doi:10.1016/j.jmb.2009.03.013 J. Mol. Biol. (2009) 388, 327–344

JMB

Available online at www.sciencedirect.com

ScienceDirect

ELSEVIER

# Indirect DNA Readout on the Protein Side: Coupling between Histidine Protonation, Global Structural Cooperativity, Dynamics, and DNA Binding of the Human Papillomavirus Type 16 E2C Domain

**Tommaso Eliseo[1]†, Ignacio E. Sánchez[2]†, Alejandro D. Nadra[1,2], Mariano Dellarole[2], Maurizio Paci[1], Gonzalo de Prat Gay[2]* and Daniel O. Cicero[1]***

[1]*University of Rome "Tor Vergata" via della Ricerca Scientifica 1, 00133 Rome, Italy*

[2]*Fundación Instituto Leloir and IIBBA-CONICET, Patricias Argentinas 435 (C1405BWE), Buenos Aires, Argentina*

*Received 24 November 2008; received in revised form 6 March 2009; accepted 6 March 2009*
Available online 12 March 2009

DNA sequence recognition by the homodimeric C-terminal domain of the human papillomavirus type 16 E2 protein (E2C) is known to involve both direct readout and DNA-dependent indirect readout mechanisms, while protein-dependent indirect readout has been deduced but not directly observed. We have investigated coupling between specific DNA binding and the dynamics of the unusual E2C fold, using pH as an external variable. Nuclear magnetic resonance and isothermal titration calorimetry show that pH titration of His318 in the complex interface and His288 in the core of the domain is coupled to both binding and the dynamics of the β-barrel core of E2C, with a tradeoff between dimer stability and function. Specific DNA binding is, in turn, coupled to the slow dynamics and amide hydrogen exchange in the entire β-barrel, reaching residues far apart from the DNA recognition elements but not affecting the two helices of each monomer. The changes are largest in the dimerization interface, suggesting that the E2C β-barrel acts as a hinge that regulates the relative position of the DNA recognition helices. In conclusion, the cooperative dynamics of the human papillomavirus type 16 E2C β-barrel is coupled to sequence recognition in a protein-dependent indirect readout mechanism. The patterns of residue substitution in genital papillomaviruses support the importance of the protonation states of His288 and His318 and suggest that protein-dependent indirect readout and histidine pH titration may regulate DNA binding in the cell.



*Keywords:* papillomavirus E2 protein; DNA sequence recognition; indirect readout; histidine; protein dynamics

***Edited by J. E. Ladbury***

## Introduction

The sequence specificity of globular DNA binding proteins arises from their three-dimensional (3D) structure in two ways. First, a number of residues contact the DNA directly by means of correct positioning in 3D space, a mechanism commonly called "direct readout."[1] Histidine side chains often take part in direct readout,[2–6] providing a plausible regulatory mechanism under physiological conditions due to their intrinsic $pK_a$ of 6.3. Second, the dynamics of the globular domain[7–13] and, in particular, the target DNA sequence[1,14–18] modulate binding through indirect effects. This "indirect

*Corresponding authors. E-mail addresses: gpg@leloir.org.ar; cicero@scienze.uniroma2.it.

† T.E. and I.E.S. contributed equally to this work.

Present address: A. D. Nadra, EMBL-CRG Systems Biology Unit, Centre de Regulació Genòmica, Dr. Aiguader 88, 08003 Barcelona, Spain.

Abbreviations used: 3D, three-dimensional; HPV, human papillomavirus; ITC, isothermal titration calorimetry; Mes, 4-morpholineethanesulfonic acid; HSQC, heteronuclear single quantum coherence; NOE, nuclear Overhauser enhancement; HX, hydrogen–deuterium exchange.

readout" can couple DNA binding to protein dimerization and other molecular functions through allosterism.[7,8,12,19,20]

Human papillomaviruses (HPVs) are highly prevalent pathogens of epithelial tropism.[21,22] Mucosal HPV types are the etiological agents of cervical cancer, the second most common cancer in women, and of vaginal, anal, penile, and head and neck cancers.[22,23] Epidemiological studies have identified HPV type 16 as the type with the highest oncogenic potential.[21] The two available prophylactic vaccines[24] are not widely applied in developing countries and cannot cure the millions of people who are already infected. Understanding how HPV proteins function may lead to efficient antiviral compounds.[25] The circular double-stranded DNA HPV genome typically codes for only two late proteins and six early proteins.[26] The early protein E2[27] is a modular polypeptide consisting of an N-terminal transactivation domain, a flexible linker, and a C-terminal domain (E2C) responsible for dimerization and binding of the p53 protein.[28,29] E2C also targets four conserved DNA sites in the regulatory region of the HPV genome,[16,17] leading to the regulation of viral replication and transcription of the early proteins, including the two oncoproteins E6 and E7.[16–18] Fine discrimination between the DNA sequences of the four sites is therefore crucial for the regulation of the viral life cycle.[16–18] Unregulated transcription of E6 and E7 upon integration into the host genome and loss of the E2 gene is associated with cancer.[26]

The native state of the E2C domain is a homodimeric β-barrel consisting of two β–α–β motifs per monomer.[16,17,30–40] The β-strands of the motifs form the conserved eight-stranded β-barrel core of the structure, arranged in two β4–β1–β3–β2 sheets. The interface between the E2C monomers comprises the hydrophobic core of the barrel and of β4–β4′ and β2–β2′ hydrogen bonds. Both α-helices pack against the outside of the barrel, with most residues in helix 1 and Lys349 at the C-terminus of helix 2 being directly involved in DNA binding[16,17,33,35,38] and with most residues in helix 1 and Lys349 at the N-terminus of helix 2 being directly involved in p53 binding.[28,29] Helix 1 is best described as a strained bipartite helix, with a canonical α-helix N-terminal half and a $3_{10}$ helix C-terminal half.[41] The dimeric topology of E2C is unusual, with the DNA binding domain of the Epstein–Barr nuclear antigen 1 being the only known structural analog.[42] In spite of the structural similarity between these two viral DNA binding domains, their amino acid sequences and DNA binding specificities show no detectable homology. It is not known if this unusual topology is related to the p53 or DNA binding functions of E2C.

The E2C domain binds a pseudopalindromic DNA target sequence with the consensus ACCgaaaxcGGT, where capital letters represent strongly conserved bases and small letters represent weakly conserved bases.[16–18] Sequence recognition by E2C is mediated by direct protein–DNA contacts between the ACCg/cGGT hemisites and the highly conserved N294, K297, C298,Y301, and R302 residues from helix 1 in the first β–α–β element of each monomer.[16–18,33,35,38] Direct sequence readout constitutes the bottleneck of complex formation along the two-state kinetic binding route.[43] Direct E2C–DNA contacts also include nonspecific interactions between the DNA backbone of the ACCg/cGGT hemisites and N294, T295, K297, R300, Y301, R302, K304, and K305 in helix1; Ser315, Thr316, and Trp3177 in $\beta_2$-strand; and K349 in helix2.[16–18,33,35,38] Direct E2C–DNA contacts contribute most of the free energy for complex formation.[44] In addition, DNA flexibility can modulate the free energy of binding through indirect readout: the four-base A-rich spacer in the DNA, which does not establish direct contacts with the protein, strongly modulates binding through its ability to bend towards the minor groove.[14–18] We call this phenomenon DNA-dependent indirect readout.

The HPV16 E2C dimeric β-barrel has been proposed to play a role in indirect sequence readout in two ways, mainly from the comparison of the structures and binding specificities of homologous E2C domains. First, its flexibility may regulate the displacement of the DNA binding helices upon binding.[16,17,33,35,38] However, the β-barrel of the HPV16 E2C domain suffers only small changes in structure upon DNA binding,[30,31] in agreement with its overall rigidity in a molecular dynamics simulation relative to its bovine papillomavirus homolog.[45,46] Interestingly, we recently reported that an engineered monomeric version of the HPV16 E2C domain shows the same structure as the dimer, but altered DNA binding properties.[47] This suggested that the dynamics of the β-barrel, rather than the structure, influences DNA binding[47]—a mechanism that has not been investigated to date. Second, the electrostatic potential of the DNA binding side of the β-barrel may also play a role in indirect readout by favoring DNA bending.[16,17,33,35,38] His318 is located in the center of this surface and provides an ideal probe to validate this hypothesis, but its contribution to binding remains uncharacterized.

We present thermodynamic and nuclear magnetic resonance (NMR) evidence for coupling between DNA binding, protonation of His288 and His318, and the global conformational dynamics of E2C.[48] Our results reveal that the β-barrel of the domain is a highly cooperative structural unit that mediates indirect DNA sequence readout. This is also the first report of the E2C–DNA complex dynamics for the HPV type 16 E2 protein, a potential drug target from the most common strain of this oncogenic pathogen.[25]

## Results

### pH dependence of specific E2C–DNA binding

We have characterized the binding of HPV16 E2C to a cognate nucleotide (site 35; sequence GTAACC-GAAATCGGTTGA) using isothermal titration calorimetry (ITC) (Table 1 and Fig. 1a). We carried out the experiments at 25 °C in a three-component buffer

**Table 1.** Specific E2C–DNA binding as a function of pH, measured by ITC in 100 mM Tris, 50 mM Mes, 50 mM sodium acetate, 200 mM NaCl, and 0.2 mM DTT at 25 °C

| pH | $\Delta H$ (kcal/mol) | $K_d$ (nM) | $\Delta G$ (kcal/mol) | $-T\Delta S$ (kcal/mol) | $N$ | $C$ |
|---|---|---|---|---|---|---|
| 4.5 | −4.01±0.04 | 26±4 | −10.40±0.09 | −6.39±0.10 | 0.94 | 360 |
| 5.0 | −6.16±0.10 | 35±9 | −10.21±0.15 | −4.05±0.18 | 0.77 | 234 |
| 5.5 | −10.30±0.06 | 52±5 | −9.98±0.05 | 0.32±0.08 | 0.95 | 206 |
| 6.0 | −14.92±0.08 | 63±5 | −9.86±0.04 | 5.06±0.09 | 0.88 | 141 |
| 6.5 | −16.86±0.12 | 76±7 | −9.75±0.05 | 7.11±0.13 | 0.86 | 112 |
| 7.0 | −18.30±0.11 | 78±5 | −9.64±0.04 | 8.66±0.12 | 0.84 | 110 |
| 7.5 | −11.47±0.10 | 111±11 | −9.52±0.06 | 1.95±0.12 | 0.90 | 75 |
| 8.0 | −10.01±0.09 | 139±12 | −9.39±0.05 | 0.62±0.11 | 0.84 | 61 |
| 8.5 | −8.19±0.08 | 165±16 | −9.29±0.06 | −1.10±0.10 | 0.78 | 46 |

The total ionic strength is 300 mM, independent of pH. The fit of the raw data yields the change in enthalpy upon binding and the dissociation constant. $\Delta G$ is calculated as $RT\ln(K_d)$, and $-T\Delta S$ is calculated as $\Delta G$ minus $\Delta H$. $N$ is the stoichiometry of the reaction, and $C$ is the ratio of the ligand concentration to the dissociation constant.[49]

[100 mM Tris, 50 mM 4-morpholineethanesulfonic acid (Mes), 50 mM sodium acetate, 200 mM NaCl, and 0.2 mM DTT] that ensures a constant ionic strength of 300 mM over the pH range studied.[50] In the pH range between 4.5 and 8.5, the observed stoichiometry is 0.8 or higher, and the ratio of the dissociation constant to the E2C concentration, or $c$ parameter,[49] is above 40 (Table 1), validating the fitting of the data to obtain a dissociation constant for the E2C–DNA interaction.[49] At higher or lower pH values, E2C was not stable in solution, leading to the observation of abnormal stoichiometries (data not shown).

The free energy for specific E2C–DNA binding shows two transitions as a function of pH (Fig. 1a), indicating that at least two functional groups change their p$K_a$ values upon binding.[51] Binding is stronger at lower pH for both transitions, showing that the p$K_a$ of both groups is higher in the complex than in the free reagents.[51] We have fitted the data in Fig. 1a to a model with two independently titrating residues in order to obtain the relevant p$K_a$ values (see Materials and Methods).[51] One of the titrating groups has a p$K_a$ of 4.95±0.08 in the free state and a p$K_a$ of 5.24±0.07 in the complex, while the other one has a p$K_a$ of 7.42±0.09 in the free state and a p$K_a$ of 7.61±0.09 in the complex.

