala, Sweden). This procedure yielded around 2 mg/L of >98% pure bE2C. The purified protein was stored as a 30 μM solution at −70 °C after snap freezing in liquid nitrogen. Protein concentration was determined using an extinction coefficient of 4.194×10^4 M^{-1} cm^{-1} for hE2C and 1.696×10^4 M^{-1} cm^{-1} for bE2C (*25*). The high salt (1.5 M) required to elute the human virus protein from the heparin column presumes a highly active dimer for DNA binding. There is also previous evidence of the correctly folded and active E2 domains, included those used in crystallization studies (*10*, *13*, *17*, *21*, *26*). However, to determine the extent of binding activity of our recombinant preparation, we mixed hE2C at 10 μg/mL with a 2-fold molar excess of biotinylated double-stranded site 35 oligonucleotide, and the complexes formed were challenged with streptavidin modified magnetic beads (Promega, Madison,

WI). After removal of the beads with the magnet, the remaining protein concentration in the solution measured by a Bradford assay was equal to the control tube without protein. Thus, we can assume with confidence that the recombinant E2C is over 90% active in specific DNA binding.

DNA Binding. Fluorescence measurements were recorded in a Aminco Bowman series 2 luminescence spectrometer assembled in "L" geometry. For fluorescein anisotropy measurements excitation was set to 495 nm with 4 nm slit and emission was recorded at 520 nm. When fluorescein concentration was below 10 nM, emission was collected through an orange shortwave cut filter OG515 (Schott, Mainz, Germany). The temperature was kept constant at 25 ± 0.1 °C through all experiments. Light scattering was measured setting excitation wavelength to 320 nm and integrating the emission from 330 to 340 nm. All titrations were performed adding small amounts of a concentrated solution of the variable ligand to a fixed amount of the other and allowed to equilibrate for 3−5 min. In all cases, maximal dilution was 20%, and the data were corrected accordingly.

Dissociation constants for E2C−DNA interaction were performed in 25 mM Bis-Tris-HCl (pH 7.0), 200 mM NaCl, and 1 mM DTT by measuring the steady-state fluorescence anisotropy and the fluorescence intensity of the Fl-DNA as a function of added E2C. Data were fitted using nonlinear least squares to the following equation:

$$[\text{E2C}-\text{DNA}] = 0.5\Delta F([\text{DNA}] + [\text{E2C}] + K_D) - [([\text{DNA}] + [\text{E2C}] + K_D)^2 - (4[\text{DNA}][\text{E2C}])]^{0.5}$$

where ΔF is the difference in the signal between the E2C−DNA complex and free DNA; [DNA] and [E2C] are the oligonucleotide and protein concentrations, respectively; and K_D is the dissociation constant for the interaction. No computational corrections for emission intensity were required, since the quantum yield did not change significantly upon binding.

Bis-ANS binding was performed in 25 mM Bis-Tris-HCl (pH 7.0), 200 mM NaCl, and 1 mM DTT. Excitation wavelength was set at 360 nm, and the emission was measured from 400 to 600 nm.

Circular Dichroism. CD spectra were monitored in the near-UV region using a Jasco J-810 equipment. Ten scans were averaged for each measurement at 25.0 ± 0.1 °C controlled by a peltier. The contribution of the protein was subtracted prior to plotting raw ellipticity at indicated wavelengths against concentration of added HPV-16 E2C. It is expected that in the 250−320 nm range, the ellipticity of the protein does not change substantially, and thus we can assume we are evaluating mainly changes in DNA conformation (*27*).

Electrophoretic Mobility Assay. Oligonucleotide band shift assay was performed in high ionic strength conditions (*28*). Reaction mixtures containing 25 mM Bis-Tris-HCl (pH 7.0), 0.2 M NaCl, 1 mM DTT, 1 μM Fl-site35, and different hE2C amounts were incubated 30 min at 25 °C in a final volume of 10 μL. Mixtures were loaded into a running 6% polyacrylamide gel containing 2.5% glycerol, 25 mM Tris-HCl (pH 8.3), 190 mM glycine, and 1 mM DTT. The gel was resolved at 4 V/cm for 2 h. Fluorescein bands were detected by UV transilumination. After visualization and documentation, the gel was silver stained.

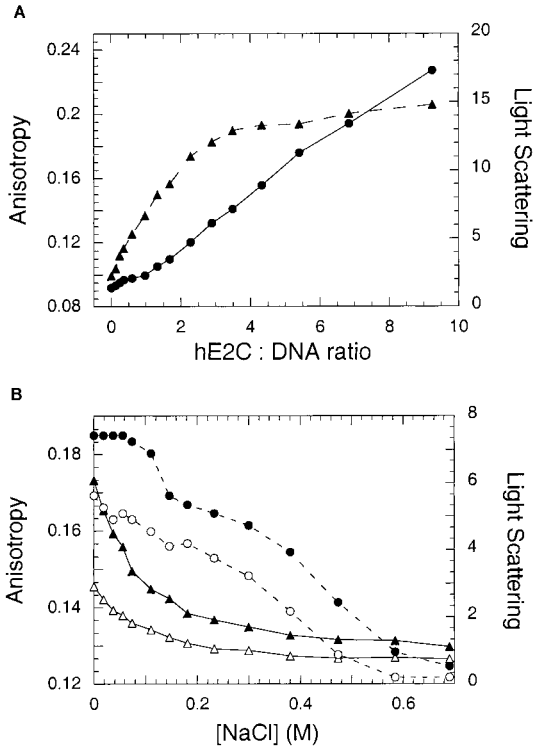

FIGURE 2: E2C−DNA complex form aggregates that are reversed at high ionic strength. (A) Fl-site35 (25 nM) was incubated in 50 mM Bis-Tris-HCl, pH 7.0, and 1 mM DTT with increasing concentrations of hE2C: (▲) anisotropy, (●) light scattering. (B) Different molar ratios of hE2C, 5:1 (black) and 2:1 (white), where incubated and concentrated NaCl was gradually added to the solution: (▲) anisotropy, (●) light scattering.

