Equilibrium dissociation and unfolding of the dimeric human papillomavirus strain-16 E2 DNA-binding domain

YU-KEUNG MOK,¹ GONZALO DE PRAT GAY,²
P. JONATHAN BUTLER,³ AND MARK BYCROFT¹

¹ MRC Unit for Protein Function and Design, Department of Chemistry, University of Cambridge, Lensfield Road, Cambridge CB2 1EW, United Kingdom

² Departamento de Bioquímica Médica, Universidade Federal do Rio de Janeiro, Cidade Universitária, Rio de Janeiro 21910, Brazil

³ MRC Laboratory of Molecular Biology, Hills Road, Cambridge CB2 2QH, United Kingdom

(RECEIVED September 22, 1995; ACCEPTED November 9, 1995)

Abstract

The equilibrium unfolding reaction of the C-terminal 80-amino-acid dimeric DNA-binding domain of human papillomavirus (HPV) strain 16 E2 protein has been investigated using fluorescence, far-UV CD, and equilibrium sedimentation. The stability of the HPV-16 E2 DNA-binding domain is concentration-dependent, and the unfolding reaction is well described as a two-state transition from folded dimer to unfolded monomer. The conformational stability of the protein, ΔG_{H_2O} , was found to be 9.8 kcal/mol at pH 5.6, with the corresponding equilibrium unfolding/dissociation constant, K_u , being 6.5×10^{-8} M. Equilibrium sedimentation experiments give a K_d of 3.0×10^{-8} M, showing an excellent agreement between the two different techniques. Denaturation by temperature followed by the change in ellipticity also shows a concomitant disappearance of secondary and tertiary structures. The K_u changes dramatically at physiologically relevant pH's: with a change in pH from 6.1 to 7.0, it goes from 5.5×10^{-8} M to 4.4×10^{10} M. Our results suggest that, at the very low concentration of protein where DNA binding is normally measured (e.g., 10^{-11} M), the protein is predominantly monomeric and unfolded. They also stress the importance of the coupling between folding and DNA binding.

Keywords: dimer; E2 DNA-binding domain; equilibrium unfolding; human papillomavirus

Many DNA-binding proteins recognize palindromic or partially palindromic DNA sequences. The twofold symmetry of the target DNA sequence is reflected in the proteins, which are often dimeric with a twofold axis of symmetry (Harrison, 1991). This arrangement allows the DNA-binding regions of each monomer to interact with DNA sequences separated by one helical turn. This effectively produces a double binding site, increasing the strength and specificity of binding. Some dimeric DNA-binding proteins are able to form heterodimers in which the two monomers have the same DNA-binding/dimerization domain but have different activation or repression domains. This provides a means of regulating and modulating gene expression. A knowledge of the strength and specificity of the dimeric interactions is therefore critical to a complete understanding of the function of these proteins.

Papillomaviruses are small DNA viruses that infect the skin and mucosa of a wide variety of mammals (Howley, 1990). Historically, bovine papillomavirus (BPV) has been the most studied of the papillomaviruses, but the discovery of a link between papillomavirus infection and cervical cancer in humans has led to much recent interest in human papillomaviruses (HPVs), especially the high-risk subtypes such as HPV 16, 18, and 33 (Gissmann, 1992). Papillomavirus E2 protein is a DNA-binding protein that regulates viral transcription and replication (Ham et al., 1991; McBride et al., 1991). The protein recognizes the sequence ACCN₆GGT, which occurs several times in papillomavirus genomes. The E2 protein consists of an N-terminal activation domain and a C-terminal DNA-binding domain joined by a linker region rich in proline residues. The protein dimerizes via the DNA-binding domain. In addition to the full-length protein, two smaller species are produced by alternative splicing. The truncated proteins have deletions in the N-terminal domain but have the C-terminal dimerization/DNA-binding domain intact. These molecules can form heterodimers with the full-length E2 protein, and these heterodimers can act as transcriptional re-

Reprint requests to: Mark Bycroft, MRC Unit for Protein Function and Design, Department of Chemistry, University of Cambridge, Lensfield Road, Cambridge CB2 1EW, United Kingdom; e-mail: mb10031@cus.cam.ac.uk.

pressors (Barsoum et al., 1992). The sequence of the DNA-binding domain of E2 protein is conserved among papilloma-viruses. The crystal structure of the DNA-binding domain of BPV E2 complexed with the self-complementary oligonucleotide 5'-CCGACCGACGTCGGTCG-3' has been determined (Hegde et al., 1992). The protein contains an antiparallel β -barrel structure in which each monomer makes up one half of the barrel. Each monomer contains two helices that pack onto the outside of the barrel, with the longer one responsible for DNA binding. The DNA is bent smoothly around the protein barrel, encompassing the pair of helices in successive major grooves.

The structural features of the BPV-1 E2-DNA complex show a large interdependence between the formation of the dimer and the binding of the DNA to its symmetric sites. This prompted us to analyze the mechanism and strength of the folding/association of the dimer and compare these results with existing studies of its interaction with DNA. In this paper, we report the results of a study on the stability of the HPV-16 E2 dimeric DNA-binding domain. We show, using a variety of biophysical techniques, that the domain unfolds in a two-state process, with no evidence of an intermediate folded monomeric form.