## pH dependence of E2C native-state structure

ITC data show a coupling between the ionization of two E2C residues and the free energy of specific E2C–DNA binding. This coupling may be due to pH titration of residues involved in direct protein–DNA contacts or through modulation of the protein structure and/or dynamics caused by protonation. Circular dichroism (CD) spectroscopy in the far-UV region showed that heparin, sodium chloride, and sodium phosphate can induce changes in the structure of E2C,[52] indicative of a plastic native-state ensemble. DNA binding induces similar changes, showing that conformational plasticity is coupled to binding.[52]

We used far-UV CD to follow the influence of protonation on the secondary structure of E2C at pH values between 5 and 9, in the three-component buffer used in the ITC experiments (Fig. 1b). While E2C shows a far-UV CD spectrum typical of a fully native protein between pH 5 and 9, we could also observe a

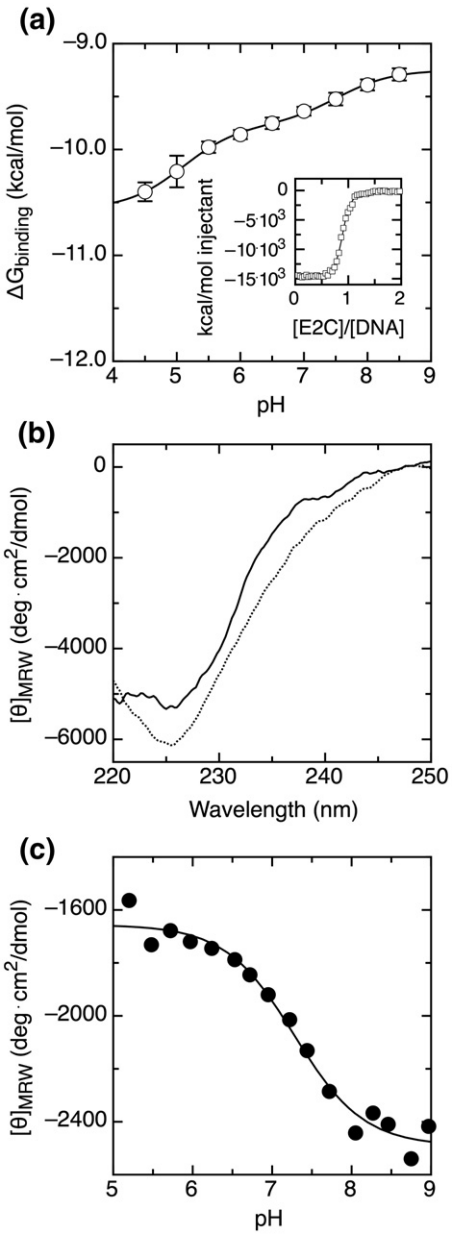


**Fig. 1.** Specific DNA binding and far-UV CD of E2C as a function of pH in 100 mM Tris, 50 mM Mes, 50 mM sodium acetate, 200 mM NaCl, and 0.2 mM DTT at 25 °C. The total ionic strength is 300 mM, independent of pH. (a) Free energy of binding as a function of pH, measured by ITC. Error bars are fitting errors given by the Microcal Origin software (Microcal Software, Northampton, MA, USA). The line is a fit to a model with two independently titrating residues with p$K_a$ values of 4.95±0.08 and 7.42±0.09 in the free protein and p$K_a$ values of 5.24±0.07 and 7.61±0.09 in the DNA-bound protein. Inset: Example of binding isotherm at pH 6.0 resulting from integration of the specific heats with respect to time. (b) Far-UV CD spectra at pH 5 (continuous line) and pH 8 (dotted line). The protein concentration is 10 μM. (c) Molar ellipticity at 235 nm and fit to a single protonation event with a p$K_a$ value of 7.3±0.1.

minor transition in the spectrum between pH 5 (continuous line) and pH 8 (dotted line) (Fig. 1b). We fitted the data at each wavelength using the equation for pH titration of a single chemical group (see Materials and Methods). All wavelengths can be fitted with the same p$K_a$ (data not shown), suggesting that the changes in secondary structure are cooperative and coupled to a single protonation event. Figure 1c shows pH titration data at 235 nm, which can be fitted with a p$K_a$ value of 7.3±0.1. This p$K_a$ for changes in the secondary structure of the domain corresponds to one of the p$K_a$ values for the pH dependence of DNA binding (Fig. 1a).

## Protonation of E2C histidine residues is governed by their molecular environment and coupled to specific DNA binding

Our thermodynamic and spectroscopic results suggest that pH titration of at least two E2C residues between pH 5 and pH 9 with p$K_a$ values of 5.0 and 7.3 is coupled with changes in the free energy for specific DNA binding, and that one of them is coupled with conformational changes in the native state of the domain. In the absence of deeply buried arginine, lysine, aspartate, or glutamate residues in the structure of E2C,[30] the obvious candidates to produce such effects are the five histidine residues in the E2C sequence (Fig. 2a),[53] although we cannot, in principle, exclude the aspartate and glutamate residues for the group with a p$K_a$ of 5.0. We have characterized in detail the molecular environment of the histidine residues of the E2C protein in the free state and in complex with the "site 35" DNA binding site using NMR, in order to identify the residues whose protonation is coupled to changes in protein conformation and binding.

We performed all NMR experiments described in this work at pH values between 5 and 9 and in 50 mM sodium phosphate and 5 mM DTT. The temperatures were 30 °C for the free protein and 45 °C for the E2C–DNA complex. These conditions optimized spectral resolution and protein stability and solubility.[31] On the other hand, ionic strength varies for this buffer in the pH range used, and the p$K_a$ values obtained should be regarded as approximations. First, we recorded a series of $^{1}H$–$^{15}N$ heteronuclear single quantum coherence (HSQC) and long-range $^{1}H$–$^{15}N$ HSQC spectra of free E2C as a function of pH to observe the correlations between the two carbon-attached hydrogens ($H^{\varepsilon 1}$ and $H^{\delta 2}$) and the two nitrogens ($N^{\delta 1}$ and $N^{\varepsilon 2}$) of the imidazole ring. Figure 2b shows sample spectra of E2C (left) and the E2C–DNA complex (right) at pH 7.5. Assignments of the five histidine imidazolic rings have already been described for the free protein.[25] Assignments of the E2C–DNA complex were based on the analysis of $^{1}H$–$^{1}H$ nuclear Overhauser enhancement (NOE) in the $^{13}C$-edited 3D NOE spectroscopy spectrum, using the available assignment of the side chains.[26] The right triangular pattern of cross-peaks in the $^{1}H$–$^{15}N$ long-range HSQC spectrum can be used to determine the composition of the equilibrium between the protonated form and the neutral form of a histidine and which of the two nitrogens is attached to the hydrogen in the neutral form[54] (Table S1). We used the chemical shift for both nuclei of the histidine imidazole ring and for the amide group of the protein backbone to determine the p$K_a$ values for the histidine residues (see Materials and Methods) (Table 2 and Fig. 3). Overall, the free E2C histidine p$K_a$ values presented here are in good agreement with a previous study carried out under different solvent conditions.[53] The imidazole ring of a deprotonated histidine can exist in two tautomeric forms, depending on whether the hydrogen is bound to the $N^{\varepsilon 2}$ or $N^{\delta 1}$ atoms.[55] The relative populations of the $N^{\varepsilon 2}$–H and $N^{\delta 1}$–H tautomers are shown in Table 2. We can combine the tautomer populations and p$K_a$ values to draw a picture of the molecular environment of the five E2C histidines in the two forms of the protein. For a solvent-exposed histidine, the expected p$K_a$ is approximately 6.3, and the populations of the $N^{\varepsilon 2}$–H and $N^{\delta 1}$–H tautomers are expected to be in a 80:20 ratio.[55]

### *His322 and His326*

These two histidine residues are located in the solvent-exposed and flexible β2–β3 loop (Fig. 2a). In the unbound E2C, they both present the canonical p$K_a$ and tautomer ratio, indicating that they are fully exposed to the solvent. The β2–β3 loop is known to remain highly flexible in the DNA complex[31] and forms only long-range contacts with the DNA.[31] The loop is close in space to the negatively charged phosphate backbone,[31] and mutation of residues Lys325 and Lys327 to alanine moderately decreases the affinity for DNA.[44] From these data, we hypothesized that the two chargeable histidine residues could establish electrostatic interactions with the target DNA. However, the chemical shifts, p$K_a$, and tautomer populations of His322 and His326 do not change upon DNA binding (Table 2 and Fig. 3a). Thus, His322 and His326 do not take part in DNA binding and remain solvent exposed in the DNA complex.

### *His306*

This histidine residue is located at the C-terminal half of the DNA binding helix (Fig. 2a). In free E2C, His306 shows a decreased p$K_a$ of 5.7 and a modest decrease in the population of the $N^{\varepsilon 2}$–H tautomer. This indicates that this residue establishes weak interactions with its tertiary environment, which disfavor the $N^{\varepsilon 2}$–H tautomer and proton binding. The p$K_a$ value and tautomer composition of His306 present little changes upon binding (Table 2 and Fig. 3a), indicating that the side chain is not directly involved in binding.

### *His318*

This histidine residue belongs to the solvent-exposed side of $\beta_2$-strand on the DNA binding surface of the domain (Fig. 2a). His318 shows an

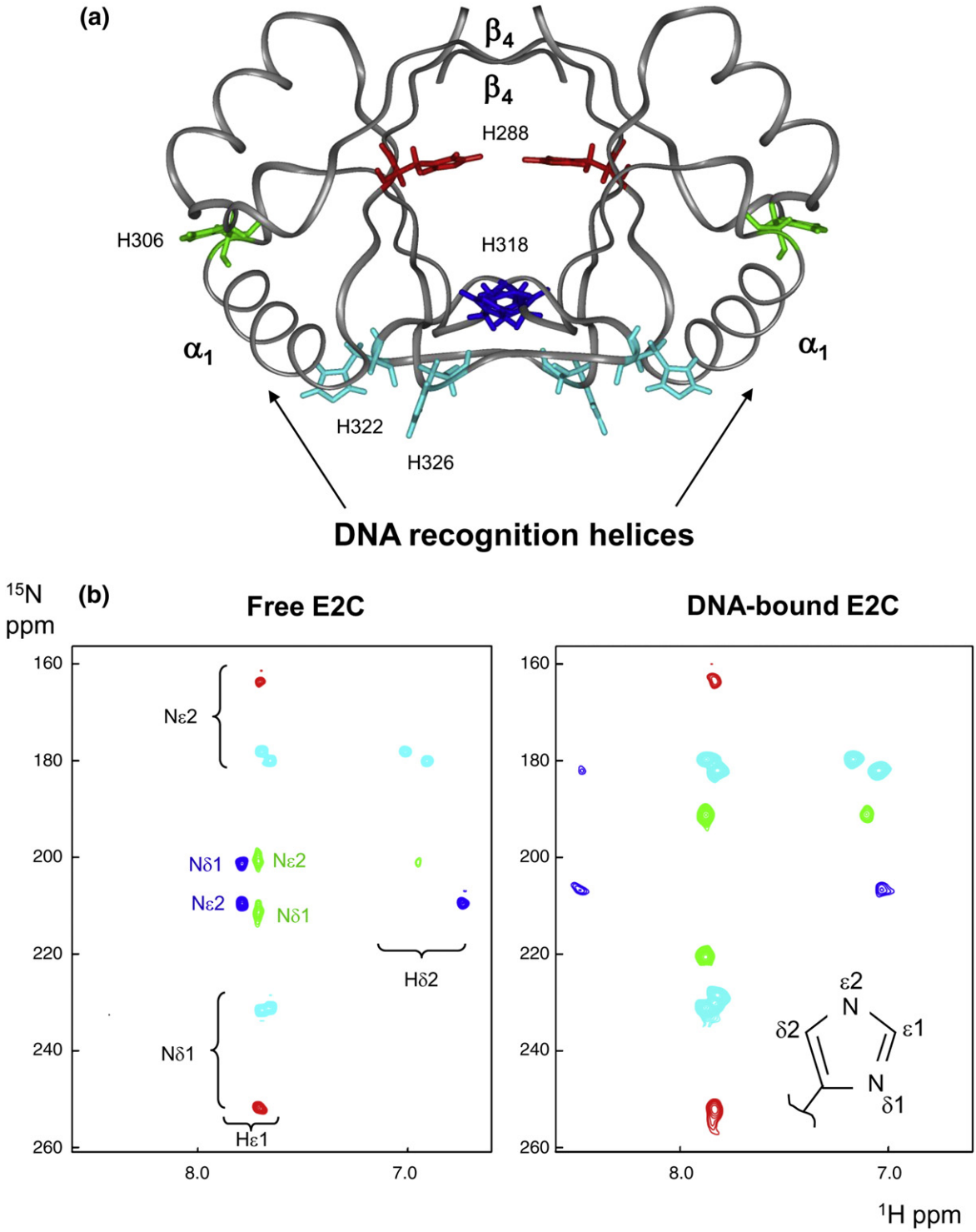


**Fig. 2.** Molecular environment of E2C histidine residues. (a) Representation of the structure of the free E2C homodimer.[30] The backbone of the protein is shown as a ribbon, and the five histidine side chains are shown in all-atom stick representation (His288 in red, His 306 in green, His318 in blue, and His322 and His326 in cyan). We also show the location of the DNA recognition helices. (b) Long-range $^{1}H$–$^{15}N$ HSQC spectra of free and DNA-bound E2C obtained at pH 7.5. The color coding used for the five histidine residues is the same as in (a).

increased p$K_a$ of 6.7 in E2C, implying that the native structure favors protonation. In agreement with this, in the neutral form of His318, the two possible tautomers are almost equally populated, with both nitrogen atoms oscillating rapidly between nonprotonated and protonated states, causing the side chain to resemble a charged histidine. DNA binding shifts the p$K_a$ of His318 to 7.8, the strongest perturbation of all E2C histidines (Table 2 and Fig. 3a). The tautomer population of the neutral His318 in complex with

**Table 2.** E2C residues showing pH-dependent chemical shifts in the pH range 5–9, p$K_a$ values, population of the $N^{\varepsilon 2}$ tautomer for deprotonated histidines, and influence of specific DNA binding

| | p$K_a$ | | $N^{\varepsilon 2}$–H (%) | |
|---|---|---|---|---|
| Residue | Free E2C | E2C–DNA | Free E2C | E2C–DNA |
| His288 | 4–5 | 4–5 | 100 | 100 |
| His306 | 5.74±0.02 | 5.6±0.1 | 56 | 68 |
| His318 | 6.74±0.06 | 7.76±0.04 | 45 | ND |
| His322 | 6.29±0.01 | 6.3±0.1 | 83 | 81 |
| His326 | 6.22±0.04 | 6.2±0.1 | 81 | 79 |

DNA could not be determined due to the shift in p$K_a$. We interpret that His318 interacts with the negatively charged phosphate backbone of the DNA via water molecules as in some[38]—but not all[33,38]—homologous complexes. Thus, we can assign His318 to the residue with a p$K_a$ of ~7.3 identified in the pH dependence of DNA binding measured by ITC at constant ionic strength (Fig. 1).