RESULTS

E2C−DNA Complex Undergoes a Reversible Ionic Strength-Dependent Aggregation in Excess of Protein. To study the hE2C−DNA interaction mechanism in solution, we start by analyzing the binding reaction with minimal additions to the buffer and measure the fluorescence anisotropy changes that take place upon binding of the fluorescein (FITC)-modified DNA oligonucleotide to the purified recombinant domain. The oligonucleotide corresponds to one of the four natural sites present in the long control region of the HPV-16 genome (site 35) (*29*). At 25 nM fixed DNA concentration and pH 7.0, the fluorescence anisotropy increases as the concentration of the E2C domain is increased (Figure 2A). At pH 5.6, the anisotropy change reaches even higher final values (not shown), and the magnitude of the observed changes at both pH values suggest the formation of a high molecularity oligomer or aggregate. This was confirmed by measuring the light scattering in parallel, which increases 10-fold the value at equimolar protein:DNA ratio (Figure 1A), supporting the idea of the formation of a large oligomer in molar excess of protein over DNA. The FITC fluorescence intensity increases up to 1:1 hE2C:DNA ratio and abruptly decreases after this point, further supporting the idea of an aggregation process (not shown). Centrifugation of the samples at 15000*g* after the titration led to a large loss of signal. Furthermore, this behavior can be observed even at concentrations as low as 2 nM protein and in the

presence of excess of reductant, suggesting that it is not coupled to redox modifications. The HPV-16 E2C domain does not aggregate in these conditions, and different crystallographic and NMR studies further support a high solubility.

The effect of ionic strength was tested at different hE2C:DNA ratios in order to eliminate the observed aggregation. By different ratios, we refer to those involving E2C as a dimer in solution combined with the double-stranded site 35 oligonucleotide. The E2C domain exists as a dimer throughout the conditions used in this work, since the dimer−monomer dissociation constant in solution in similar conditions was calculated to be 10^{-13} M (*21*). At 5:1 and 2:1 E2C:DNA ratios, there is a large decrease in both anisotropy and light scattering, as the salt concentration was increased, and the end of the transition corresponds to complete dissociation of the protein−DNA complex (Figure 2B). An evident pretransition in anisotropy change can be observed up to around 0.15−0.2 M NaCl (0.02 unit), and there is a further decrease in anisotropy (0.045 unit) from there to the end of the complete transition at both 5:1 and 2:1 ratios (Figure 2B). We performed titrations of site 35 DNA with increasing concentrations of E2C at 0.2 M NaCl and found that the change in light scattering is very small and parallels a discrete change in anisotropy (not shown), compatible with absence of aggregation. Therefore, we use a concentration of 0.2 M NaCl in experiments throughout this work. No transition was observed when titrations were conducted at 0.6 M NaCl, indicating the absence of binding, where the anisotropy sloping value is coincident with the end of the dissociation transition described in Figure 2B.

Stoichiometry of Binding. The absence of aggregation at 0.2 M NaCl prompted us for a detailed titration experiment at higher concentrations in order to determine the binding stoichiometry. At 200 nM oligonucleotide and pH 7.0, the anisotropy changes on titration with increasing protein concentration reaches values beyond 1:1 ratio and saturates at 2:1 hE2C:DNA ratio with a total change of 0.045 anisotropy unit (Figure 3A). There is a small but measurable change in the FITC fluorescence (less than 15%) that takes place in absolute parallel with the anisotropy change. By lowering the pH to 5.5 in otherwise similar buffer conditions, the stoichiometry is clearly 1:1 hE2C:DNA (Figure 3A). These two observations confirm a high DNA binding activity of the recombinant domain and the unexpected 2:1 stoichiometry at pH 7.0.

An alternative approach for the analysis of the binding stoichiometry is by monitoring possible changes in the intrinsic fluorescence of tryptophan residues in hE2C upon binding to the site 35 DNA oligonucleotide. A detailed fluorescence spectral analysis indicated an increase in fluorescence caused by DNA binding (*23*). A titration experiment monitoring the tryptophan fluorescence in response to added nonfluoresceinated DNA showed a break at 0.5:1 DNA:hE2C ratio, but the post-transition baseline showed a significant slope that complicated the analysis (Figure 3B, inset). The hE2C domain has two buried tryptophan residues per monomer at the center of the barrel and one tryptophan completely exposed to the solvent, located at the opposite side of the molecule with respect to the DNA binding site (*13*). We replaced this latter tryptophan for a phenylalanine residue, and the resulting mutant is identical to the wild type protein in terms of stability and