Results

Recombinant expression and purification of the E2 DNA-binding domain

The C-terminal 80-amino-acid (286–365) DNA-binding domain of the HPV-16 E2 protein was overexpressed in *Escherichia coli* as a soluble, folded, and dimeric form. This domain is linked to the rest of the protein through a highly flexible hinge region, as shown in Figure 1A, and can exist as a structured domain retaining the ability to dimerize and bind DNA (Gauthier et al.,

1991). The C-terminal 89 amino acids of HPV-16 E2 (279–368) have also been expressed and purified and shown to bind to their conserved sequence ACC(N)₆GGT with high affinity (Sanders & Maitland, 1994).

The HPV-16 E2 DNA-binding domain has a very high affinity for the Heparin Hyper D column even in the presence of 0.6 M NaCl at pH 8.0. Protein eluted from this column is >90% pure (Fig. 1B, lane 4) and has an OD 280/260 ratio of about 1.7. A further gel filtration step in high-salt buffer removed remaining impurities and produced essentially homogeneous E2 protein for denaturation experiments (Fig. 1B, lanes 6–8). Purified protein migrated as the fastest band (MW $\approx$ 10) on a 15% SDS polyacrylamide gel.

Characterization of the dimeric domain by fluorescence and CD spectroscopy

As a first approach to characterizing the conformation of the HPV16 E2 DNA-binding domain, we obtained a fluorescence emission spectrum at pH 5.6, where the dimer is folded (Fig. 2). It shows a maximum of intensity at 340 nm owing to its three Trp residues. In the presence of 8 M urea at the same pH, the intensity is about the same, but a significant shift of the emission maximum to 350 nm is observed. At pH 2.6, the fluorescence drops to about 50% of the original intensity, with a concomitant shift of the emission maximum to 350 nm.

The far-UV CD spectrum of the E2 DNA-binding domain displays characteristics of α helix, with a maximum of ellipticity at 195 nm and two minima at 210 and 224 nm, respectively (Fig. 3A). These minima differ slightly from the typically observed values at 208 and 222 nm. This could arise from the high proportion of β -sheet content of the protein as well as slight variations in the helical conformations, because the pair of DNA-

Fig. 1. **A:** Schematic representation of the HPV-16 E2 protein. The C-terminal DNA-binding/dimerization domain is linked to the N-terminal transactivation domain via a middle "hinge region," as indicated by the zigzag line. Black cylinders represent α helices, gray rectangles represent β strands. **B:** Expression and purification of HPV-16 E2 DNA-binding domain in BL21 (DE3) pLys S strain. Lane 1, crude extract from induced bacterial cells; lane 2, 60% ammonium sulfate precipitation cut of the crude extract; lane 3, flow-through from Heparin Hyper D column; lane 4, pooled fractions after Heparin Hyper D affinity chromatography; lane 5, pooled fractions after gel filtration chromatography; lanes 6–8, the gel was overloaded with 8.6, 17.2, and 25.7 μ g of the protein, respectively, to show the purity.

Fig. 2. Fluorescence spectra of the native and unfolded HPV-16 E2 DNA-binding domain. Unfolding of HPV-16 E2 DNA-binding domain was obtained by adding 8 M urea or reducing the pH to 2.6 using 100 mM citrate-phosphate buffer with 1 mM DTT. Excitation wavelength, 280 nm; buffer, sodium acetate with 1 mM DTT, pH 5.6.

binding helices are wrapped around the β barrel. The protein contains three Trp residues and, of these, two are closely packed together in the dimeric interface. The aromatic side chain may also contribute in the far-UV region of the spectrum (Johnson, 1988). On addition of 6 M urea, all evidence of folded secondary structure disappeared. A comparable change is also observed at pH 2.6, except that there is a typical negative band at about 200 nm, which is a characteristic of acid-denatured proteins (Johnson, 1988).

The near-UV CD spectrum of the E2 DNA-binding domain (Fig. 3B) shows an overall maximum of ellipticity at around 280 nm, indicating that at least some of the aromatic residues, especially tryptophan, are in an asymmetric environment and that the protein is folded into its tertiary structure. Several bands can be observed, but they cannot be resolved because there are six aromatic residues that could contribute to the spectrum in this region.

Fig. 3. Far-UV and near-UV CD spectra of HPV-16 E2 DNA-binding domain. **A:** Far-UV CD spectra of the native and unfolded HPV-16 E2 DNA-binding domain in 6 M urea and at 100 mM citrate-phosphate buffer, pH 2.6. Concentration of the protein is 20 μ M in 50 mM sodium acetate buffer, pH 5.6, with 1 mM DTT, in a 0.05-cm path-length cuvette. **B:** Near-UV CD spectrum of the native HPV-16 E2 DNA-binding domain. Concentration of the protein is 10 μ M in 50 mM sodium acetate buffer, pH 5.6, with 1 mM DTT, in a 1.0-cm path-length cuvette.

Unfolding of the E2 DNA-binding domain by low pH and urea

In order to analyze the stability of the dimeric domain in relation to pH, we performed equilibrium acid denaturation experiments monitored by fluorescence. Figure 4 shows the pH denaturation transition at 10 μ M concentration of the dimer. A typical pH titration is observed, with a midpoint at around pH 3.5. When the pH denaturation is followed by ellipticity in the far-UV region, a transition caused by unfolding is also observed below pH 4.0 at 40 μ M protein concentration (Fig. 4, insert). However, the deviation at the intermediate pH points, which is caused by an intense negative signal at 215 nm, suggests that aggregation (not observed by fluorescence) might be occurring. Control equilibrium centrifugation experiments showed that an aggregation process is indeed taking place at pH 3.2 and 3.4 (data not shown). Nevertheless, similar sedimentation experiments conducted at pH 2.0 and 2.8 indicated the presence of the unfolded monomeric forms exclusively, whereas at pH 5.6 and above, only dimers were observed. This is a clear indication that the pH denaturation involves unfolding and dissociation into monomers at low pH.