### *His288*

His288 belongs to $\beta_1$-strand and is buried in the hydrophobic core of the β-barrel, far from the DNA binding surface (Fig. 2a). It presents a sharply decreased p$K_a$ in the range of 4–5 (Fig. 3b). The exact value could not be accurately determined by using the $^1H$–$^{15}N$ cross-peaks due to both line broadening and overall protein instability below pH 5. We interpret that partial protonation of His288 renders a more dynamic E2C, while full protonation leads to global unfolding of the domain. The changes in chemical shift reported for E2C NH groups in the dimeric interface between free and complexed E2C are small.[31] This is in agreement with the small differences in tautomer populations and chemical shifts between the free and DNA-bound forms for the H and N nuclei of His288 (Table 2; Table S1). Both data sets support a small conformational change in the β-barrel dimeric interface upon DNA binding. Similar to the free form of the protein, we could not determine the p$K_a$ of His288 in the E2C–DNA complex due to incomplete pH titration (Fig. 3b), but could estimate that it falls in the pH range 4–5 (Table 2). Interestingly, the curve for $N^{\delta 1}$ shows a double sigmoidal shape that can be fitted with p$K_a$ values of 4.4 and 7.7. Comparison with the results in Table 2 suggests that His288 suffers a change in conformation upon protonation of His318 in the E2C–DNA complex. By comparison with the ITC data, taking into account the changes in buffer and temperature, we can assign His288 to the E2C residue with a p$K_a$ of around 5.0, which influences DNA binding (Fig. 1).

## Coupling between histidine protonation and E2C structure

Next, we determined which regions of E2C change conformation upon protonation of histidine residues. The amide $^{15}N$ chemical shifts of many residues in the free form of the domain change in the pH range from 5

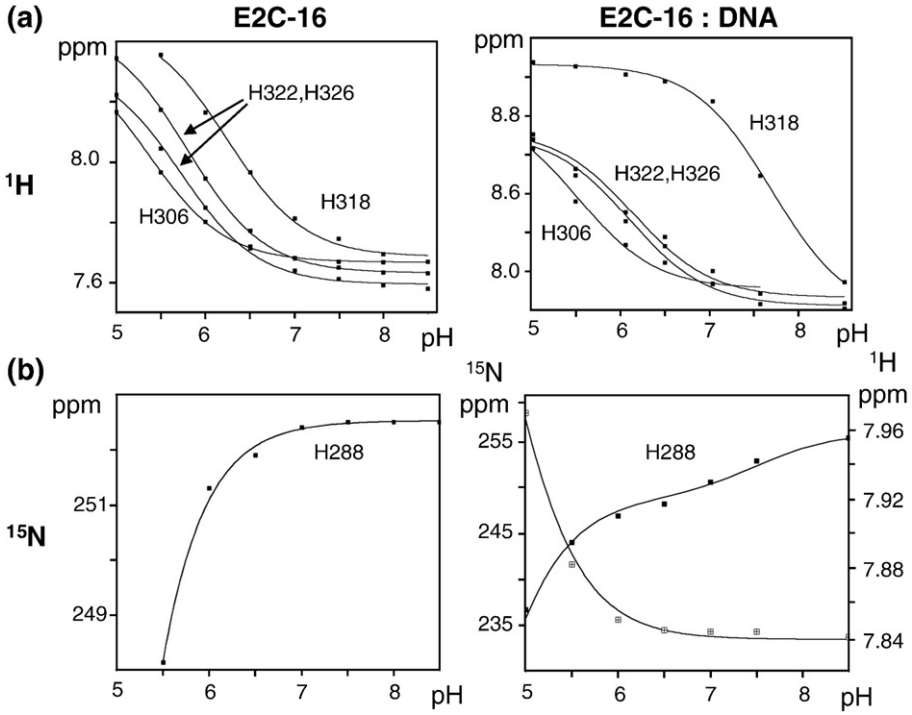


**Fig. 3.** E2C chemical shifts as a function of pH. (a) pH dependence of $H^{\varepsilon 1}$ chemical shifts for His318, His306, His322, and His 326, and fits to single protonation events. Left: Free E2C. Right: DNA-bound E2C. (b) pH dependence of $N^{\varepsilon 2}$ and $H^{\varepsilon 1}$ chemical shifts on His288, and fits to single or double protonation events. Left: Free E2C. Right: DNA-bound E2C.

to 9. This is the result of the pH titration of side chains showing p$K_a$ values in this range that affect indirectly the structure, or more generally the electronic environment, of the nuclei. Fitting the curves of chemical shift *versus* pH (see Materials and Methods) allowed the extraction of the p$K_a$ values, as well as the differences in chemical shift between the structures with a given histidine protonated or nonprotonated (see Fig. S1 for examples). Some of the pH titrations were biphasic, indicating that some $^{15}$N chemical shifts change upon pH titration of more than one residue. In these cases, the data were fitted using a double sigmoidal curve, obtaining two p$K_a$ values and corresponding changes in chemical shift (Table 3). An additional residue, Cys298, titrates within this pH range with a p$K_a$ of 8.8 in free E2C (data not shown; the ionization and redox properties of this residue will be the subjects of future work).

Most E2C backbone amide nitrogens titrate with apparent p$K_a$ values that are almost identical with those determined for one of the five histidine residues (Table 3 and Wetzler *et al.*[41]). We could not measure the influence of His322 and His326 protonation on the structure of the β2–β3 loop because most residues in this region cannot be monitored above pH 5. Only a few residues from the rest of the domain cotitrate with these two histidines (Table 3), indicating that protonation of His322 and His326 does not influence residues that are far away in the structure. Similarly, only residues in the DNA binding helix or close in space cotitrate with His306 (11 residues; average p$K_a$ = 5.6 ± 0.1).

**Table 3.** Cotitrating residues in free E2C in the pH range 5–9 from amide $^{15}$N chemical shifts

| Residue | Secondary structure element | p$K_a$ | Δppm amide $^{15}$N |
|---|---|---|---|
| G291 | β1 | 6.28 ± 0.08 | 0.33 |
| W317 | β2 | 6.3 ± 0.2 | ND |
| T320 | β2 | 6.3 ± 0.2 | ND |
| T284 | β1 | 6.84 ± 0.03 | 0.96 |
| L289 | β1 | 6.51 ± 0.05 | 0.27 |
| K290 | β1 | 6.7 ± 0.1 | 0.26 |
| T311 | β2 | 6.65 ± 0.05 | 0.52 |
| H318 | β2 | 6.73 ± 0.05 | 1.46 |
| W319 | β2 | 6.81 ± 0.05 | 0.93 |
| I330 | β3 | 6.5 ± 0.1 | 0.14 |
| T332 | β3 | 6.73 ± 0.05 | 0.65 |
| T334 | β3 | 6.72 ± 0.09 | 0.28 |
| S337 | β3 | 6.7 ± 0.2 | 0.18 |
| S357 | β4 | 6.5 ± 0.1 | 0.24 |
| F360 | β4 | 6.76 ± 0.03 | 0.44 |
| M361 | β4 | 6.70 ± 0.07 | 0.95 |
| S362 | β4 | 6.72 ± 0.01 | 0.62 |
| I363 | β4 | 6.5 ± 0.2 | 0.86 |
| T284[a] | β1 | 4.5 ± 0.2 | 0.91 |
| L289 | β1 | 4.5 ± 0.2 | 0.22 |
| K290 | β1 | 4.7 ± 0.2 | 0.78 |
| H318 | β2 | 4.5 ± 1.0 | 1.21 |
| W319 | β2 | 4.5 ± 0.2 | 1.45 |
| I330 | β3 | 4.5 ± 0.2 | 0.43 |
| T332 | β3 | 4.7 ± 0.2 | 0.41 |
| D336[a] | β3 | 4.5 ± 1.0 | 1.15 |
| S337[a] | β3 | 4.5 ± 1.0 | 0.74 |
| S357 | β4 | 4.7 ± 0.2 | 0.83 |
| F360[a] | β4 | 4.5 ± 0.2 | 0.69 |
| M361 | β4 | 4.5 ± 0.2 | 0.57 |

The data for His306 have been published elsewhere.[41]

[a] Residues that are close to D336 and can, to some extent, be perturbed by aspartate titration.

The backbone amide nitrogen of His318 cotitrates with that of 15 residues, with an average p$K_a$ value of 6.7 ± 0.1, in agreement with the data presented in Table 2. In contrast with other histidines, His318 protonation leads to $^{15}$N chemical shift changes in residues belonging to the four strands of the dimeric β-barrel (Fig. 4 and Table 3). This explains the change in secondary structure observed by CD and assigned to this side chain by comparison with the ITC data. Interestingly, pH titration of His318 extends its effect far away from its position in the structure, including residues 360–363 in $\beta_4$-strand on the opposite side of the barrel that are, on average, 16.5 Å apart ($C^\alpha$–$C^\alpha$ distance) and residues Ala329, Val331, and Thr332, which are 4.9, 7.6, and 10.4 Å apart, respectively. Most of the cotitrating residues are involved in contacts between the two monomers. Comparison with the ITC and CD data (Fig. 1) indicates that protonation of His318 is coupled with both specific DNA binding and a change in the secondary structure of the domain, which we can now assign to the β-barrel.

The backbone amide nitrogen of His288 cotitrates with that of 12 residues, with an average p$K_a$ value of 4.6 ± 0.1 (Table 3). This value is within the range anticipated from the tautomer populations and chemical shifts of this side chain (Fig. 3b and Table 2). The value of the p$K_a$ is also compatible with the protonation of an acidic side chain of an aspartate (the p$K_a$ of a solvent-exposed residue is 3.8) or glutamate (4.5) of the protein. Inspection of the observed shifts, in fact, reveals a large change in chemical shift for the amide $^{15}$N of Asp336 (1.15 ppm), a possible indication that this perturbation is the consequence of the pH titration of its own side chain. Other residues that show their side chains close to this aspartate are Thr284, Ser337, and Phe360; so, it is possible that the observed chemical shift perturbation is related, to some extent, to the protonation of Asp336 (Fig. 4). On the other hand, eight residues belonging to the dimeric β-barrel and the dimerization interface do not show any close acidic residue, and their perturbation can be assigned to the protonation of His288 (Fig. 4 and Table 3). The perturbation reaches the $\beta_2$-strand in the DNA binding surface, providing an explanation for the coupling between His288 protonation and DNA binding. Altogether, the perturbation of amide $^{15}$N chemical shifts by protonation of E2C histidine side chains confirms the key role of His288 and His318 protonation in E2C native-state structure. Protonation of His288 and His318 leads to changes in the dimeric β-barrel conformation, regardless of the site of perturbation. The α-helices of E2C, including the DNA binding helix, are not affected by protonation of these two histidines, as is the β-barrel. pH titration of His288 and His318 is also coupled to specific DNA binding, strongly suggesting a linkage between DNA binding and the global structure and dynamics of the domain.