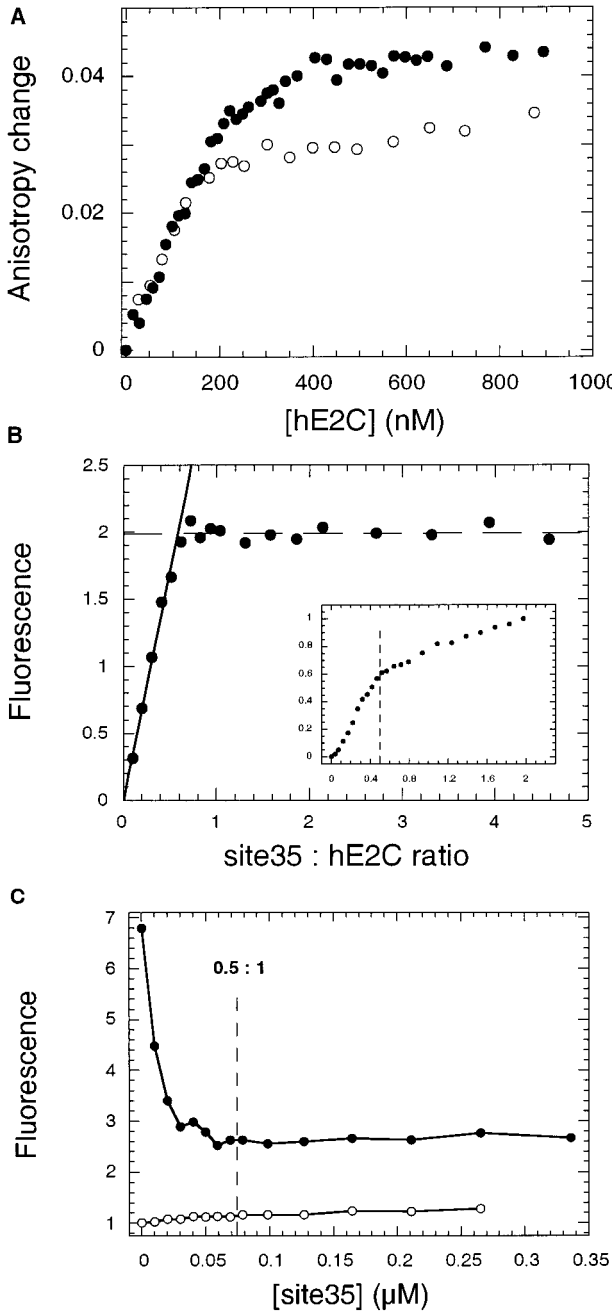

FIGURE 3: Stoichiometry of the hE2C−DNA interaction. (A) Fl-site35 (200 nM) in 200 mM NaCl, 1 mM DTT, and 25 mM Bis-Tris-HCl, pH 7.0 (●), or 25 mM acetate buffer pH 5.5 (○) was titrated with increasing amounts of hE2C, and the fluorescein fluorescence anisotropy was measured. (B) Titration of the W339F hE2C mutant (200 nM) with nonlabeled site 35 DNA monitored by intrinsic tryptophan fluorescence (inset, same titration for wild-type E2C). (C) bis-ANS (600 nM) solution was incubated with increasing concentrations of site 35 either in the presence (●) or in the absence (○) of 150 nM hE2C. DNA binding activity of the E2C preparations is >98% (see Experimental Procedures).

DNA binding (L.G.A., D.U.F., and G.P.G., unpublished results). This mutant also displays a tryptophan fluorescence change upon DNA binding, and the flat post-transition slope shows that the brake is effectively at 0.5:1 DNA:hE2C, confirming the proposed stoichiometry (Figure 3B) and ruling out the possibility that the stoichiometry is due to the presence of the FITC modification.

An independent method for evaluating the stoichiometry is the eventual displacement of the hydrophobic dye bis-

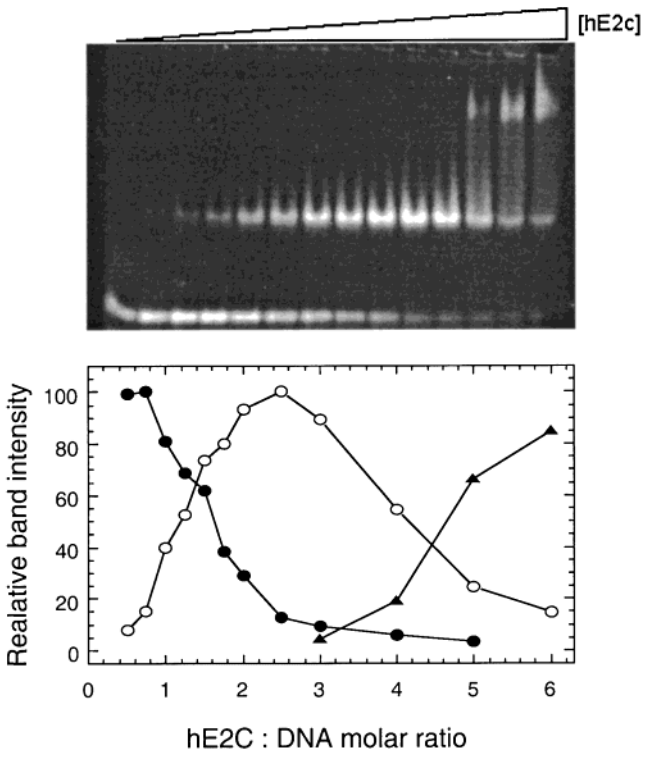

FIGURE 4: hE2C−DNA interaction analyzed by native gel electrophoresis. Fl-site35 DNA double-stranded oligonucleotide was incubated at 2 μM with different molar ratios of hE2C and subjected to native poliacrylamide gel electrophoresis as described in experimental procedures. (A) Direct visualization of the fluorescein label. (B) Quantification of the fluorescent bands: (●) free DNA, (○) 1:1 E2C−DNA complex, and (▲) 2:1 E2C−DNA complex. Silver staining of the gel confirms that all shifted bands contain protein (not shown).

ANS from the DNA binding surface. We have previously reported that hE2C binds two molecules of bis-ANS, one displaced by DNA and the other elsewhere in the structure (*23*). The bis-ANS was effectively displaced in a titration experiment with site 35 DNA (Figure 3C). The transition in bis-ANS fluorescence is complete at a ratio of 0.5:1 DNA:hE2C, confirming the stoichiometry obtained from two other independent methods. A control experiment shows that the DNA has no effect on the fluorescence of bis-ANS in the absence of hE2C (Figure 3C).

To test our stoichiometry results with a nonspectroscopic technique, we analyzed the E2C−DNA interaction by a gel retardation experiment in native polyacrylamide electrophoresis. In our simplified buffer conditions (pH 7.0, 0.2 M NaCl and 1 mM DTT), we mixed hE2C and Fl-site35 DNA and monitored the formation of protein−DNA complexes by native gel electrophoresis (Figure 4A). Lane 1 corresponds to the free oligonucleotide, which is retarded upon the gradual addition of hE2C. The 1:1 hE2C:DNA complex saturates at a 2.5:1 protein:DNA ratio, with an apparent K_D around 1 μM, which clearly indicates that the electrophoretic conditions are dissociating, perhaps due to the lack of a number of buffer additives (*20*). However, the 1:1 complex disappears concomitantly with the appearance of a slower migrating band at high protein concentration, which we ascribe to the formation of a 2:1 hE2C−DNA complex. Figure 4B also shows a quantitative analysis of the electrophoretic experiment where, although the affinity is lowered drastically by the electrophoretic conditions, the equilibrium between populations 1:1 and 2:1 hE2C−DNA complexes is evident. The 1:1 complex increases as the free DNA disappears and subsequently disappears as the 2:1 complex accumulates, strongly suggesting that the 2:1 complex is formed at the expense of the 1:1 species.

Analysis of Binding. The stoichiometric titration experiment monitored by anisotropy of the FITC site 35-DNA indicated two binding events. At 2 nM DNA, the total anisotropy change corresponded to 0.02 unit, roughly half of the change observed in the titration at 200 nM (not shown). This smaller change in anisotropy at a 1:1 hE2C:DNA ratio led us to consider that the high-affinity binding could be analyzed separately in experiments at lower concentrations. Preliminary titrations showed that the dissociation range for the high-affinity site requires very low concentration of DNA, which precludes the use of anisotropy measurements. Figure 5A shows the fluorescence intensity data used for the analysis at three different concentrations for defining the high-affinity binding event. Assuming a 1:1 stoichiometry, we obtained an average K_D of 0.18 ± 0.03 nM for the high-affinity binding event. The random distribution of the residuals confirms the binding model used (Figure 5A). To measure the second binding event, we mixed a hE2C−DNA complex at a 1:1 ratio and added increasing amounts of protein (Figure 5B). A similar analysis yielded a K_D of 1.8 ± 0.5 nM for the lower-affinity binding event, averaged from three experiments at different concentrations. We assume this corresponds to the binding of a second molecule of hE2C dimer to the 1:1 complex; nevertheless, the precise structural arrangement of this remains to be elucidated.