To gain insight into the dissociation/unfolding of this dimeric DNA-binding domain, we conducted chemical denaturation experiments using urea. We followed the change in fluorescence intensity at 323 nm, where the maximum change occurs, with increasing urea concentration; a transition was observed from 1.0 to 2.5 M urea at 10 μ M dimer concentration (Fig. 5A). In order to evaluate the reversibility of the process, the protein was unfolded at a high concentration of urea and diluted into gradually decreasing concentrations of the denaturant. The process is largely reversible, as the comparison with the unfolding curve indicates (Fig. 5B). When a similar unfolding/dissociation experiment was performed at 1 μ M concentration of protein, the $[U]_{50\%}$ (the concentration of urea at the midpoint of the transition) was found to be 0.97 M (Fig. 5C; Table 1). This concentration dependence indicates that the transition involves a dissociation as well as an unfolding process. Changes in secondary structure with increasing denaturant concentration were studied by following the molar ellipticity change at 225 nm (Fig. 5D). At 10 μ M protein concentration, a transition is observed over the same range of urea concentration as for the flu-

Fig. 4. Stability of HPV-16 E2 DNA-binding domain toward acid denaturation. Fluorescence emission at 331 nm of 2 μ M and 10 μ M protein were monitored at different pHs. Buffer, 100 mM citrate-phosphate buffer with 1 mM DTT. Insert shows the pH denaturation of 40 μ M E2 by monitoring the far-UV CD signal at 228 nm at different pH's. The deviation at intermediate pH's was due to a highly negative broad-band at 215 nm, which probably indicates an aggregation process.

orescence change (Fig. 5D, insert). This indicates that secondary and tertiary structures are lost in a concerted manner and that there is no detectable folded monomer accumulated at equilibrium, because dissociation parallels unfolding.

We can, therefore, analyze the transition as a two-state model, where only folded dimer and unfolded monomers are populated. The $[U]_{50\%}$ is found to be 1.53 M for the ellipticity change and 1.51 M for the fluorescence experiment, both at a protein concentration of 10 μ M dimer. The free energy change for unfolding, ΔG_{H_2O} , from folded dimer to unfolded monomer is also similar in both cases (Table 1). As expected for a dissociation/unfolding process, the $[U]_{50\%}$ of the transition at 1 μ M protein concentration is significantly lower than the value at 10 μ M, that is, 0.97 M. Nevertheless, the calculated free energy changes are similar at different protein concentrations (Table 1). The free energy change from the equilibrium urea denaturation experiments is 9.80 ± 0.29 kcal mol⁻¹, equivalent to a K_u of 6.5×10^{-8} M.

The slight difference in m values between fluorescence and ellipticity could result either from the different fluorescence and CD spectroscopic characteristics of the unfolded state or from

residual structure in the otherwise unfolded state. Neither case would involve significant deviations from the $N_2 \leftrightarrow U$ model.

Thermal denaturation experiments followed by far- and near-UV CD

We monitored the thermal denaturation of the E2 DNA-binding domain by CD in both the far- and near-UV regions. At 214 nm, where the largest change in ellipticity takes place, a sharp symmetric transition can be observed with an apparent T_m of 57.05 °C (Fig. 6A). A similar transition occurs at 278 nm, where the aromatic residues in asymmetric environments probing the tertiary structure are responsible for the changes in ellipticity, with an apparent T_m of 54.33 °C (Fig. 6B). This indicates that the changes in secondary and tertiary structure occur in parallel, similarly to what happens in the urea denaturation. However, above 75 °C, the protein tends to aggregate, and the native spectrum is not recovered upon cooling (not shown). Under these conditions, we cannot obtain reliable thermodynamic parameters. The aggregation is likely to be caused by residual secondary structures, probably because the dimer is not completely dissociated at high temperatures.

Characterization of dimer-monomer equilibrium by sedimentation measurements

The solvent density (ρ_o) and the density of a solution of E2 protein at dialysis equilibrium with that solvent were determined for a range of concentrations of urea. These values, together with the protein concentration in each solution, were used to calculate values for the density increment $[(\partial\rho/\partial c_2)_m]$ and hence the apparent partial specific volume (ϕ'). The measured and calculated values are shown in Table 2. Values for ρ_o and ϕ' at additional intermediate urea concentrations, as required, were estimated by interpolation.

Plots of $\bar{M}_{w,app}$ against concentration were obtained by analysis of sedimentation equilibrium experiments with E2 solutions at a number of urea concentrations in sodium acetate buffer, pH 5.6. The protein appeared to be essentially dimeric at 0.0 M urea, with no visible evidence of dissociation (Fig. 7A), and almost exclusively monomeric at 3.0 or 5.0 M urea (Fig. 7A). However, at intermediate urea concentrations (0.5–2.7 M), $\bar{M}_{w,app}$ decreased in the range from dimer toward monomer. The data could be fitted to a monomer/dimer association equilibrium (Fig. 7A), and values of K_d obtained from this fitting are shown, plotted (as the logarithm) against urea concentration, in Figure 7B.