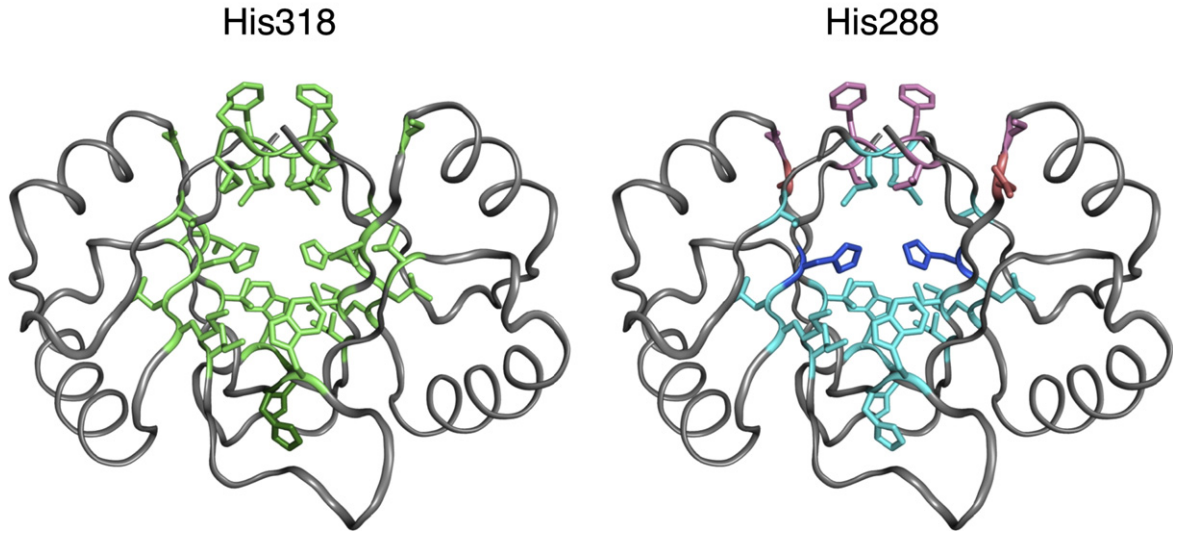


**Fig. 4.** Representation of cotitrating residues in the structure of the free E2C homodimer.[30] The backbone of the protein is shown as a ribbon, and the side chains of the cotitrating residues are shown in all-atom stick representation. Left: His318 and cotitrating residues. Right: His288 and cotitrating residues. Asp336 is shown in red, and Thr284, Ser337, and Phe360 are shown in pink.

### Coupling between histidine protonation, specific E2C–DNA binding, and the global conformational dynamics of E2C

The pH dependence of the backbone amide $^{15}N$ chemical shifts shows that the β-barrel of E2C undergoes a small but cooperative change in conformation upon protonation of His318 and His288. We will now examine the effect of histidine protonation on the dynamics of the domain using the pH dependence of the cross-peaks in the $^{1}H$–$^{15}N$ HSQC spectrum (Fig. 5). A remarkable effect of pH is observed also for the intensity of NH cross-peaks. At pH 6.5–7.0, almost all residues give rise to a visible signal in the HSQC spectrum,[30] while certain peaks disappear at higher pH values (8.0–8.5) and other peaks disappear or exhibit a very reduced intensity at lower pH values (5.0–5.5). The HSQC spectrum of the DNA-free HPV16 E2C protein recorded at 30 °C and pH 5.0 is shown in Fig. S2 of Supplementary Material, where the corresponding HN cross-peak assignment is reported. The assignment at pH 5.0 was confirmed by analysis of triple-resonance experiments (HNCO, HNCA, and $^{15}N$-edited NOE spectroscopy). This assignment confirms that those peaks that disappear at acidic pH values do not appear in other regions of the spectrum. Relative intensities of cross-peaks in the HSQC are determined by the relaxation properties of the two nuclei. Fast internal flexibility would cause an increment of the intensity as the signal becomes sharper. On the contrary, both chemical exchange and conformational exchange (in the microsecond-to-millisecond timescale) will diminish it, as the signal becomes broader. In the case of amide HN nuclei, exchange with the solvent, which occurs by an acid-catalyzed reaction below pH 4 or by a base-catalyzed reaction at higher pH,[56] is expected to decrease the cross-peak intensity of solvent-exposed residues at pH values higher than 7.5. This is the case for Gly353 (Fig. 5a), located in the solvent-exposed α2–β4 loop. The intensity of a cross-peak may also decrease because of the appearance of protein motions in the microsecond-to-millisecond timescale following the protonation of a side chain, as for Gly359 (Fig. 5a).

Figure 5c shows the logarithm of the ratio between the peak intensities at pH 5 and pH 9 (filled circles) for all characterized E2C amides in the free form of the domain. In this plot, residues exchanging with the solvent present a positive value for $\log(I_{5.0}/I_{9.0})$. Among the regions with regular secondary structure, helix α2 and strands β1 and β3 show little exchange of the amide proton with water. Strands β2 and β4, both closing the dimeric β-barrel, and the recognition helix α1 show a high degree of water exchange. Last, the exchange effect is strongest for the solvent-exposed β2–β3 loop, leading to the disappearance of most HN cross-peaks of these residues in the $^{1}H$–$^{15}N$ HSQC at pH values above 5. On the other hand, a number of residues show a large decrease in intensity upon going to acidic pH (Fig. 5c). The residues are spread over the four strands of the barrel: Val287, Leu 289, and Gly291 in β1; Ser314 in β2; Val331 and Thr332 in β3; and Gly359 and Ile363 in β4. We conclude that protonation of His288 and His318 not only induces conformational changes in the β-barrel but also increases its dynamics in the microsecond-to-millisecond timescale.

We have also measured the pH dependence of the cross-peaks in the $^{1}H$–$^{15}N$ HSQC spectrum for DNA-bound E2C (Fig. 5c, empty squares). In contrast with the free E2C, there are no cross-peaks that decrease in intensity at acidic pH values. One example is Gly359, which was not detected at pH 5 in the free state (Fig. 5a) but did not change its intensity as a function of pH in the E2C–DNA complex (Fig. 5b). This behavior extends to the abovementioned residues of the barrel

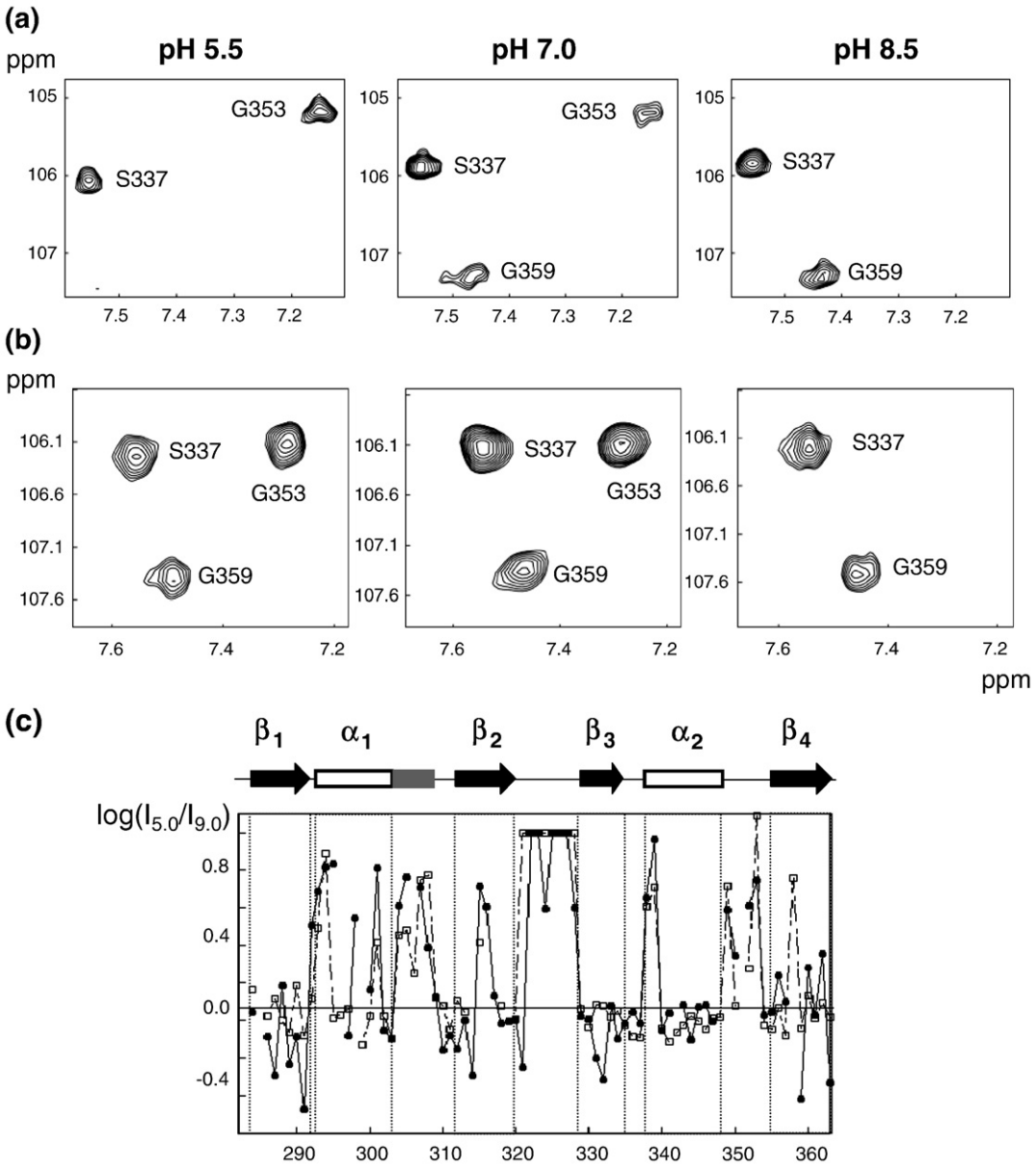


**Fig. 5.** $^{1}H$–$^{15}N$ HSQC spectrum of E2C as a function of pH in the range 5–9. (a) Example cross-peaks for free E2C. (b) Example cross-peaks for DNA-bound E2C. (c) Logarithm of the cross-peak intensity at pH 5 and pH 9, $\log(I_{5.0}/I_{9.0})$, for the free (black circles) and DNA-bound (empty squares) forms of the domain. β-Strands are represented as arrows, α-helices are represented as white rectangles, and $3_{10}$ helices are represented as gray rectangles.

strands, as observed in Fig. 5c. We interpret that DNA binding quenches the microsecond-to-millisecond dynamics of the β-barrel triggered by protonation of His288 and His318.

We have further investigated the coupling between specific DNA binding and E2C dynamics in the picosecond-to-nanosecond and microsecond-to-millisecond timescales by measuring $^{15}N$ relaxation data at pH 6.5 for E2C in the free and DNA-bound forms. Although the free domain becomes less dynamic at this pH value than at lower pH values, this condition ensures the measurement of relaxation times for a higher number of $^{15}N$ nuclei. Backbone $^{15}N$ $T_1$, $T_2$, and steady-state $^{15}N$-[$^{1}H$] NOE were measured at a $^{1}H$ resonance frequency of 700 MHz (Fig. S3 in Supplementary Material). We were able to obtain relaxation parameters for 63 of 80 residues at 30 °C for the DNA-free protein. On the other hand, only 48 of 80 accurate relaxation rates could be measured for the DNA-complexed form at the same temperature, while 72 values could be measured at 45 °C (Fig. S3, central panel). However, all cross-peaks detectable at 45 °C in the $^{1}H$–$^{15}N$ HSQC were also visible in the spectrum recorded at 30 °C, although they exhibited a somewhat broader linewidth (see Fig. S4 in Supplementary Material). The very small difference in chemical shifts shows that no significant changes occur in the E2C–DNA protein when the temperature is raised from 30 to 45 °C. Most of the missing residues belong to the β2–β3 loop. In fact, only two residues of

this loop were visible in the $^{1}H$–$^{15}N$ HSQC spectra for the free form (Val324 and Ser328), and three residues of this loop were visible in the $^{1}H$–$^{15}N$ HSQC spectra for the E2C–DNA complex (Val324, Ser328, and Lys325). In addition, all $T_2$ values measured at 30 °C for the DNA-bound form are low (26.1 ms, on average), except for the central flexible β2–β3 loop. These values can be translated into a rotational correlation time $\tau_c$ of 26 ns, estimated using an isotropic model of rotational motion for the complex. The assumption of isotropic motion is confidently valid if one considers the DNA-bound conformation of E2C.[31,43] The observed $\tau_c$ value is not compatible with the conditions employed to study the system and corresponds to a molecular size much higher than that of the E2C–DNA complex (28 kDa). When relaxation data are collected at 45 °C, the resulting $\tau_c$ obtained from relaxation analysis is 18.6 ns, which leads to a viscosity-corrected value of 13.2 ns at 30 °C. This value is compatible with the DNA–E2C molecular size and indicates that relaxation data measured at 30 °C are affected by factors other than the global tumbling and internal dynamics (aggregation). These observations suggested that experiments be carried out at a higher temperature in the case of the complex without compromising conclusions about the protein internal dynamics.