The ability of DNA binding proteins to recognize their specific targets among the variety of sequences present in genomes relies on their capacity to discriminate specific from nonspecific DNA. As the source for nonspecific DNA binding, we measured the affinity for an oligonucleotide corresponding to the palindromic site of the Epstein−Barr nuclear antigen 1 (EBNA1), an OBP with a similar folding topology (*11*) to E2C. This OBP displays no amino acid sequence homology and recognizes a completely different DNA target. The values for all the dissociation constants determined in the same equilibrium conditions are summarized in Table 1.

Comparative Binding Analysis of the Bovine (BPV-1) and Human (HPV-16) E2C Domains. The only structural analysis on E2C−DNA interaction was obtained from the BPV1 domain, for which there is also a wealth of information regarding transcription regulation (*3*, *14*). We intend to analyze the binding of the BPV-1 E2C domain in similar buffer conditions to those used for HPV-16 E2C and the same oligonucleotide target corresponding to an HPV-16 site. We subcloned, recombinantly expressed, and purified the domain used for the crystallographic studies (*10*, *14*).

Titration of an FITC-modified, bovine-derived site with BPV1-E2C at 0.2 M salt and pH 7.0 produces a discrete change in anisotropy, and unlike the human counterpart, a second binding event is not detectable within the concentration range reached in the experiment (Figure 6). A similar behavior is observed when the bovine domain was challenged with the HPV-16 site 35 oligonucleotide (not shown). The values for the dissociation constants for these two sites were determined at near-dissociation conditions and are shown in Table 1. The binding properties of the HPV16 E2C domain

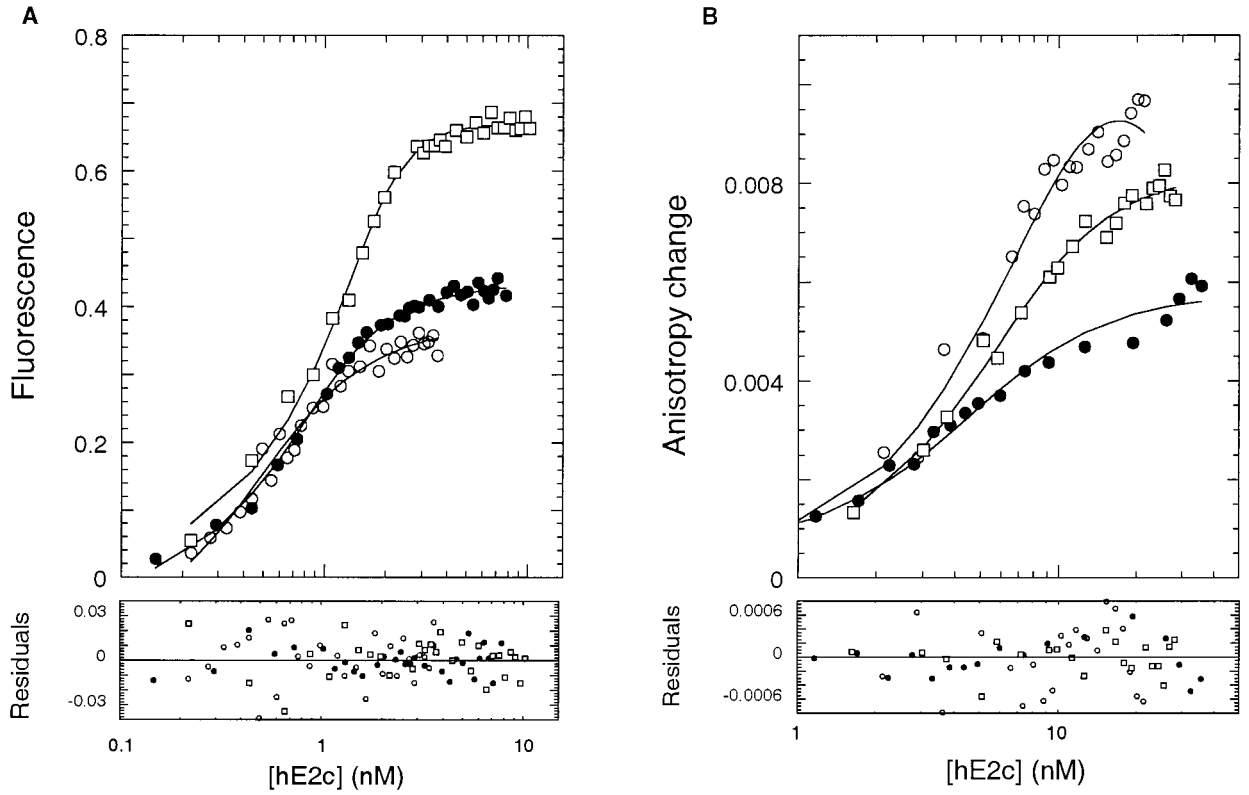

FIGURE 5: Determination of dissociation constants. Binding experiments were performed by adding hE2C to a fixed amount of oligonucleotide as described in experimental procedures. (A) High-affinity binding was carried out at the following fixed DNA concentrations: (○) 0.5, (●) 1.0, and (□) 2.0 nM. (Lower panel) Residuals for the fittings to eq 1. The calculated K_D were 0.19 ± 0.04, 0.17 ± 0.02, and 0.18 ± 0.02 nM, respectively. (B) Binding of the second hE2C dimer to site 35 was carried out as in panel A, starting at a 1:1 hE2C−DNA complex at the following fixed concentrations: (●) 5.0, (□) 7.0, and (○) 10.0 nM. (Lower panel) Residuals. Calculated K_D were 1.6 ± 0.4, 1.4 ± 0.2, and 2.1 ± 0.3 nM, respectively.