This plot shows a linear variation of $\ln(K_d)$, and, because

$$\Delta G = -RT \ln(K_d),$$

Table 1. Urea denaturation of the HPV-16 E2 DNA-binding domain analyzed by fluorescence and far-UV CD

	Protein conc. (μ M)	ΔG_{H_2O} (kcal/mol)	m (kcal/mol)	$[U]_{50\%}$ (M)	K_u (M)
Fluorescence	10	11.03 ± 0.23	3.07 ± 0.14	1.51	8.0×10^{-9}
Fluorescence	1	10.86 ± 0.50	3.20 ± 0.40	0.97	1.0×10^{-8}
Far-UV CD	10	9.80 ± 0.29	2.21 ± 0.16	1.53	6.5×10^{-8}

Fig. 5. Stability of HPV-16 E2 DNA-binding domain toward urea denaturation. **A:** Denaturation of HPV-16 E2 DNA-binding domain monitored by fluorescence at 323 nm. Protein concentration, 10 μ M; buffer, 50 mM sodium acetate, pH 5.6, with 1 mM DTT. The curve was fitted as described in the Materials and methods. **B:** Reversibility of HPV-16 E2 DNA-binding domain denaturation. The protein was fully denatured in 4.0 M urea and then diluted to lower the urea concentration to allow renaturation. Protein concentration, 10 μ M. Data were normalized to show the fraction of protein unfolded. **C:** Protein concentration dependence of the HPV-16 E2 DNA-binding domain to urea denaturation. Two protein concentrations of 1 μ M and 10 μ M E2 were used. Curves were generated by a simultaneous fitting of the two-state model at two different protein concentrations. **D:** Denaturation of HPV-16 E2 DNA-binding domain monitored by far-UV CD at 225 nm. Protein concentration, 10 μ M; buffer, 50 mM acetate, pH 5.6, with 1 mM DTT. Insert in Figure 5D shows a comparison between data obtained from fluorescence and far-UV CD. Data were normalized to the fraction of protein unfolded so that the value in the absence of urea was always 0.

ΔG must be linearly dependent upon the urea concentration. This justifies extrapolating the plot for $\ln(K_d)$ to 0.0 M urea and estimating K_d for the E2 protein association under these conditions from this intercept. This gives a value of $K_d = 3 \times 10^{-8}$ M and $\Delta G = 10.2$ kcal mol $^{-1}$.

Effect of pH on the unfolding/dissociation of HPV-16 E2 DNA-binding domain

Because the dimeric HPV-16 E2 DNA-binding domain is crucial for regulating the expression of viral genes in the host cells,

Fig. 6. Thermal denaturation of HPV-16 E2 DNA-binding domain. Denaturation was followed by measuring the ellipticity change in (A) the far-UV region at 214 nm and (B) the near-UV region at 277 nm. Protein concentration, 10 μ M dimer in both cases; buffer, 50 mM sodium acetate, pH 5.6, with 1 mM DTT. Apparent T_m 's were obtained by fitting the data below 75 $^{\circ}$ C to a two-state equation as described in Ruiz-Sanz et al. (1995).

Table 2. Measurements of solvent density and apparent partial specific volume for HPV-16 E2 DNA-binding domain at various concentrations of urea

Urea (M)	Solvent density (ρ_0) (g/mL)	E2 concentration (c_2) (mg/mL)	$\Delta\rho$ for solution (mg/mL)	Density increment ($\partial\rho/\partial c_2$) _{μ}	Apparent partial specific volume (ϕ') (mL/g)
0.0	0.999870	10.122856	2.2612	0.2234	0.7767
1.3	1.020650	10.534048	2.2895	0.2173	0.7668
1.7	1.027055	10.318926	2.2074	0.2139	0.7654
3.0	1.047359	10.826651	2.3229	0.2146	0.7499
5.0	1.077516	11.394192	2.3919	0.2099	0.7332

the effects of pH on its unfolding/dissociation at the physiological range of pH 5.0–8.0 were investigated. The curves obtained from urea unfolding experiments at different pHs were fitted to Equations 5 and 6 (see Materials and methods), and the data are presented in Table 3. A two-state transition is observed at each of the pH values tested.

At pH 5.0 in acetate buffer, the dimer is very unstable, with the consequence that an accurate value of its stability cannot be determined owing to the lack of a folded-state baseline (not shown). At pH 5.6 and 6.1, the protein has a ΔG_{H_2O} value of

around 10 kcal mol⁻¹, which is relatively low when compared to other dimeric proteins of similar size (Neet & Timm, 1994). When the pH was increased to 7 or 8, the stability of the protein was increased substantially, by about 3.0 kcal mol⁻¹, and the K_u value was increased by two orders of magnitude (Table 3). The m values were essentially constant over the pH range tested.

Citrate-phosphate buffer has a particular stabilizing effect on the protein. The ΔG_{H_2O} value of the protein in citrate-phosphate buffer at pH 7 is increased by a further 3 kcal mol⁻¹, taking the K_u to 1.4×10^{-12} M.

Discussion

We have expressed the C-terminal 80-amino-acid dimeric DNA-binding domain of HPV-16 E2 in a soluble folded form and studied its unfolding/dissociation reaction. The structure of the HPV-16 E2 DNA-binding domain has not been determined, but we can use the X-ray crystal structure of BPV-1 E2 DNA-binding domain (Hegde et al., 1992) to interpret our denaturation results, because there is more than 30% sequence identity between the two proteins. The BPV-1 E2 DNA-binding domain represents a new fold for a DNA-binding protein (Hegde et al., 1992). It is a dyad-symmetric eight-stranded antiparallel β barrel made up of two identical "half-barrel" subunits. Each subunit has two α helices that make crossover connections on the outside of the barrel. The better conserved α helix connecting β strands 1 and 2 serves as the recognition helix for DNA binding. The dimeric interface is formed by intersubunit β -sheet hydrogen bonding and by the packing of hydrophobic residues in the center of the barrel.