We analyzed the $^{15}N$ relaxation data using the reduced spectral density approach, which makes no assumption about the nature of molecular global rotational diffusion.[57] In its reduced version,[58,59] relaxation rates are related to motions of the N–H bond at 0, $\omega_N$, and $0.87\omega_H$ frequencies. The corresponding spectral densities at these three frequencies $J$

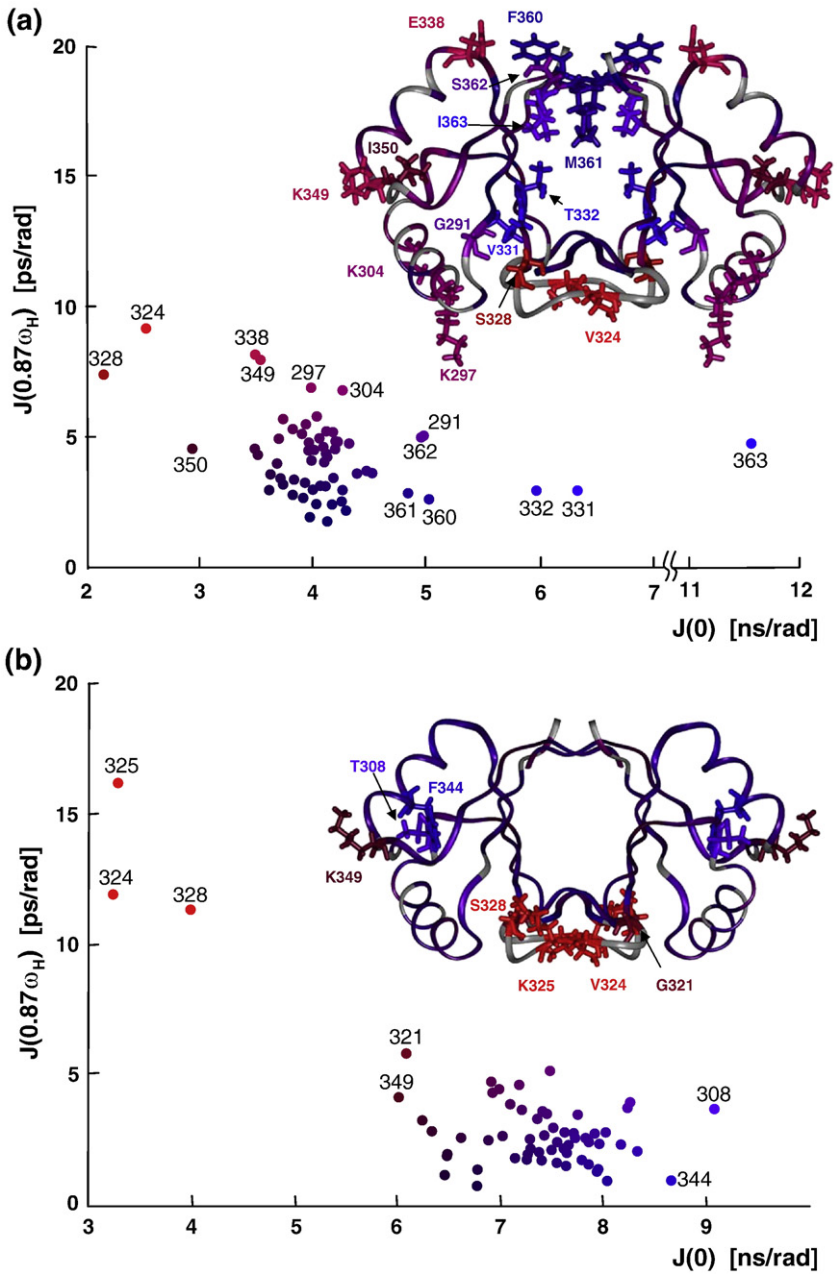


**Fig. 6.** $^{15}N$ relaxation dynamics of E2C using the reduced spectral density approach, shown in a $J(0.87\omega_H)$-*versus*-$J(0)$ plot and represented on the structure of the domain. To make visualization easier, we have used color coding combining red and blue for $J(0.87\omega_H)$ and $J(0)$, respectively. (a) Free E2C. (b) DNA-bound E2C.

(ω) can be calculated from relaxation times $T_1$ and $T_2$ and the heteronuclear NOE using standard equations (see Materials and Methods). We can use a plot of $J(0.87\omega_H)$ *versus* $J(0)$ to detect different contributions from both fast and slow motions (Fig. 6). Higher-than-average $J(0)$ values may reflect contributions from chemical or conformational exchange, whereas $J(0.87\omega_H)$ is insensitive to these motions. On the other hand, high values of $J(0.87\omega_H)$ and low values of $J(0)$ are indicative of fast local motions. Due to these relationships between internal motions and spectral density function values, it is expected that regions with flexibility in the fast timescale (picosecond-to-nanosecond timescale) will appear in the upper-left regions of the plot, whereas residues showing microsecond-to-millisecond flexibility will show up to the right of the average values. To make visualization easier, we have used color coding combining red (for $J(0.87\omega_H)$) and blue (for $J(0)$) (Fig. 6).

In free E2C, we detect fast motions for the β2–β3 loop and at the beginning and end of the α2-helix (Fig. 6a). This includes Lys297, Lys304, and Lys349, all of which play a role in the interaction with DNA,[31] and Glu338 in the p53 binding region. Slow motions in free E2C causing shortened $T_2$ relaxation times are concentrated in the β-barrel and, in particular, involve residues forming the dimeric interface such as Thr332, Met361, and Ile363 (Fig. 6a). In DNA-bound E2C, the dispersion of values in the $J(0.87\omega_H)$-*versus*-$J(0)$ plot is clearly smaller than that for the free protein (Fig. 6), indicating quenching of several internal motions. Upon DNA binding, most fast motions are attenuated, including the residues in the recognition helix (Glu338 and Lys349). The β2–β3 loop shows significant flexibility, confirming that this loop does not acquire a rigid structure by contacting the spacer region of the DNA.[31] DNA binding also leads to a drastic reduction of slow motions in the microsecond-to-millisecond timescale for the β-barrel (Fig. 6b), in agreement with our measurements as a function of pH (Fig. 5c). This fact can be quantified by the reduction in standard deviation in the $J(0)$ values calculated for the E2C–DNA complex (0.71 ns/rad) with respect to that for the free protein (1.14 ns/rad). In summary, we found evidence for fast motions in the E2C–DNA binding surface and slow motions in the β-barrel, both of which disappear largely upon DNA binding.

It is important to note that the dynamic studies of the E2C–DNA complex were conducted at a temperature (45 °C) higher than and at a pH (5.5) lower than those of the free E2C protein (30 °C; 6.8). Although the higher temperature used in the complex studies would enhance the conformational flexibility of the protein, we found a more rigid structure, reinforcing the observation that DNA binding quenches internal motions. In addition, we have observed that the free E2C protein is more flexible at pH 5.5, as different signals disappear in the $^{1}H$–$^{15}N$ HSQC due to conformational exchange in the microsecond-to-millisecond timescale. At this pH, the DNA-bound form does not show any short $^{15}N$ $T_2$. In fact, $T_2$ values measured both at 30 °C and at 45 °C for the DNA-complexed protein (Fig. S3, central panel) show that transverse relaxation rates are not significantly shortened for residues experimenting slow dynamics in the free state, neither for residues making direct contacts with the DNA bases. These observations give a second indication of the diminished flexibility of the protein caused by DNA binding.

## Very slow dynamics of E2C as measured by solvent exchange

We have also probed the dynamics of E2C in the hours timescale by monitoring hydrogen–deuterium exchange (HX) of backbone amides. We dissolved the protein in $D_2O$ at pH∗ 6.5 and 30 °C and collected $^{1}H$–$^{15}N$ HSQC spectra after 0.5, 2, 4, 12, 24, and 48 h. The cross-peaks that remain detectable after each time

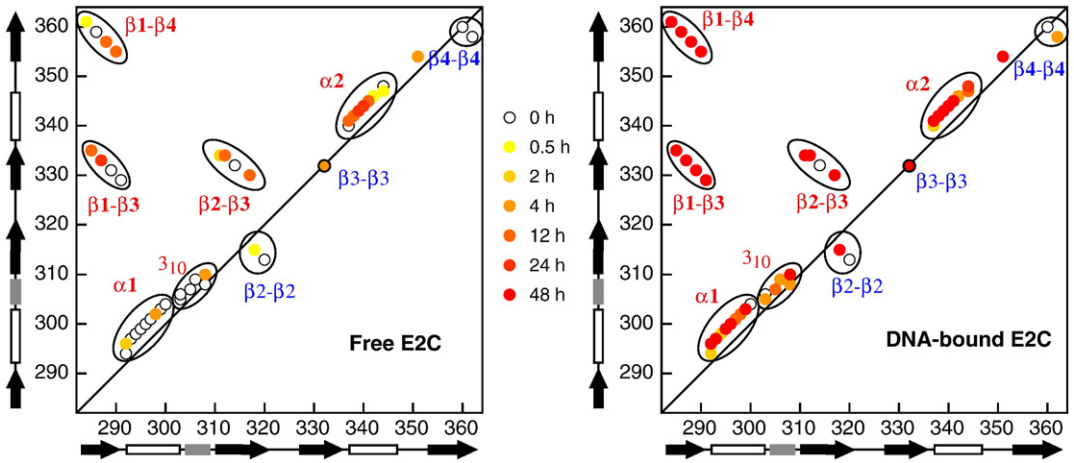


**Fig. 7.** HX of E2C backbone amide hydrogens. Left: Free E2C. Right: DNA-bound E2C. The protected amides at each time point and their hydrogen bond partners in the native homodimer are shown as a contact map. For clarity, contacts are grouped according to interacting secondary structure elements. Tertiary and quaternary contacts are in red and blue fonts, respectively. β-Strands are represented as arrows, α-helices are represented as white rectangles, and $3_{10}$ helices are represented as gray rectangles.

are shown in Fig. 7 as a contact map that highlights both the backbone hydrogen bond donor and acceptor residues and the regular secondary structures. Points above the diagonal correspond to tertiary intramonomer contacts (red), and points below the diagonal correspond to quaternary intermonomer contacts (blue). The HX kinetics are commonly on the hours timescale and report on the conformer population-weighted solvent accessibility of each backbone amide hydrogen. In free E2C, HX may take place from the fully folded conformation, from regions undergoing local fluctuations, from cooperatively unfolded secondary structure elements, and from the fully unfolded domain.[60] HX in the complex may happen not only from the residual population ensemble of free E2C but also from the ensemble of DNA-bound E2C conformations, which may include nonspecific complexes and molecules in which one or more secondary structure elements unfold.[10] Given this potential complexity, we interpret the E2C HX kinetics as an ensemble-averaged measure of the domain dynamics and its changes upon binding.

For free E2C, only six NH cross-peaks are detectable after 24 h (Fig. 7, left), indicating that this structure has an unusually high solvent accessibility compared to most proteins of similar size and stability. The fastest exchange is observed for the DNA binding helix and the β2–β2 and β4–β4 quaternary interactions that hold the barrel together, in agreement with the data in Fig. 5. On the other side, the β4–β1, β1–β3, and β3–β2 hydrogen bonds between strands that hold together

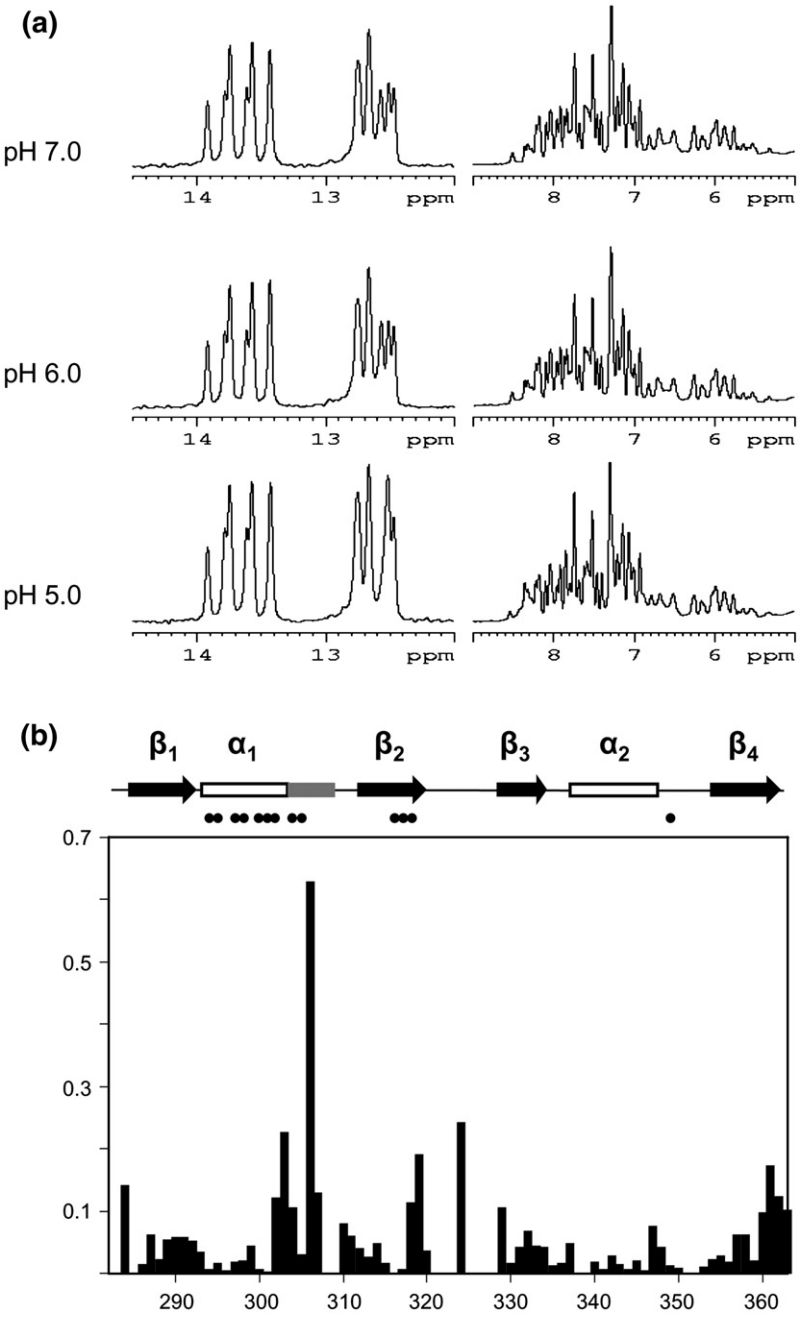


**Fig. 8.** (a) Regions of the $^{1}$H spectra of the 18-bp double-stranded DNA corresponding to site 35 at different pH values. (b) Chemical shift differences for the HN group for the E2C–DNA complex between resonances observed at pH 7.0 and pH 5.0. Secondary structure elements are represented as in Fig. 5. Residues close to the DNA are evidenced by black circles.

the half β-barrel of each monomer and helix α2 show a higher degree of protection and may be considered the "slowly exchanging core" of the E2C domain.