Table 1: Affinities and Sequence Discrimination of HPV-16 and BPV-1 E2C Domains

	site 35[a] (HPV-16 cognate)	site BPV (BPV-1 cognate)	EBNA (nonspecific)	cognate ratio[c]	specificity ratio[d]
HPV-16 E2C	1: 0.2 ± 0.02[b] 2: 2 ± 0.3	1: 30 ± 15 2: 3000 ± 1200	2000 ± 80	150	10000
BPV-1 E2C	700 ± 70	200 ± 23	11000 ± 1600	3.5	55

[a] Dissociation constants (K_D) are all expressed in nanomolar. [b] K_D values for the two binding modes of HPV-16 E2C are shown. Site 1 was used for calculation of the ratios. [c] $K_{Dcognate}/K_{Dnoncognate}$. [d] $K_{Dspecific}/K_{Dnonspecific}$.

for the bovine-derived site were also evaluated for comparison (Table 1). Once more, binding isotherms in identical buffer conditions and concentration show that the human domain is capable of two binding events, similar to what was observed for the cognate site 35 (Figure 6). It should be noted that while the affinity of the second binding event of HPV-16 E2C is 10-fold lower than the highest affinity binding for the human site 35 DNA, this ratio is 100-fold lower for the bovine site (Table 1).

The ratios between the different dissociation constants provide an idea of the sequence discrimination capacity of the domains against noncognate and nonspecific DNA sites of equal length. The HPV16 E2C discriminates 150-fold against the noncognate (BPV1) site, while it has a 10 000-fold higher affinity to its cognate site with respect to a nonspecific DNA oligonucleotide. The discrimination capacity of the bovine domain is only 3.5- and 55-fold against noncognate and nonspecific sequences, respectively.

Conformational Changes of Bound DNA. The near-UV region of the CD spectra is particularly useful when analyzing DNA conformation in solution (*27*). In similar buffer conditions to the binding experiments but at concentrations well above the K_D, the conformation of the site 35 DNA

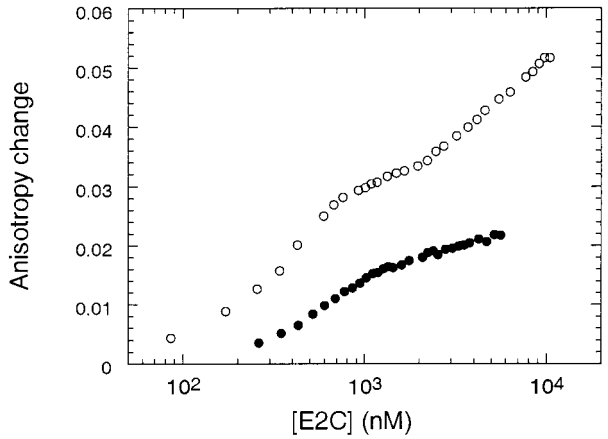

FIGURE 6: Interaction of BPV-1 and HPV-16 E2C with BPV-derived DNA. 500 nM Fl-site BPV was incubated in 25 mM Bis-Tris-HCl, pH 7.0, 0.2 M NaCl, and 1 mM DTT. BPV-1 E2C (●) or HPV-16 E2C (○) was added. The change in fluorescein fluorescence anisotropy was monitored.

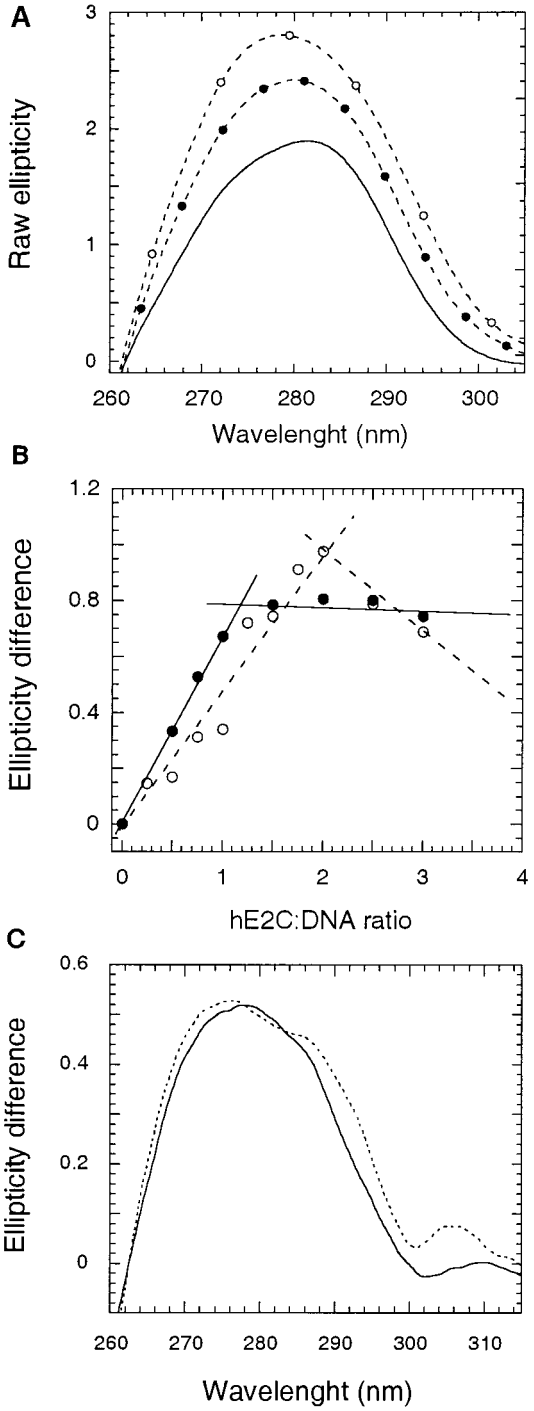

FIGURE 7: DNA conformational changes induced by E2C binding. (A) 1 μM site 35 and different concentrations of HPV16-E2C were incubated in 25 mM Bis-Tris-HCl buffer, pH 7.0, 0.2 M NaCl, and 1 mM DTT. CD spectra were recorded. Solid line, site 35; plus 1 (●) or 2 μM (○) hE2C. (B) Titrations of 1 μM site 35 with hE2C followed by the molar ellipticity at 270 (○) or 292 nm (●). (C) 1 μM site 35 was incubated in equimolar concentrations with either HPV-16 E2C (dotted line) or BPV-1 E2C (solid line). CD spectra were recorded, and the differential spectrum are shown.

changes upon binding of 1 mol of E2C (Figure 7A). The increase in the CD band in this region can be ascribed to DNA unwinding and partial base unstacking (*27*). Further conformational changes take place when a second molar equivalent of hE2C dimer is added, supporting the existence of a second DNA binding event, in agreement with the stoichiometry observed in previous experiments. Careful titration experiments based on difference spectra show that the band at 270 nm saturates at a 1:1 hE2C:DNA stoichiometry, while at 292 nm the ellipticity changes up to a 2:1 ratio (Figure 7B).