Unfolding experiments performed on the HPV-16 E2 DNA-binding domain using different spectroscopic techniques show that the protein unfolds in a two-state process. A concomitant loss of secondary and tertiary structures is observed by far-UV CD and fluorescence, respectively. The thermal unfolding also shows a simultaneous disappearance of secondary and tertiary structure, with virtually identical apparent T_m values obtained from far-UV and near-UV CD experiments.

There are obvious differences in the fluorescence spectra of the urea-unfolded state and the acid-unfolded state of the protein (Fig. 2). These differences may arise from a difference in the environment of at least one of the Trp residues in the acid- and urea-unfolded states of the protein, which could mean different conformations, at least locally at the level of the Trp residues. The far-UV CD spectra at high urea concentration and

Fig. 7. Equilibrium sedimentation studies on the urea dissociation of HPV-16 E2 DNA-binding domain. **A:** Plots of $\bar{M}_{w,app}$ against protein concentration at various urea concentrations, chosen to show the range of aggregation observed. Curves were fitted either for nonideality (0.0 and 3.0 M urea) or for a monomer/dimer equilibrium (0.5 and 2.5 M urea). **B:** Plot showing the linear dependence of $\ln(K_d)$ upon urea concentration over the range from 0.5 to 2.7 M urea.

Table 3. Effect of pH on the urea unfolding/dissociation of the HPV-16 E2 DNA-binding domain

	pH	ΔG_{H_2O} (kcal/mol)	m (kcal/mol)	$[U]_{50\%}$ (M)	K_u (M)
Sodium acetate	5.6	9.80 ± 0.29	2.21 ± 0.16	1.53	6.5×10^{-8}
MES	6.1	9.89 ± 0.33	1.82 ± 0.13	1.39	5.5×10^{-8}
Bis-Tris	7.0	12.75 ± 0.38	2.07 ± 0.12	2.61	4.4×10^{-10}
Tris-HCl	8.0	12.74 ± 0.60	1.77 ± 0.17	3.04	4.5×10^{-10}
Citrate-phosphate	7.0	16.18 ± 0.63	1.98 ± 0.14	4.44	1.4×10^{-12}

low pH is also different, supporting the idea of structural differences in the unfolded states of E2.

The midpoint of the urea unfolding transition is dependent upon protein concentration, showing that unfolding is coupled to the dissociation of the dimer. This conclusion is supported by the observation that the ΔG_{H_2O} is constant at the different dimer concentrations used in the urea-unfolding experiments (Grant et al., 1992).

The coincidence of unfolding and dissociation was confirmed by the results of equilibrium sedimentation experiments at different urea concentrations. Equilibrium sedimentation has the advantage over spectroscopic techniques that it determines the association state of the dimer without being affected by the shape or state of unfolding of the protein. The K_d determined in this way differs from K_u in that it represents only the dissociation step of the dimer. The data clearly show that the protein is a dimer at pH 5.6 in the absence of urea, but is a monomer in 3 M urea. The ΔG_{H_2O} value obtained in the sedimentation experiment is $10.2 \text{ kcal mol}^{-1}$, in excellent agreement with the ΔG_{H_2O} value obtained from the urea denaturation experiments.

On the basis of the above observations, we can conclude that the unfolding of the HPV-16 E2 DNA-binding domain takes place by a mechanism in which the dimer is dissociated and unfolded simultaneously without any significant amount of folded monomeric intermediate being formed:

The instability of the isolated monomers is not surprising, because, in the crystal structure of the BPV-1 E2 DNA-binding domain, the dimeric interface is an essential component of the overall β -barrel fold.

Other small DNA-binding proteins also appear to be folded only as dimers. The λ Cro repressor (Anderson et al., 1981) and Arc repressor (Breg et al., 1990) both contain a dimeric interface formed by a region of antiparallel β sheet. Both of them unfold via a strict two-state behavior, where dissociation and unfolding occur in a concerted manner (Bowie & Sauer, 1989; Griko et al., 1992). In these cases, most of the stabilization energy of the protein comes from the antiparallel β sheet dimeric interface, and the dissociated monomeric domain is not stable enough to stand alone as a folded structure. Larger DNA-binding proteins with separate dimerization domains, such as λ repressor (Banik et al., 1992), display a three-state unfolding behavior.

The HPV-16 E2 DNA-binding domain has been shown to bind its specific DNA sequence with a high affinity of about

$8 \times 10^{-11} \text{ M}$ (Sanders & Maitland, 1994). This is very close to our measured K_u (K_d) of the dimer at pH 7. At low protein concentrations, the DNA-binding process will, therefore, be linked to the monomer-dimer equilibrium.

pH has a significant effect on the equilibrium unfolding/dissociation of the HPV-16 E2 DNA-binding domain. The dimer is destabilized by almost $3.0 \text{ kcal mol}^{-1}$ when the pH is decreased from 7.0 to 6.1. Within this pH range, this effect is most likely to be the result of the protonation of a histidine residue. There are five pairs of histidine residues in the dimeric HPV-16 E2 DNA-binding domain. On the basis of the crystal structure of the BPV-1 E2 DNA-binding domain, we can assume that one pair of histidines is buried inside the hydrophobic core. The protonation of these residues would lead to the formation of a pair of buried positive charges that would destabilize the protein.