The situation is largely different for the E2C–DNA complex (Fig. 7, right). Forty-eight hours after transferring the sample to $D_2O$, most of the backbone hydrogens belonging to regions with defined secondary structures are still visible. Quaternary interactions still show less protection than tertiary hydrogen bonds, and the $3_{10}$ segment of the DNA recognition helix exchanges faster than the α-helix segment, as expected from its malleability.[41] The general slowdown of exchange is in agreement with the reduction of motions in the DNA binding helix and the β-barrel observed upon DNA binding (Figs. 5c and 6b).

### Free DNA and E2C–DNA conformations are unperturbed by pH changes

The results presented point towards a large change in free E2C conformation and dynamics upon protonation of two key histidine residues. However, changes in the protonation state and/or conformational properties of free DNA or the protein–DNA complex between pH 7.0 and pH 5.0 can also contribute to the observed increase in affinity. We have used chemical shift as a sensitive probe for this type of changes in the free DNA and the E2C–DNA complex. Figure 8a shows a series of $^1H$ spectrum of the 18-bp double-stranded DNA corresponding to the site 35 used in this study. No significant chemical shift changes were detected for the imino proton resonances of T and G bases that are present at very low fields due to the double-helix formation (Fig. 8a, left spectra). In addition, no remarkable changes were detected in the region between 5.0 and 9.0 ppm, where $NH_2$ groups of A, C, and G, as well as aromatic protons of the bases, are observed (Fig. 8a, right spectra). This analysis shows that a change of pH in this range does not perturb DNA conformation significantly.

Next, we explored the possible difference in conformation for the E2C–DNA complex formed at pH 7.0 and pH 5.0. Figure 8b shows the HN chemical shift changes between these two pH values. Large shifts are circumscribed to the region close to His306, which, according to our p$K_a$ measurements, changes its protonation state between 7.0 and 5.0. However, the effect of this change on histidine protonation is only local.[41] On the contrary, residues that contact the DNA (evidenced by black circles in the figure) show an almost identical chemical shift, indicating that the protein–DNA interface conformation is invariant between pH 5.0 and neutral pH.

These two observations leave the changes in protein conformation and dynamics as the only significant contributions to the increase in DNA affinity going towards acidic pH values.

## Discussion

This work is the first report of data on the dynamics and amide hydrogen exchange of an E2C–DNA complex. Additionally, we have studied the coupling between DNA binding and histidine protonation. In other protein–DNA binding reactions, histidine residues play a role in binding through a direct readout mechanism.[2–6] If this were the case for E2C, the extent of coupling between protonation and binding should correlate with proximity to the DNA in the complex, with His306 in the DNA binding helix and His318 in β2 being the most important, with His322 and His326 from the flexible β2–β3 loop playing a secondary role, and with His288 in the core of the β-barrel being decoupled from function. Surprisingly, we found that coupling is strongest for His288 and His318, weak for His306, and absent for His322 and His326. This hierarchy does not correlate with direct protein–DNA contacts in the complex, but with the molecular environment of each histidine in the free state of E2C (Figs. 2 and 3, and Table 2). His288 and His318, which deviate most from a canonical solvent-exposed side chain, are the ones with the strongest influence on binding. His288 exerts its influence in spite of being far from the binding surface, and His318 contacts the DNA directly, but its protonation is clearly coupled with the structure and dynamics of E2C through long-range effects (Figs. 1 and 3–7). Thus, the dynamics of the network of residue–residue contacts in the E2C domain is coupled to DNA sequence recognition. This result confirms for HPV16 E2C the "protein-dependent indirect readout" that was previously deduced from structural considerations and thermodynamic measurements on homologous E2C–DNA complexes[16,17,33,35,38] and observed in other systems.[7–13]

Interestingly, the different conservation levels of the five E2C histidines in human genital papillomaviruses (Bose *et al.*[53] and Fig. S5) correlate with the degree of coupling between DNA binding and side-chain protonation. Histidines 288 and 318 are conserved in most human genital papillomavirus types, His306 is present in around half of the viral types, and histidines 322 and 326 appear only in a minority of types. Moreover, naturally occurring substitutions for His288 and His318 (Fig. S5) parallel the dominant protonation and tautomeric states of the side chains in free HPV16 E2C (Table 2). The only other residue observed at position 288 is a glutamine, which is the closest mimic possible for the $N^{\varepsilon 2}$–H tautomer of a neutral histidine, thanks to the ε amide nitrogen. Similarly, in nature, His318 is only substituted by arginine, which is structurally similar to a protonated histidine due to the neutral $N^{\varepsilon}$ and the partially charged $N^{\gamma}$. We interpret that the conservation/substitution patterns observed in nature support the functional significance of our *in vitro* p$K_a$ values and tautomer populations and suggest that histidine protonation may actually play a regulatory role *in vivo*.

The long-range effects of histidine protonation on E2C structure make the β-barrel of the domain more dynamic upon lowering of the pH from 9 to 5 (Fig. 4), which in turn leads to an 8-fold increase in the association constant for the E2C–DNA complex (Fig. 1). One possibility is that, at higher pH values, the

protein must change conformation to bind its target DNA. This conformational transition, even if subtle,[31] would impose a free-energy penalty. In this model, loosening the structure of free E2C at more acidic pH lowers the free-energy cost of the rearrangement and leads to a higher affinity in a "protein-dependent indirect readout" mechanism. It will be interesting to test whether the same mechanism modulates the kinetics of binding along one or both of the parallel reaction pathways.[61] Interestingly, the increase in the dynamics at lower pH values is quenched upon DNA binding (Figs. 5 and 6). This implies that the increase in E2C dynamics at lower pH values is correlated with the change in conformation upon binding, since an uncorrelated loosening of the structure would lead to a higher entropic cost for binding and a lower—not higher—affinity. From the point of view of E2C stability, upon going from pH 5.6 to pH 8, the unfolding/dissociation constant at equilibrium for the E2C homodimer increases by about 200-fold,[62] unfolding becomes 5-fold faster, and the rate-limiting refolding reactions become about 100-fold slower.[63] Thus, there is a tradeoff between function and the folding thermodynamics and kinetics of the globular E2C domain, similar to other DNA binding proteins.[12,64,65] This tradeoff limits how far the dynamics of E2C can be modulated for function without becoming a partially or fully disordered protein that folds upon binding[66–68] and suggests why the domain evolved a stable but highly dynamic native state as opposed to a maximally stable structure.

The HPV16 E2C β-barrel responds to histidine protonation in a highly concerted manner, with protonation of His288 and His318 at different locations leading to effects on both the structure (Fig. 5) and the dynamics of the domain, as measured by amide $^{15}N$ relaxation, pH dependence of the cross-peaks in the $^{1}H$–$^{15}N$ HSQC spectrum, and HX (Figs. 5–7). Moreover, His318 and DNA binding clearly influence sites far away from the binding surface. The behavior of His318 shows that two of the previously proposed mechanisms for protein-dependent indirect sequence readout, namely, barrel flexibility[16,17,33,35,38] and the electrostatic potential of the DNA binding surface,[16,17,33,35,38] are coupled and need to be considered simultaneously. The changes in chemical shift upon histidine pH titration are small, as observed for DNA binding,[31] and point at an overall robust yet dynamic structure of the barrel. Our view is that the quaternary interface between strands β4–β4′ and β2–β2′ is pivotal in the malleability of the β-barrel by pH and may modulate DNA binding. In fact, the dynamics of the barrel in the microsecond-to-millisecond (Fig. 6) and hours (Fig. 7) timescales shows that the $\beta_4$-strands are the most dynamic in the barrel, followed by the $\beta_2$-strands. This observation of a "breathing" interface connecting two more stable monomers may be related to the monomeric intermediate observed in the refolding kinetics[63,69] and the amyloid route of the domain.[70] In agreement with our view, quenching of β-barrel breathing is the largest change in the dynamics of the HPV16 E2C domain upon DNA binding (Figs. 6 and 7; see also Falconi *et al.*[45,46]).

The dynamics of the HPV16 E2C domain show a remarkable partition into a few cooperative units. The most prominent feature is that changes in the highly cooperative β-barrel do not propagate to the β2–β3 loop or either helix (Fig. 4 and Table 3). Additionally, the β2–β3 loop and the DNA binding helix show only local changes upon protonation of His306, His322, and His326 (Table 3 and Wetzler *et al.*[41]), showing that they are independent cooperative units dominated by local interactions. These data are in agreement with the view of the E2C β-barrel as a sensitive hinge that modulates the relative position of the p53 and DNA binding helices, while the β2–β3 loop behaves independently.[16,17] We would like to add that such a role for the β-barrel requires that the two monomers behave as rigid flaps and, in particular, that the DNA binding helices are firmly anchored to the β-sheet of each monomer.[41] Indeed, the lack of coupling between the barrel and the helices corroborates the rigidity of the monomer "flaps." It is known that the global dynamics of a protein are mainly determined by the overall arrangement of secondary structure elements within its tertiary and quaternary structures.[71–73] We propose that this topology–activity relationship is one of the reasons behind the rare fold of the DNA-binding domain of the E2 HPV master regulator and its homodimeric nature.

Finally, we would like to propose a hypothetical role for selective histidine protonation under physiological conditions. In the cell, the HPV16 E2 protein has to search for its target sequences in the context of the viral and host genomes, and it is very likely that the E2C domain forms a low-affinity complex with nontarget DNA for at least part of the search time. His318 would face the negative charges of the DNA backbone in such a complex and become protonated, similar to the shift in $pK_a$ that we have observed for the specific complex (Fig. 1 and Table 2). Protonation of the His318 side chain and—due to the cooperative nature of the two effects—partial protonation of His288 make the E2C β-barrel more flexible, facilitating the fit of the recognition helices into the DNA major grooves. Thus, we can understand His318 as a sensor that detects the presence of DNA, triggers His288 protonation, and uses some of the stability of the protein to induce a conformation that is better suited for strong binding (Fig. 1). Other functions of the E2 protein that may be influenced by histidine protonation and the coupled changes in dynamics include oligomerization,[70,74] E1 binding, and recruitment of transcription machinery.[27]

## Materials and Methods

The HPV16 E2C domain was expressed and purified as described before.[14,47] Purity after chromatography in a heparin Hyper D column and in a size-exclusion column is higher than 95%. We determined E2C concentration by UV

spectroscopy, with a molar extinction coefficient of 41,290 $M^{-1}$ $cm^{-1}$ at 280 nm. The 18-mer single-stranded DNA oligonucleotide of sequence GTAACC-GAAATCGGTTGA, incorporating the HPV16 E2 binding site 35 (reference), and its complementary strand were purchased from MWG Biotech, Inc., (Germany). 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-site 35, and no detectable single-stranded oligonucleotide was present as judged by PAGE (data not shown).

### ITC and CD measurements

The three-component buffer (100 mM Tris, 50 mM Mes, 50 mM sodium acetate, 200 mM NaCl, and 0.2 mM DTT) was prepared by dissolving appropriate weights of each component in water. DTT was added from a 1 M stock solution. The resulting solution had a pH of 8.3–8.5, which was taken to the desired value with concentrated NaCl or NaOH and used within 12 h. pH was measured in an Orion 420A pH meter (Orion Research, Inc., USA). We used VP-ITC from Microcal for ITC measurements and the proprietary Microcal Origin software for fitting of the data. We titrated 10 μM E2C in the cell with several injections of 100 μM DNA. The injection volume was 5 μl, except for a first injection of 2 μl. We continued the injections beyond saturation to determine the heat of ligand dilution, which was subtracted from the data prior to fitting with a single-site model. We have fitted the data to a model with two independently titrating residues:[51]

$$\Delta G_{\text{binding}}(\text{pH}) = \Delta G_{\text{binding}}(\text{pH 14}) + 2RT\ln\left(\frac{1+10^{\text{p}K^{\text{A}}_{\text{free}}-\text{pH}}}{1+10^{\text{p}K^{\text{A}}_{\text{complex}}-\text{pH}}}\right) + 2RT\ln\left(\frac{1+10^{\text{p}K^{\text{B}}_{\text{free}}-\text{pH}}}{1+10^{\text{p}K^{\text{B}}_{\text{complex}}-\text{pH}}}\right) \quad (1)$$

where A and B are the two titrating groups and the factor 2 comes from the homodimeric nature of E2C. We carried out CD measurements on a Jasco J-810 spectropolarimeter using a Peltier temperature-controlled sample holder and a 0.1-cm pathlength cuvette We recorded and averaged five scans for each condition, with a bandwidth of 2 nm and a data point of 0.2 nm each. All data were buffer subtracted. No smoothing procedure was applied. We fitted the data at each wavelength to a model with a single titrating residue:[51]

$$\theta_{\text{MRW}} = \frac{\theta^{\text{deprotonated}}_{\text{MRW}} + \theta^{\text{protonated}}_{\text{MRW}} \times 10^{(\text{p}K-\text{pH})}}{1+10^{(\text{p}K-\text{pH})}} \quad (2)$$

### NMR spectroscopy

NMR experiments were all performed at 30 °C (free protein) or 45 °C (DNA complex) on a Bruker Avance 700 spectrometer equipped with a triple-resonance probe incorporating self-shielded gradient coils with uniformly $^{15}$N-labeled E2C. The buffer was composed of 50 mM sodium phosphate and 5 mM DTT, prepared as described above for the three-component buffer. Pulsed field gradients were appropriately employed to achieve suppression of the solvent signal and spectral artifacts. Quadrature detection in the indirectly detected dimensions was obtained using the States-TPPI method. The NMR data were processed on Silicon Graphics workstations using NMRPipe[75] and analyzed using NMRView.[76] Linear prediction and apodization 90°-shifted squared sine-bell functions were typically applied before Fourier transformation. The final concentration of the complex for NMR studies was 0.6 mM.