We have shown that the BPV1 E2C domain binds to HPV 16 site 35 DNA and even for its cognate BPV1 site with lower affinity than the human virus domain. A recent detailed investigation on structure-binding properties of E2 DNA sites showed that the HPV 16 E2C domain displays increased affinity for targets with a decreased flexibility and that it is more sensitive to sequence-dependent conformational variations of the DNA depending on the spacer regions (*30*). Furthermore, based on the HPV16-E2C crystal structure, Hegde et al. (*13*) hypothesized that the DNA must be differentially distorted in bovine and human domains in order to accommodate their recognition helices. We wanted to test whether there are further experimental structural basis for the differential DNA binding properties observed for the two domains in solution, using a spectroscopic method. For this purpose, site 35 DNA was mixed with BPV1 E2C, and the near-UV spectrum was analyzed in comparison with that for HPV-16 E2C−DNA. The difference spectrum of the BPV1 E2C bound DNA shows a red-shift in the band near 275 nm and a decrease in the band at 292 nm (Figure 7c). A small but clear band at 305 nm in the HPV-16 E2C bound DNA is not present in the DNA−BPV1 E2C complex. Taking into account the fact that the DNA is the same, the observed spectral changes are significant and indicate that the bovine and human domains inflict a different conformation to the HPV-16 site 35 DNA when bound.

DISCUSSION

In the present work, we investigated the binding mechanism of the E2C domain of HPV-16 to DNA in solution. The large complexity of this type of reaction within the cell nucleus involves numerous variables such as molecular crowding, ionic strength, pH, redox state, interaction with other proteins, macromolecule concentration, etc. Our aim was to analyze the process in a simplified buffer, keeping additives to a minimum and using an equilibrium technique that yields a reversible binding reaction in solution that will hopefully allow a detailed characterization of the underlying molecular mechanism.

At low ionic strength, an aggregation process predominates in excess of protein over DNA. This process was shown to be fully and instantly reversible by the addition of salt. The tendency to aggregation also appears more pronounced at lower pH. Thus, two macroscopic solvent variables strongly influence the oligomeric state of the DNA-binding domain of this transcriptional regulator bound to its specific DNA, within physiological range. BPV1 E2C appears not to aggregate in similar conditions as the human domain does (not shown). However, this observation is not conclusive since experiments should be 100-fold above its already high K_D, too high for a reliable interpretation of fluorescence spectroscopic experiments.

In a previous work, we reported a large stabilization of the free dimeric hE2C from pH 5.6 to pH 7.0 (*21*), a strong stabilizing effect by ionic strength (*23*), and a pronounced resistance to reversible pressure denaturation on going from pH 5.5 to pH 7.0 (*31*). We have discussed and continue to

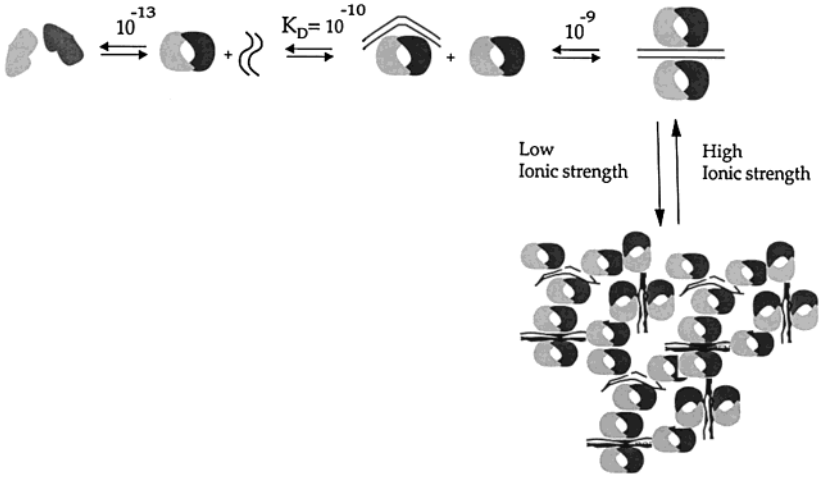

FIGURE 8: Proposed model for the DNA binding mechanism of E2C in solution. hE2C forms a dimer in solution with a very low dissociation constant. This dimer binds the specific target DNA sequence in two successive binding modes, which can lead to an aggregation depending on the solution ionic strength. The structural basis for the 2:1 species and for the aggregate remain to be elucidated; therefore, the drawing is purely schematic.

add evidence in the present work that the E2C domain displays a substantial structural plasticity that allows it to accommodate conformational changes that result in stabilization with DNA, ionic strength, and pH, to mention the more relevant parameters (*23*). Although we cannot confirm at this stage which of these variables can be regulatory relevant in vivo, it becomes clear that the biochemical properties of this domain exceed those of a DNA-binding reaction.

At 0.2 M salt and pH 7.0, a second binding event is evident in the hE2C−site 35 DNA interaction, confirmed by different approaches such as fluorescence anisotropy change of the FITC−DNA, intrinsic fluorescence of the trp residues, displacement of bound bis-ANS by DNA, and CD. Supporting this observation, a second shifted band was evident in a native electrophoresis band retardation experiment, and this band increased concomitantly with the disappearance of the first shifted band, corresponding to the 1:1 complex.

By lowering the concentration of the cognate site 35 DNA in the binding experiments, we were able to measure separately the higher affinity 1:1 binding event with a K_D of 0.18 nM for the HPV-16 E2C domain. The second binding event shows a K_D of 1.8 nM, which we believe is too high to be considered nonspecific, still 17-fold the affinity for the noncognate bovine E2 site. Binding of HPV-16 E2C to the palindromic site of the EBNA1 DNA binding domain (*11*) was 3 orders of magnitude higher, setting the range for nonspecific DNA binding by hE2C.

Spectroscopic binding analysis of HPV11-E2C interaction with a similar E2 site showed a single binding event, but this could reflect a difference in experimental conditions. More likely, this discrepancy could also arise from intrinsic differences in binding properties of the domains from these two human strains, as there is also a difference in the number and distribution of E2 binding sites in the respective genomes (*26*). Both viruses produce mucosal infections, but while HPV-11 is a low risk strain, HPV-16 is high risk for cervical carcinomas (*5*). One can speculate that these differences reflect general features of E2C−DNA recognition in low and high risk viruses, an hypothesis well worth testing.