Citrate-phosphate buffer increases the ΔG_{H_2O} value by another $3.0 \text{ kcal mol}^{-1}$ when compared with bis-Tris buffer at the same pH. This dramatic increase in stability is likely to be the result of the interaction of phosphate anions with the DNA-binding site, a phenomenon frequently observed in proteins that interact with nucleic acids.

In summary, the HPV-16 E2 DNA-binding domain dissociates/unfolds in a two-state manner which is very sensitive to pH and phosphate anions. Our results strongly suggest that the domain is largely monomeric (and therefore unfolded) at the concentrations at which its DNA-binding capacity is normally determined. This conclusion has direct implications for how the folding and DNA-binding mechanisms are coupled in this viral gene-expression regulatory domain.

Materials and methods

Materials

The restriction endonucleases and T4 DNA ligase used in the subcloning of DNA fragments were purchased from USB. The PFU polymerase and buffer used in PCR amplification were from Stratagene. Urea was of analytical grade and purchased from Fisons. Deionized and purified water was used to make up fresh buffer and denaturant before each experiment. Ammonium sulfate was of enzyme grade and was from BDH. All other buffers and chemicals used in purification and denaturation experiments were purchased from Sigma.

Construction of the plasmid for bacterial expression of the E2 DNA-binding domain

A pair of oligonucleotides was used to amplify a 243-bp fragment (nucleotides 3610–3852 of the HPV-16 genome) encoding

the C-terminal DNA-binding domain of HPV-16 E2 (Seedorf et al., 1985) from a clone kindly given by Dr. L. Crawford, Pathology Department, University of Cambridge. A ribosome binding site of the sequence AAAGAGGAGAAATTATCG was incorporated into the gene in front of the start codon. The PCR fragment was treated with *Eco*RI and *Bam*HI and then cloned into a pTz18U-based vector (Pharmacia, Uppsala, Sweden) under the control of the bacteriophage T7 promoter. The resulting plasmid, pTz18U-E2, was transformed into the BL21 (DE3) pLysS *E. coli* strain for expression.

Cells were grown in LB medium to $A_{600} = 0.4$, and the expression was induced by adding IPTG to 1 mM. After overnight incubation at 37 °C, cells were spun down and sonicated in E2 extraction buffer (100 mM HEPES, pH 6.8, 0.6 M NaCl, 5 mM β -mercaptoethanol, 1 mM PMSF, and 20 μ g/mL DNase). The supernatant was subjected to 60% ammonium sulfate precipitation, and the precipitate was resuspended and dialyzed in a medium of 50 mM HEPES, pH 6.8, 0.6 M NaCl, and 1 mM DTT. The dialysate was loaded onto a Heparin Hyper D affinity column (Sepracor) equilibrated with 100 mM HEPES, pH 8.0, 0.6 M NaCl, 1 mM DTT, and 1 mM EDTA. The column was washed with five column volumes of buffer, and the bound protein was eluted with a 0.6–2 M NaCl gradient. Fractions that were >90% pure were pooled and dialyzed in 50 mM sodium acetate buffer, pH 5.6, with 1 mM DTT. The dialysate was subjected to a second 80% ammonium sulfate precipitation. The precipitate was redissolved in a minimum amount of gel filtration buffer (50 mM sodium acetate buffer, pH 5.6, 0.5 M NaCl, and 1 mM DTT) before loading onto a preparative Superdex G75 gel filtration column (Pharmacia). The main peak collected was dialyzed against 50 mM sodium acetate buffer, pH 5.6, with 1 mM DTT, and then concentrated and the buffer exchanged into water using a YM10 Centriprep concentrator (Amicon). The purified protein was stored as a 100 μ M solution at –70 °C after snap freezing in liquid nitrogen. Protein concentration of the dimer (MW = 18.49 kDa) was determined using the Protein assay kit (Bio-Rad). A yield of around 30 mg E2 per liter of culture was obtained.

Fluorescence and CD spectroscopy

Fluorescence spectra and single measurements were recorded on an LS-5B Perkin-Elmer luminescence spectrometer with a slit width of 2.5 nm and an excitation wavelength of 280 nm under the conditions described in the figures. Urea equilibrium denaturation was performed in 50 mM sodium acetate buffer, pH 5.6, and 1 mM DTT unless otherwise stated. Acid denaturation was performed in 100 mM citrate-phosphate buffer at various pH's, from 2.4 to 7.0, with 1 mM DTT. All the solution volumes were measured and delivered using a Hamilton Micro LAB M instrument. The temperature in all experiments was 25 ± 0.1 °C, and the protein/denaturant solutions were pre-equilibrated at 25 °C for 2 h before measurements.

The urea renaturation experiment was carried out by dialyzing a 500 μ M protein solution overnight in buffer containing 4.0 M urea at 4 °C to attain complete unfolding. The unfolded protein was then allowed to refold by diluting into increasingly lower concentrations of urea and incubating at 25 °C for 2 h before measurements.

CD spectra were obtained with a Jasco J720 spectropolarimeter, using a 0.1-cm or 1.0-cm path-length cuvette for far-UV

and near-UV measurements, respectively. Spectra were the average of 10 scans at a 50 nm/min scan speed, and the buffer baselines were subtracted. For urea denaturation experiments, only the far-UV region from 250 to 210 nm was analyzed, owing to increased noise at shorter wavelengths. For the thermal denaturation experiment, a Neslab RTE 110 water bath was used to generate a temperature gradient from 20 to 75 °C at a speed of 50 °C per hour.