### Histidine and amide pH titration

To investigate the tautomeric and protonation states of E2C and E2C–DNA histidine residues, a series of $^{1}$H–$^{15}$N long-range HSQC spectra was acquired at different pH values (from 4.5, 5.0, 5.5, 6.0, 6.5, 7.5, 8.0, 8.5, and 9.0). For this experiment, the delay during which $^{1}$H–$^{15}$N antiphase is produced was set to 22 ms to refocus single-bond correlations. The observed chemical shifts of $H^{\varepsilon 1}$ or $N^{\varepsilon 2}$ were plotted as a function of pH and fitted to the following equation[77] in order to evaluate p$K_a$ values:

$$\delta_{\text{obs}} = \delta_{\text{d}} + \frac{\delta_{\text{p}} - \delta_{\text{d}}}{1+10^{(\text{pH}-\text{p}K_a)}} \quad (3)$$

where $\delta_p$ and $\delta_d$ are the chemical shifts of the protonated and deprotonated forms, respectively. In cases where a double exponential statistically fitted the data better, the following equation was used:

$$\delta_{\text{obs}} = \delta_{\text{d}} + A\left(\frac{10^{\text{p}K_{a1}-\text{pH}}}{1+10^{\text{p}K_{a1}-\text{pH}}}\right) + B\left(\frac{10^{\text{p}K_{a2}-\text{pH}}}{1+10^{\text{p}K_{a2}-\text{pH}}}\right) \quad (4)$$

where A and B are the respective numerical weights of the two pH titrations and represent the Δδ contribution of each transition, and p$K_{a1}$ and p$K_{a2}$ are the corresponding p$K_a$ values for the two titrating groups.

These equations were used to analyze the chemical shift change on the amide $^{15}$N nuclei, and the data are reported in Table 2. Regular $^{1}$H–$^{15}$N HSQC spectra were acquired at the same pH values reported for the histidine analysis.

Curve fits were performed using KaleidaGraph (Synergy Software) and ProFit (Quantumsoft).

### Relaxation data and backbone dynamics analysis

Relaxation experiments were carried out at 30 °C on a Bruker Avance 700 spectrometer. Measurements of $^{15}$N $T_1$, $T_2$, and $^{1}$H–$^{15}$N NOE were made by performing established $^{1}$H-detected pulse schemes[78,79] in an interleaved manner. Six points were acquired using delays of 16, 240, 465, 689, 913, and 1123 ms for $T_1$ at 40.5 MHz, and delays of 14, 210, 420, 700, 1191, and 1542 ms for $T_1$ at 70.9 MHz. Six points were acquired with delays of 8.3, 33.1, 82.2, 99.4, 132.5, and 165.6 ms for $T_2$ at 40.5 MHz, and delays of 8.2, 24.5, 40.8, 57.1, 73.4, and 97.9 ms for $T_2$ at 70.9 MHz. Integrated cross-peak volumes of nonoverlapped resonances were fitted to two-parameter monoexponential decays. The uncertainties of cross-peak intensities were evaluated as the standard deviation of spectral noise measured in a region free of cross-peaks. The heteronuclear NOE values were determined by the ratio of cross-peak volumes of spectra recorded with and without $^{1}$H saturation, employing a net relaxation delay of 5 s for each scan in both experiments.

Analysis of $^{15}$N relaxation data of E2C protein in its DNA-free and DNA-bound forms was performed using reduced spectral density mapping analysis.[58,59] The spectral densities at the three relevant frequencies 0, $\omega_N$, and $0.87\omega_H$ can be obtained from relaxation times

$T_1$, $T_2$, and heteronuclear NOE using the following equations:

$$\sigma_{NH} = (1/T_1)(\mathrm{NOE} - 1)\gamma_N/\gamma_N \quad (5)$$

$$J(0) = (6/T_2 - 3/T_1 - 2.72\sigma_{NH})/(3d^2 + 4c^2) \quad (6)$$

$$J(\omega_N) = (4/T_1 - 5\sigma_{NH})/(3d^2 + 4c^2) \quad (7)$$

$$J(0.87\omega_H) = (4\sigma_{NH})/(5d^2) \quad (8)$$

where $d = (\mu_0 h\gamma_N\gamma_H/(8\pi^2))/\langle r_{NH}^3\rangle$; $c = \omega_N(\sigma_{||} - \sigma_\perp)/3^{1/2}$; $\mu_0$ is the permeability of free space; $h$ is Planck's constant; $\gamma_H$ and $\gamma_N$ are the gyromagnetic ratios of $^1$H and $^{15}$N, respectively; $r_{NH} = 1.02$ Å is the average amide bond length; and $(\sigma_{||} - \sigma_\perp) = -160$ ppm is the chemical shift anisotropy for $^{15}$N nuclei.

### Amide exchange studies

A series of $^1$H–$^{15}$N HSQC was acquired at 0.5, 2, 4, 12, and 24 h (midpoints of the experiments) after dissolving the lyophilized free protein in (or transferring the complex to) $D_2O$ to monitor the disappearance of signals of amide-exchangeable protons.

## Acknowledgements

We acknowledge funding from Welcome Trust grant GR077355AYA and Agencia Nacional de Promoción Científica y Tecnológica PICT (2000 01-08959); doctoral fellowship from Consejo Nacional de Investigaciones Científicas y Técnicas (CONICET) to A.D.N. and M.D.; and MUTIS postdoctoral fellowship from the Agencia Española de Cooperación Internacional to I.E.S. G.d.P.G. is a career investigator from Consejo Nacional de Investigaciones Científicas y Técnicas.

## Supplementary Data

Supplementary data associated with this article can be found, in the online version, at doi:10.1016/j.jmb.2009.03.013

## References

1. Lesser, D. R., Kurpiewski, M. R. & Jen-Jacobson, L. (1990). The energetic basis of specificity in the EcoRI endonuclease–DNA interaction. *Science*, **250**, 776–786.
2. Luscombe, N. M., Laskowski, R. A. & Thornton, J. M. (2001). Amino acid–base interactions: a three-dimensional analysis of protein–DNA interactions at an atomic level. *Nucleic Acids Res.* **29**, 2860–2874.
3. Lundback, T., van Den Berg, S. & Hard, T. (2000). Sequence-specific DNA binding by the glucocorticoid receptor DNA-binding domain is linked to a salt-dependent histidine protonation. *Biochemistry*, **39**, 8909–8916.
4. Deleeuw, L., Tchernatynskaia, A. V. & Lane, A. N. (2008). Thermodynamics and specificity of the Mbp1–DNA interaction. *Biochemistry*, **47**, 6378–6385.
5. Bradshaw, E. M., Sanford, D. G., Luo, X., Sudmeier, J. L., Gurard-Levin, Z. A., Bullock, P. A. & Bachovchin, W. W. (2004). T antigen origin-binding domain of simian virus 40: determinants of specific DNA binding. *Biochemistry*, **43**, 6928–6936.
6. Bergqvist, S., Williams, M. A., O'Brien, R. & Ladbury, J. E. (2004). Heat capacity effects of water molecules and ions at a protein–DNA interface. *J. Mol. Biol.* **336**, 829–842.
7. Holmbeck, S. M., Dyson, H. J. & Wright, P. E. (1998). DNA-induced conformational changes are the basis for cooperative dimerization by the DNA binding domain of the retinoid X receptor. *J. Mol. Biol.* **284**, 533–539.
8. Wikstrom, A., Berglund, H., Hambraeus, C., van den Berg, S. & Hard, T. (1999). Conformational dynamics and molecular recognition: backbone dynamics of the estrogen receptor DNA-binding domain. *J. Mol. Biol.* **289**, 963–979.
9. Zhu, L., Hu, J., Lin, D., Whitson, R., Itakura, K. & Chen, Y. (2001). Dynamics of the Mrf-2 DNA-binding domain free and in complex with DNA. *Biochemistry*, **40**, 9142–9150.
10. Kalodimos, C. G., Boelens, R. & Kaptein, R. (2002). A residue-specific view of the association and dissociation pathway in protein–DNA recognition. *Nat. Struct. Biol.* **9**, 193–197.
11. Jarymowycz, V. A. & Stone, M. J. (2006). Fast time scale dynamics of protein backbones: NMR relaxation methods, applications, and functional consequences. *Chem. Rev.* **106**, 1624–1671.
12. Okamura, H., Makino, K. & Nishimura, Y. (2007). NMR dynamics distinguish between hard and soft hydrophobic cores in the DNA-binding domain of PhoB and demonstrate different roles of the cores in binding to DNA. *J. Mol. Biol.* **367**, 1093–1117.
13. Rippin, T. M., Freund, S. M., Veprintsev, D. B. & Fersht, A. R. (2002). Recognition of DNA by p53 core domain and location of intermolecular contacts of cooperative binding. *J. Mol. Biol.* **319**, 351–358.
14. Ferreiro, D. U., Lima, L. M., Nadra, A. D., Alonso, L. G., Goldbaum, F. A. & de Prat Gay, G. (2000). Distinctive cognate sequence discrimination, bound DNA conformation, and binding modes in the E2 C-terminal domains from prototype human and bovine papillomaviruses. *Biochemistry*, **39**, 14692–14701.
15. Zhang, Y., Xi, Z., Hegde, R. S., Shakked, Z. & Crothers, D. M. (2004). Predicting indirect readout effects in protein–DNA interactions. *Proc. Natl Acad. Sci. USA*, **101**, 8337–8341.
16. Hegde, R. S. (2002). The papillomavirus E2 proteins: structure, function, and biology. *Annu. Rev. Biophys. Biomol. Struct.* **31**, 343–360.
17. de Prat Gay, G., Gaston, K. & Cicero, D. O. (2008). The papillomavirus E2 DNA binding domain. *Front. Biosci.* **13**, 6006–6021.
18. Sanchez, I. E., Dellarole, M., Gaston, K. & de Prat Gay, G. (2008). Comprehensive comparison of the interaction of the E2 master regulator with its cognate target DNA sites in 73 human papillomavirus types by sequence statistics. *Nucleic Acids Res.* **36**, 756–769.
19. Lefstin, J. A. & Yamamoto, K. R. (1998). Allosteric effects of DNA on transcriptional regulators. *Nature*, **392**, 885–888.