Our working model of the DNA-binding mechanism of hE2C in solution in our simplified buffer conditions is shown in Figure 8. With a reliable method in solution to measure the DNA binding, we can now compare it with the dissociation/unfolding of the free E2C dimer. At similar ionic strength and pH at which the binding experiments were performed, the dissociation constant of the free dimer is 0.14 pM (*21*). Our present E2C:DNA binding results at pH 7.0 indicate that the dimerization and DNA binding processes appear not to be coupled as we previously suggested (*21*). The affinity of the second binding event is 350-fold higher than that for nonspecific binding determined in this work. It is not possible to know at this stage the structural basis for the 2:1 complex nor that of the aggregate; therefore, the representation is purely schematic. We have no evidence suggesting that the two E2C dimer molecules that bind to the DNA form a tetramer, and there is no precedent that the E2C can form tetramers in solution. Instead, we believe that the second E2C dimer binds to the DNA in a different mode that remains to be elucidated in structural terms.

As Figure 8 summarizes, the aggregation is readily reversible by changes in ionic strength. Several reports mentioned the presence of higher order complexes in gel retardation experiments, but these were not analyzed or discussed in detail (*29*, *32*). Evidence from different systems indicate that aggregation could be a physiological process taking place at DNA transcription and replication origins rather than an unwanted artifact to be avoided with additives (*33*−*36*). Under similar conditions, a second binding mode was not observed in the BPV1 domain for either the human or bovine DNA sites. However, the existence of a second binding event at protein concentrations higher than those used in this work cannot be ruled out. The fact that the HPV-16 domain displays a second binding mode with both human and bovine DNA sites suggests that it is a property of the protein rather than of the DNA. Two binding modes were also observed for other model DNA-binding proteins (*37*, *38*).

The binding of the BPV-1 E2C domain to the HPV site 35 DNA showed a much weaker binding than the human domain. Even for a bovine-derived site, the affinity is 7-fold lower than that for the HPV-16 E2C challenged with the same site (Table 1). Furthermore, the ability to discriminate

against a noncognate sequence is 50-fold higher for the human domain, and the latter is 180-fold better at discriminating against a nonspecific DNA sequence. Thus, the lower affinity of BPV1-E2C for a bovine-derived site appears to be linked to the low discrimination capacity not only for noncognate but also for nonspecific sequences. This may vary somehow for all the different bovine sites in the genome, but the evidence is convincing if we combine ours with those from other groups (*30*).

There are four E2 binding sites in HPV-16 whereas the BPV-1 genome bears 17 of these sites (*15*), and the BPV domain shows lower affinity for both HPV-16 and BPV-1 type sites (this work and ref *30*). It is believed that the lower affinity is compensated by a larger number of sites in the BPV-1 genome, and it is also generally accepted that the more electrostatic the nature of the protein−DNA binding, the weaker and less specific the interaction. While the more electropositive surface of the BPV-1 E2C may help bending the DNA through interactions with the phosphate backbone (*13*), this is counter-balanced by a weaker binding. A close comparison of the crystal structures of the free HPV-16 and BPV-1 E2C domains showed that despite the close structural similarity there are significant differences. One of these differences is the electrostatic potential distribution at the DNA interaction surface (*13*); HPV-16 E2C is less (positively) charged than the bovine counterpart.

From the structural comparison of the free domains and the bovine E2C−DNA complex, it was proposed that the DNA must be in a different conformation in the HPV-16 E2C−DNA complex in order to maintain symmetric contacts with both helices (*13*). The ability to discriminate among different specific and nonspecific sequences and the apparent sensitivity to salt effects agree very well with what was proposed from the crystallographic results (*13*). In any case, we expect that the differences will not only take place on the protein but also on the conformation of bound DNA.

No information on the DNA conformation is available for a human E2C−DNA complex to date neither in a crystal or in solution, something that prompted us for a spectroscopic analysis in our experimental conditions. Near-UV CD spectra show a considerable conformational change in site 35 DNA when it binds to its cognate HPV-16 E2C domain. The nature of the change suggests partial unwinding and base unstacking, indicating a significant degree of deviation from the canonical B-conformation of duplex DNA. More importantly, the conformation of the same HPV-16 site 35 DNA bound to the bovine domain is markedly different. Although structural details cannot be extracted from these spectra, this further argues for differences in the DNA binding strategy of both homologous domains. Previous studies showed that HPV-16 E2C displays greater affinity for more rigid targets, where the spacer DNA sequence favors prebending through compression of the minor groove (*39*). The bovine domain shows much less discrimination capacity for sequences in the noncontacted spacers. The bovine E2C must therefore bend both flexible and rigid targets, resulting in a lower overall affinity. DNA bending by BPV1 E2C must involve change of winding angle and base pair twist generating a distinctive CD spectrum.

The DNA binding domain of HPV-16 E2C has particular biochemical properties besides its DNA binding activity, i.e., oligomerization or aggregation through protein−protein contacts with itself or other viral and nuclear factors. It is not impossible to picture a recruiting activity at origin binding sites that include aggregation as a key step for its biological activity. In addition, full-length BPV1 E2 protein was shown to participate in DNA loping by electron microscopy, and an aggregation process was hypothesised to be a necessary step for this activity (*40*). Although significantly lower than the full-length protein, the E2C domain also displayed DNA looping activity.

Human and bovine domains appear to differ significantly in their DNA recognition strategies, despite the strong overall similarities. Recently, the repression of E6/E7 oncogenes under HPV16 and 18 regulatory regions (LCR) containing the respective human E2 binding sites was shown to be mediated by HPV-16 E2 TA, but the BPV1 TA had similar effects on the same construct (*9*). In light of the 150-fold difference in DNA binding affinity we now describe for human and bovine proteins, it should be evaluated whether the presence of the N-terminal transactivation domains can have such a large effect on affinity of cognate versus noncognate discrimination. Alternatively the transient expression plasmids used in this type of studies could lead to E2 TA levels well above those found in infected host epithelial cells and saturate even low-affinity binding sites. Finally, while the BPV E2-TA is a potent activator of early gene expression, no positive effect on gene expression has been reported for HPV-16 E2-TA.