Denaturation/dissociation of the E2 DNA-binding domain: Analysis of the data using a two-state model

In a two-state model of denaturation, where only the native dimer and the denatured monomer are in equilibrium ($N_2 \leftrightarrow 2U$), the equilibrium constant for unfolding, K_u , and the free energy of unfolding, ΔG_u , in the presence of denaturant can be represented by the equation

$$K_u = \exp(-\Delta G_u/RT) = [U]^2/[N_2], \quad (1)$$

where $[N_2]$ and $[U]$ are the concentrations of the native dimer and the unfolded monomer, respectively. The total monomeric protein concentration, P_T , equals $2[N_2] + [U]$.

The free energy of unfolding ΔG_u is linearly related to the concentration of urea (Pace, 1986; Tanford, 1968):

$$\Delta G_u = \Delta G_{H_2O} - m[\text{urea}]. \quad (2)$$

The unfolding data obtained from fluorescence or CD measurements can be converted to the fraction of unfolded protein, F_u , using Equation 3 (Greene & Pace, 1974):

$$F_u = (I_N - I_{\text{obs}})/(I_N - I_U) = [U]/(2[N_2] + [U]), \quad (3)$$

where I_{obs} is the observed intensity and I_N and I_U are the intensities of the native and unfolded states, respectively.

Combining Equations 1 and 3 gives the quadratic equation $2F_u^2 P_T + K_u F_u - K_u = 0$, in which F_u is solved to be

$$F_u = (K_u/4P_T) [(1 + 8P_T/K_u)^{1/2} - 1]. \quad (4)$$

By combining Equations 1–4, Equation 5 was obtained. The observed fluorescence and CD intensities were fitted to this equation by using the computer program Kaleidagraph (Abelbeck software).

$$I_{\text{obs}} = I_N - (I_N - I_U) \{ \exp[(m[\text{urea}] - \Delta G_{H_2O})/RT] \} \times \frac{\{1 + 8P_T/\exp[(m[\text{urea}] - \Delta G_{H_2O})/RT]\}^{1/2} - 1}{4P_T}. \quad (5)$$

The denaturant concentration at which half of the protein was unfolded, $[U]_{50\%}$, is dependent on protein concentration in an equilibrium between the native dimer and its unfolded monomer. By substituting $F_u = 1/2$ into Equation 4, $[U]_{50\%}$ can be obtained from the following equation:

$$[U]_{50\%} = [(RT \ln P_T) + \Delta G_{H_2O}]/m. \quad (6)$$

Measurement of apparent partial specific volume

The densities of the buffer and of solutions of E2 protein at dialysis equilibrium in 50 mM sodium acetate, pH 5.6, and different concentrations of urea were measured with a Precision Density Meter (model DMA06; Anton Paar) (Kratky et al., 1973), calibrated using air and water, at 25 °C. The barometric pressure was measured with a mercury barometer with no correction to sea level, as the actual density of air was required. The density increment $[(\partial\rho/\partial c_2)_\mu]$, which was calculated from these measured values and the protein concentration (c_2), was used to calculate the apparent partial specific volume (ϕ') from the equation:

$$\phi' = [1 - (\partial\rho/\partial c_2)_\mu] / \rho_o,$$

where ρ_o is the density of the solvent (Casassa & Eisenberg, 1964).

Analytical ultracentrifugation

Solutions of E2 protein prepared as above were diluted to appropriate A_{280} values. Sedimentation measurements were made using a Beckman An-G analytical rotor in a Beckman L8-70 ultracentrifuge equipped with a Prep UV scanner, interfaced to an Apple III microcomputer, which recorded the absorbance at 280 nm for approximately 400 equally spaced points along the cell during each scan.

Sedimentation equilibrium measurements were made using the long-column method, and the samples were centrifuged to equilibrium at 20,000 rev/min and 25 °C, after overspeeding (van Holde & Baldwin, 1958) for approximately 5 h at 30,000 rev/min. Scans were taken at intervals of approximately 24 h and, when no difference could be detected between successive scans, the later scan was taken to be operationally at equilibrium. The baseline was determined by overspeeding at 45,000 rev/min until all the material had sedimented away from the meniscus, then slowing down to 20,000 rev/min before taking a further scan.

Apparent weight average molecular masses ($\bar{M}_{w,app}$) were determined for successive sets of 11 datum points from the equation (Casassa & Eisenberg, 1964)

$$\bar{M}_{w,app} = \frac{2RT[d \ln(c_2)/dr^2]}{(1 - \phi' \rho_o) \bar{\omega}^2},$$

where r and c_2 are the radius and concentration at each point and ω is the angular velocity. Plots of $\bar{M}_{w,app}$ (expressed as monomers; $M_r = 9,245$) against the concentration (as molarity of monomer) at the midpoint of each set of 11 datum points were made for each data set, to show the dependence of $\bar{M}_{w,app}$ upon protein concentration.