20. Laine, O., Streaker, E. D., Nabavi, M., Fenselau, C. C. & Beckett, D. (2008). Allosteric signaling in the biotin repressor occurs via local folding coupled to global dampening of protein dynamics. *J. Mol. Biol.* **381**, 89–101.
21. Munoz, N., Bosch, F. X., de Sanjose, S., Herrero, R., Castellsague, X., Shah, K. V. *et al.* (2003). Epidemiologic classification of human papillomavirus types associated with cervical cancer. *N. Engl. J. Med.* **348**, 518–527.
22. Bosch, F. X., Sanjosé, S., Castellsagué, X., Moreno, V. & Muñoz, N. (2006). Epimediology of human papillomavirus infections and associations with cervical cancer: new opportunities for prevention. In *Papillomavirus Research: From Natural History to Vaccines and Beyond* (Saveria Campo, M., ed.), *Papillomavirus Research: From Natural History to Vaccines and Beyond*, chapt. 3, pp. 19–40Caisreir Academic Press, Wymondham.
23. D'Souza, G., Kreimer, A. R., Viscidi, R., Pawlita, M., Fakhry, C., Koch, W. M. *et al.* (2007). Case–control study of human papillomavirus and oropharyngeal cancer. *N. Engl. J. Med.* **356**, 1944–1956.
24. Cohen, J. (2005). Public health. High hopes and dilemmas for a cervical cancer vaccine. *Science*, **308**, 618–621.
25. Hegde, R. S. (2006). Papillomavirus proteins and their potential as drug design targets. *Future Med.* **1**, 795–810.
26. Kalantari, M. & Bernard, H. -U. (2006). Gene expression of papillomaviruses. In *Papillomavirus Research: From Natural History to Vaccines and Beyond* (Saveria Campo, M., ed.), *Papillomavirus Research: From Natural History to Vaccines and Beyond*, chapt. 4, pp. 41—52Caisreir Academic Press, Wymondham.
27. McBride, A. A., Romanczuk, H. & Howley, P. M. (1991). The papillomavirus E2 regulatory proteins. *J. Biol. Chem.* **266**, 18411–18414.
28. Parish, J. L., Kowalczyk, A., Chen, H. T., Roeder, G. E., Sessions, R., Buckle, M. & Gaston, K. (2006). E2 proteins from high- and low-risk human papillomavirus types differ in their ability to bind p53 and induce apoptotic cell death. *J. Virol.* **80**, 4580–4590.
29. Brown, C., Kowalczyk, A. M., Taylor, E. R., Morgan, I. M. & Gaston, K. (2008). p53 represses human papillomavirus type 16 DNA replication via the viral E2 protein. *Virol. J.* **5**, 5.
30. Nadra, A. D., Eliseo, T., Mok, Y. K., Almeida, C. L., Bycroft, M., Paci, M. *et al.* (2004). Solution structure of the HPV-16 E2 DNA binding domain, a transcriptional regulator with a dimeric beta-barrel fold. *J. Biomol. NMR*, **30**, 211–214.
31. Cicero, D. O., Nadra, A. D., Eliseo, T., Dellarole, M., Paci, M. & de Prat Gay, G. (2006). Structural and thermodynamic basis for the enhanced transcriptional control by the human papillomavirus strain-16 E2 protein. *Biochemistry*, **45**, 6551–6560.
32. Hegde, R. S. & Androphy, E. J. (1998). Crystal structure of the E2 DNA-binding domain from human papillomavirus type 16: implications for its DNA binding-site selection mechanism. *J. Mol. Biol.* **284**, 1479–1489.
33. Hegde, R. S., Grossman, S. R., Laimins, L. A. & Sigler, P. B. (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.
34. Hegde, R. S., Wang, A. F., Kim, S. S. & Schapira, M. (1998). Subunit rearrangement accompanies sequence-specific DNA binding by the bovine papillomavirus-1 E2 protein. *J. Mol. Biol.* **276**, 797–808.
35. Kim, S. S., Tam, J. K., Wang, A. F. & Hegde, R. S. (2000). The structural basis of DNA target discrimination by papillomavirus E2 proteins. *J. Biol. Chem.* **275**, 31245–31254.
36. Liang, H., Petros, A. M., Meadows, R. P., Yoon, H. S., Egan, D. A., Walter, K. *et al.* (1996). Solution structure of the DNA-binding domain of a human papillomavirus E2 protein: evidence for flexible DNA-binding regions. *Biochemistry*, **35**, 2095–2103.
37. Bussiere, D. E., Kong, X., Egan, D. A., Walter, K., Holzman, T. F., Lindh, F. *et al.* (1998). Structure of the E2 DNA-binding domain from human papillomavirus serotype 31 at 2.4 Å. *Acta Crystallogr. Sect. D*, **54**, 1367–1376.
38. Hooley, E., Fairweather, V., Clarke, A. R., Gaston, K. & Brady, R. L. (2006). The recognition of local DNA conformation by the human papillomavirus type 6 E2 protein. *Nucleic Acids Res.* **34**, 3897–3908.
39. Dell, G., Wilkinson, K. W., Tranter, R., Parish, J., Leo Brady, R. & Gaston, K. (2003). Comparison of the structure and DNA-binding properties of the E2 proteins from an oncogenic and a non-oncogenic human papillomavirus. *J. Mol. Biol.* **334**, 979–991.
40. Veeraraghavan, S., Mello, C. C., Androphy, E. J. & Baleja, J. D. (1999). Structural correlates for enhanced stability in the E2 DNA-binding domain from bovine papillomavirus. *Biochemistry*, **38**, 16115–16124.
41. Wetzler, D. E., Gallo, M., Melis, R., Tommaso, E., Nadra, A. D., Ferreiro, D. U., et al. (2009). A strained DNA binding helix is conserved for site recognition, folding nucleation, and conformational modulation. *Biopolymers*, in press. doi:10.1002/bip.21146.
42. Bochkarev, A., Barwell, J. A., Pfuetzner, R. A., Bochkareva, E., Frappier, L. & Edwards, A. M. (1996). Crystal structure of the DNA-binding domain of the Epstein–Barr virus origin-binding protein, EBNA1, bound to DNA. *Cell*, **84**, 791–800.
43. Ferreiro, D. U., Sanchez, I. E. & de Prat Gay, G. (2008). The transition state for protein–DNA recognition. *Proc. Natl Acad. Sci. USA*, **105**, 10797–10802.
44. Ferreiro, D. U., Dellarole, M., Nadra, A. D. & de Prat Gay, G. (2005). Free energy contributions to direct readout of a DNA sequence. *J. Biol. Chem.* **280**, 32480–32484.
45. Falconi, M., Santolamazza, A., Eliseo, T., de Prat Gay, G., Cicero, D. O. & Desideri, A. (2007). Molecular dynamics of the DNA-binding domain of the papillomavirus E2 transcriptional regulator uncover differential properties for DNA target accommodation. *FEBS J.* **274**, 2385–2395.
46. Falconi, M., Oteri, F., Eliseo, T., Cicero, D. O. & Desideri, A. (2008). MD simulations of papillomavirus DNA–E2 protein complexes hints at a protein structural code for DNA deformation. *Biophys. J.* **95**, 1108–1117.
47. Dellarole, M., Sanchez, I. E., Freire, E. & de Prat Gay, G. (2007). Increased stability and DNA site discrimination of "single chain" variants of the dimeric beta-barrel DNA binding domain of the human papillomavirus E2 transcriptional regulator. *Biochemistry*, **46**, 12441–12450.
48. Whitten, S. T., Garcia-Moreno, E. B. & Hilser, V. J. (2005). Local conformational fluctuations can modulate the coupling between proton binding and global structural transitions in proteins. *Proc. Natl Acad. Sci. USA*, **102**, 4282–4287.
49. Wiseman, T., Williston, S., Brandts, J. F. & Lin, L. N. (1989). Rapid measurement of binding constants and heats of binding using a new titration calorimeter. *Anal. Biochem.* **179**, 131–137.

50. Ellis, K. J. & Morrison, J. F. (1982). Buffers of constant ionic strength for studying pH-dependent processes. *Methods Enzymol.* **87**, 405–426.
51. Baker, B. M. & Murphy, K. P. (1996). Evaluation of linked protonation effects in protein binding reactions using isothermal titration calorimetry. *Biophys. J.* **71**, 2049–2055.
52. Lima, L. M. & de Prat Gay, G. (1997). Conformational changes and stabilization induced by ligand binding in the DNA-binding domain of the E2 protein from human papillomavirus. *J. Biol. Chem.* **272**, 19295–19303.
53. Bose, K., Yoder, N. C., Kumar, K. & Baleja, J. D. (2007). The role of conserved histidines in the structure and stability of human papillomavirus type 16 E2 DNA-binding domain. *Biochemistry*, **46**, 1402–1411.
54. Sudmeier, J. L., Ash, E. L., Gunther, U. L., Luo, X., Bullock, P. A. & Bachovchin, W. W. (1996). HCN, a triple-resonance NMR technique for selective observation of histidine and tryptophan side chains in $^{13}C/^{15}N$-labeled proteins. *J. Magn. Reson. B*, **113**, 236–247.
55. Bachovchin, W. W. (1986). $^{15}N$ NMR spectroscopy of hydrogen-bonding interactions in the active site of serine proteases: evidence for a moving histidine mechanism. *Biochemistry*, **25**, 7751–7759.
56. Woodward, C., Simon, I. & Tüchsen, E. (1982). Hydrogen exchange and the dynamic structure of proteins. *Mol. Cell. Bichem.* **48**, 135–160.
57. Farrow, N. A., Zhang, O., Szabo, A., Torchia, D. A. & Kay, L. E. (1995). Spectral density function mapping using $^{15}N$ relaxation data exclusively. *J. Biomol. NMR*, **6**, 153–162.
58. Farrow, N. A., Zhang, O., Forman-Kay, J. D. & Kay, L. E. (1995). Comparison of the backbone dynamics of a folded and an unfolded SH3 domain existing in equilibrium in aqueous solution. *Biochemistry*, **34**, 868–878.
59. Peng, J. W. & Wagner, G. (1995). Frequency spectrum of NH bonds in eglinc from spectral density mapping at multiple fields. *Biochemistry*, **34**, 16733–16752.
60. Bai, Y., Sosnick, T. R., Mayne, L. & Englander, S. W. (1995). Protein folding intermediates: native-state hydrogen exchange. *Science*, **269**, 192–197.
61. Ferreiro, D. U. & de Prat Gay, G. (2003). A protein–DNA binding mechanism proceeds through multi-state or two-state parallel pathways. *J. Mol. Biol.* **331**, 89–99.
62. Mok, Y. K., de Prat Gay, G., Butler, P. J. & Bycroft, M. (1996). Equilibrium dissociation and unfolding of the dimeric human papillomavirus strain-16 E2 DNA-binding domain. *Protein Sci.* **5**, 310–319.
63. de Prat Gay, G., Nadra, A. D., Corrales-Izquierdo, F. J., Alonso, L. G., Ferreiro, D. U. & Mok, Y. K. (2005). The folding mechanism of a dimeric beta-barrel domain. *J. Mol. Biol.* **351**, 672–682.
64. Ogata, K., Kanei-Ishii, C., Sasaki, M., Hatanaka, H., Nagadoi, A., Enari, M. *et al.* (1996). The cavity in the hydrophobic core of Myb DNA-binding domain is reserved for DNA recognition and trans-activation. *Nat. Struct. Biol.* **3**, 178–187.
65. Dupureur, C. M. & Dominguez, M. A., Jr (2001). The PD…(D/E)XK motif in restriction enzymes: a link between function and conformation. *Biochemistry*, **40**, 387–394.
66. Cordier, F., Hartmann, B., Rogowski, M., Affolter, M. & Grzesiek, S. (2006). DNA recognition by the Brinker repressor—an extreme case of coupling between binding and folding. *J. Mol. Biol.* **361**, 659–672.
67. Dyson, H. J. & Wright, P. E. (2002). Coupling of folding and binding for unstructured proteins. *Curr. Opin. Struct. Biol.* **12**, 54–60.
68. Spolar, R. S. & Record, M. T., Jr (1994). Coupling of local folding to site-specific binding of proteins to DNA. *Science*, **263**, 777–784.
69. Mok, Y. K., Bycroft, M. & de Prat Gay, G. (1996). The dimeric DNA binding domain of the human papillomavirus E2 protein folds through a monomeric intermediate which cannot be native-like. *Nat. Struct. Biol.* **3**, 711–717.
70. Wetzler, D. E., Castano, E. M. & de Prat Gay, G. (2007). A quasi-spontaneous amyloid route in a DNA binding gene regulatory domain: the papillomavirus HPV16 E2 protein. *Protein Sci.* **16**, 744–754.
71. Keskin, O., Jernigan, R. L. & Bahar, I. (2000). Proteins with similar architecture exhibit similar large-scale dynamic behavior. *Biophys. J.* **78**, 2093–2106.
72. Maguid, S., Fernandez-Alberti, S. & Echave, J. (2008). Evolutionary conservation of protein vibrational dynamics. *Gene*, **422**, 7–13.
73. Maguid, S., Fernandez-Alberti, S., Parisi, G. & Echave, J. (2006). Evolutionary conservation of protein backbone flexibility. *J. Mol. Evol.* **63**, 448–457.
74. Sim, J., Ozgur, S., Lin, B. Y., Yu, J. H., Broker, T. R., Chow, L. T. & Griffith, J. (2008). Remodeling of the human papillomavirus type 11 replication origin into discrete nucleoprotein particles and looped structures by the E2 protein. *J. Mol. Biol.* **375**, 1165–1177.
75. Delaglio, F., Grzesiek, S., Vuister, G. W., Zhu, G., Pfeifer, J. & Bax, A. (1995). NMRPipe: a multidimensional spectral processing system based on UNIX pipes. *J. Biomol. NMR*, **6**, 277–293.
76. Johnson, B. A. & Blevins, R. A. (1994). NMRView: a computer program for the visualization and analysis of NMR data. *J. Biomol. NMR*, **4**, 603–614.
77. Hass, M. A., Thuesen, M. H., Christensen, H. E. & Led, J. J. (2004). Characterization of μs–ms dynamics of proteins using a combined analysis of $^{15}N$ NMR relaxation and chemical shift: conformational exchange in plastocyanin induced by histidine protonations. *J. Am. Chem. Soc.* **126**, 753–765.
78. Stone, M. J., Fairbrother, W. J., Palmer, A. G., III, Reizer, J., Saier, M. H., Jr & Wright, P. E. (1992). Backbone dynamics of the *Bacillus subtilis* glucose permease IIA domain determined from $^{15}N$ NMR relaxation measurements. *Biochemistry*, **31**, 4394–4406.
79. Kay, L. E., Torchia, D. A. & Bax, A. (1989). Backbone dynamics of proteins as studied by $^{15}N$ inverse detected heteronuclear NMR spectroscopy: application to staphylococcal nuclease. *Biochemistry*, **28**, 8972–8979.