The use of steady-state and dynamic spectroscopic methods in solution, in addition to detailed structural characterization of the HPV16 E2C−DNA complex, will uncover the fine but fundamental differences in protein−DNA sequence discrimination and will hopefully help elucidate the molecular mechanism of transcriptional regulation and DNA replication in this widespread human pathogen.

REFERENCES

1. Tan, S., and Richmond, T. J. (1998) *Curr. Opin. Struct. Biol. 8*, 41−48.
2. Edwards, A. M., Bochkarev, A., and Frappier, L. (1998) *Curr. Opin. Struct. Biol. 8*, 49−53.
3. McBride, A., Romanczuk, H., and Howley, P. (1991) *J. Biol. Chem. 266*, 18411−18414.
4. Piirsoo, M., Ustav, E., Mandel, T. A. S., Stenlund, A., and Ustav, M. (1996) *EMBO J. 15*, 1−11.
5. Laimins, L. A. (1998) *Human Tumor Viruses*, ASM Press, Washington, DC.
6. Giri, I., and Yaniv, M. (1988) *EMBO J. 7*, 2823−2829.
7. Gauthier, J. M., Dillner, J., and Yaniv, M. (1991) *Nucleic Acids Res. 19*, 7073−7079.
8. Skiadopoulos, M. H., and McBride, A. A. (1998) *J. Virol. 72*, 2079−2088.
9. Nishimura, A., Ono, T., Ishimoto, A., Dowhanick, J. J., Frizzell, M. A., Howley, P. M., and Sakai, H. (2000) *J. Virol. 75*, 3752−3760.
10. Hegde, R., Grossman, S., Laimins, L., and Sigler, P. (1992) *Nature 359*, 505−512.
11. Bochkarev, A., Barwell, J. A., Pfuetzner, R. A., Bochkareva, E., Frappier, L., and Edwards, A. M. (1996) *Cell 84*, 791−800.
12. Linag, H., Petros, A. M., Meadows, R., Yoon, H. S., Egan, D. A., Walter, K., Holzman, T. F., Robins, T., and Fesik, S. W. (1996) *Biochemistry 35*, 2095−2103.
13. Hegde, R. S., and Androphy, E. J. (1998) *J. Mol. Biol. 284*, 1479−1489.
14. Hegde, R. S., Wang, A. F., Kim, S. S., and Schapira, M. (1998) *J. Mol. Biol. 276*, 797−808.

15. Li, R., Knight, J., Bream, G., Stenlund, A., and Botchan, M. (1989) *Genes Dev. 3*, 510−526.
16. Monini, P., Grossman, S. R., Pepinsky, B., Androphy, E. J., and Laimins, L. A. (1991) *J. Virol. 65*, 2124−2130.
17. Sanders, C. M., and Maitland, N. J. (1994) *Nucleic Acids Res. 22*, 4890−4897.
18. Tan, S., Leong, L. E., Walker, P. A., and Bernard, H. (1994) *J. Virol. 68*, 6411−6420.
19. Chao, S. F., Rocque, W. J., Daniel, S., Czyzyk, L. E., Phelps, W. C., and Alexander, K. A. (1999) *Biochemistry 38*, 4586−4594.
20. Hill, J. J., and Royer, C. A. (1997) *Methods Enzymol. 278*, 390−416.
21. Mok, Y. K., Prat Gay, G. de, Butler, P. J., and Bycroft, M. (1996) *Protein Sci. 5*, 310−319.
22. Mok, Y. K., Bycroft, M., and Prat Gay, G. de (1996) *Nat. Struct. Biol. 3*, 711−717.
23. Lima, L. M., and Prat Gay, G. de (1997) *J. Biol. Chem. 272*, 19295−19303.
24. Clackson, T., Detlef, G., and Jones, P. (1991) in *PCR, a Practical Approach* (McPherson, M., Quirke, P., and Taylor, G., Eds.) pp 202, IRL Press, Oxford.
25. Pace, C. N., Vajdos, F., Lanette, F., Grimsley, G., and Gray, T. (1995) *Protein Sci. 4*, 2411−2423.
26. Alexander, K. A., and Phelps, W. C. (1996) *Biochemistry 35*, 9864−9872.
27. Gray, D. (1996) *Circular Dichroism and the Conformational Analysis of Biomolecules*, Plenum Press, New York.
28. Ausbel, F., Brent, R., Kingston, R. E., Moore, D. D., Siedman, J. G., Smith, J. G., and Struhl, K. (1997) *Short Protocols in Molecular Biology*, John Wiley & Sons, New York.
29. Thain, A., Webster, K., Emery, D., Clarke, A. R., and Gaston, K. (1997) *J. Biol. Chem. 272*, 8236−8242.
30. Hines, C. S., Meghoo, C., Shetty, S., Biburger, M., Brenowitz, M., and Hegde, R. S. (1998) *J. Mol. Biol. 276*, 809−818.
31. Foguel, D., Silva, J. L., and Prat Gay, G. de (1998) *J. Biol. Chem. 273*, 9050−9057.
32. Pepinsky, R. B., Prakash, S. S., Corina, K., Grossel, M. J., Barsoum, J., and Androphy, E. J. (1997) *J. Virol. 71*, 828−831.
33. Sevenich, F. W., Langowski, J., Weiss, V., and Rippe, K. (1998) *Nucleic Acids Res. 26*, 1373−1381.
34. Dekker, J., Kanellopoulos, P. N., Loonstra, A. K., van Oosterhout, J. A., Leonard, K., Tucker, P. A., and van der Vliet, P. C. (1997) *EMBO J. 16*, 1455−1463.
35. Wyman, C., Rombel, I., North, A. K., Bustamante, C., and Kustu, S. (1997) *Science 275*, 1658−1661.
36. Margeat, E., Le Grimellec, C., and Royer, C. A. (1998) *Biophys. J. 75*, 2712−2720.
37. Czernik, P. J., Shin, D. S., and Hurlburt, B. K. (1994) *J. Biol. Chem. 269*, 27869−27875.
38. Grillo, A. O., Brown, M. P., and Royer, C. A. (1999) *J. Mol. Biol. 287*, 539−554.
39. Rozenberg, H., Rabinovich, D., Frolow, F., Hegde, R. S., and Shakked, Z. (1998) *Proc. Natl. Acad. Sci. U.S.A. 95*, 15194−15199.
40. Knight, J. D., Li, R., and Botchan, M. (1991) *Proc. Natl. Acad. Sci. U.S.A. 88*, 3204−3208.

BI001694R