Curves were fitted to the data using proFit (Cherwell Scientific Publishing Ltd.) to fit the equation for a monomer/dimer association equilibrium:

$$\bar{M}_{w,app} = \frac{K_d + 8[M_t] - \sqrt{K_d(K_d + 8[M_t])}}{4[M_t]},$$

where K_d is the dissociation constant and $[M_t]$ is the total protein concentration (expressed as molarity of monomer). The data obtained at 0.0 M urea and at ≥ 3.0 M urea could not be fitted in this way, because the parameters tend to their limits (i.e., the equilibrium consisted of essentially just dimer or monomer, respectively) and these data were fitted for nonideality with the equation

$$1/\bar{M}_{w,app} = 1/M + Bc_2$$

where M is the ideal molecular mass and B the second virial coefficient.

Acknowledgments

This work has been submitted by Y.-K.M. to the Graduate School of the University of Cambridge in partial fulfillment of the requirements for the Ph.D. degree in Physics/Chemistry. Y.-K.M. was supported by a scholarship from the Croucher Foundation. M.B. was supported by a ZENeca/MRC/DTI LINK program.

References

- Anderson WF, Ohlendorf DH, Takeda Y, Matthews BW. 1981. Structure of the cro repressor from bacteriophage λ and its interaction with DNA. *Nature* 290:754–758.
- Banik U, Saha R, Mandal NC, Bhattacharyya B, Roy S. 1992. Multiphasic denaturation of the λ repressor by urea and its implications for the repressor structure. *Eur J Biochem* 206:15–21.
- Barsoum J, Prakash SS, Han P, Androphy EJ. 1992. Mechanism of action of the papillomavirus E2 repressor: Repression in the absence of DNA binding. *J Virol* 66:3941–3945.
- Bowie JU, Sauer RT. 1989. Equilibrium dissociation and unfolding of the Arc repressor dimer. *Biochemistry* 28:7139–7143.
- Breg JN, van Opheusden JHL, Burgering MJM, Boelens R, Kaptein R. 1990. Structure of Arc repressor in solution: Evidence for a family of β -sheet DNA-binding proteins. *Nature* 346:586–589.
- Casassa EF, Eisenberg H. 1964. Thermodynamic analysis of multicomponent solutions. *Adv Protein Chem* 19:287–395.
- Gauthier JM, Dillner J, Yaniv M. 1991. Structural analysis of the human papillomavirus type 16-E2 transactivator with antipeptide antibodies reveals a high mobility region linking the transactivation and the DNA-binding domain. *Nucleic Acids Res* 19:7073–7079.
- Gissmann L. 1992. Papillomaviruses and human oncogenesis. *Curr Opin Genet Dev* 2:97–102.
- Grant SK, Deckman IC, Culp JS, Minnich MD, Brooks IS, Hensley P, Debouck C, Meek TD. 1992. Use of protein unfolding studies to determine the conformational and dimeric stabilities of HIV-1 and SIV proteases. *Biochemistry* 31:9491–9501.
- Greene RF, Pace CN. 1974. Urea and guanidine hydrochloride denaturation of ribonuclease, lysozyme, α -chymotrypsin, and β -lactoglobulin. *J Biol Chem* 249:5388–5393.
- Griko YV, Rogov VV, Privalov PL. 1992. Domains in λ Cro repressor. A calorimetric study. *Biochemistry* 31:12701–12705.
- Ham J, Dostatni N, Gauthier JM, Yaniv M. 1991. The papillomavirus E2 protein: A factor with many talents. *Trends Biochem Sci* 16:440–444.
- Harrison SC. 1991. A structural taxonomy of DNA-binding domains. *Nature* 353:715–719.
- Hegde RS, Grossman SR, Laimins LA, Sigler PB. 1992. Crystal structure at 1.7 Å of the bovine papillomavirus-1 E2 DNA-binding domain bound to its DNA target. *Nature* 359:505–512.
- Howley PM. 1990. Papillomavirinae and their replication. In: Fields BN, Knipe DM, eds. *Virology 2nd ed*. New York: Raven Press, Ltd. pp 1625–1676.
- Johnson WCJ. 1988. Secondary structure of proteins through circular dichroism spectroscopy. *Annu Rev Biophys Biophys Chem* 17:145–166.
- Kratky O, Leopold H, Stabinger H. 1973. The determination of the partial specific volume of proteins by the mechanical oscillator technique. *Methods Enzymol* 27:98–108.
- McBride AA, Romanczuk H, Howley PM. 1991. The papillomavirus E2 regulatory proteins. *J Biol Chem* 266:18411–18414.
- Neet KE, Timm DE. 1994. Conformational stability of dimeric proteins:

- Quantitative studies by equilibrium denaturation. *Protein Sci* 3:2167–2174.
- Pace CN. 1986. Determination and analysis of urea and guanidine hydrochloride denaturation curves. *Methods Enzymol* 131:266–279.
- Ruiz-Sanz J, Prat Gay Gd, Otzen DE, Fersht AR. 1995. Protein fragments as models for events in protein folding pathways: Protein engineering analysis of the association of two complementary fragments of the barley chymotrypsin inhibitor 2 (CI-2). *Biochemistry* 34:1695–1701.
- Sanders CM, Maitland NJ. 1994. Kinetic and equilibrium binding studies of the human papillomavirus type-16 transcription regulatory protein E2 interacting with core enhancer elements. *Nucleic Acids Res* 22:4890–4897.
- Seedorf K, Kraemmer G, Duerst M, Suhai S, Roewekamp WG. 1985. Human papillomavirus type 16 DNA sequence. *Virology* 145:181–185.
- Tanford C. 1968. Protein denaturation. *Adv Prot Chem* 23:121–282.
- van Holde KE, Baldwin RL. 1958. Rapid attainment of sedimentation equilibrium. *J Phys Chem* 62:734–743.