A strained DNA binding helix is conserved for site recognition, folding nucleation, and conformational modulation

Diana E. Wetzler[1]*, Mariana Gallo[2]*, Riccardo Melis[2], Tomasso Eliseo[2], Alejandro D. Nadra[1,2,3], Diego U. Ferreiro[1], Maurizio Paci[2], Ignacio E. Sánchez[1], Daniel O. Cicero[2] and Gonzalo de Prat Gay[1‡]

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

[2] University of Rome "Tor Vergata" via della Ricerca Scientifica 00133, Rome, Italy.

[3] Present address: EMBL-CRG Systems Biology Unit, Centre de Regulació Genòmica. Dr. Aiguader 88, 08003 Barcelona, Spain

[‡] Corresponding author, e-mail: gpg@leloir.org.ar, Phone: +54 11 5238 7500, ext. 3209; Fax: +54 11 5238 7501

* These authors contributed equally to this work.

This work was performed at Fundación Instituto Leloir and IIBBA-CONICET, Patricias Argentinas 435 (C1405BWE) Buenos Aires, Argentina.

Running title: A strained DNA binding helix

Keywords: E2, papillomavirus, DNA recognition, alpha helix, polyproline, 3_{10} helix

Abstract

Nucleic acid recognition is often mediated by α-helices or disordered regions that fold into α-helix upon binding. A peptide bearing the DNA recognition helix of HPV16 E2 displays type II polyproline (PII) structure as judged by pH, temperature, and solvent effects on the CD spectra. NMR experiments indicate that the canonical α-helix is stabilized at the N-terminus, while the PII forms at the C-terminus half of the peptide. Reexamination of the dihedral angles of the DNA binding helix in the crystal structure and analysis of the NMR chemical shift indexes confirm that the N-terminus half is a canonical α-helix, while the C-terminal half adopts a 3_{10} helix structure. These regions precisely match two locally driven folding nucleii, which partake in the native hydrophobic core and modulate a conformational switch in the DNA binding helix. The peptide shows only weak and unspecific residual DNA binding, 10^4-fold lower affinity and 500-fold lower discrimination capacity compared to the domain. Thus, the precise side chain conformation required for modulated and tight physiological binding by HPV E2 is largely determined by the non-canonical strained α-helix conformation, "presented" by this unique architecture.

INTRODUCTION

DNA binding proteins often recognize their target sequences and exert their functions via α-helices [1,2]. These helices may be held into a fixed structural arrangement in the unbound state by the tertiary and quaternary structure of the domain [1,2] or, alternatively, fold upon binding [3-5]. The observation of helices that fold upon binding [6-9], hints that DNA-binding helices within globular domains may behave similarly in isolation as long as they have helical propensity.

The C-terminal domain of the E2 protein from human papillomavirus type 16 (E2C) is a well characterized model for α-helix-mediated DNA sequence recognition [10]. Binding of E2C to its target DNA regulates the expression and replication of the papillomavirus genome [10], and the differences between E2C domains from high- and low-risk human papillomavirus types may lead to differenctial effects on its described multiple funcions in HPV's life cycle [11,12]. The E2C sequence folds into a homodimeric eight-stranded barrel composed of 4 strands from each symmetric monomer, with two helices per monomer packing against the outside of the barrel [10,13,14]. The topology can be described by two repetitions of β-α-β secondary structure elements. This topology is unusual, being shared only by the DNA binding domain of the Epstein Barr nuclear antigen 1 [15], which presents neither homology nor DNA binding specificity or phylogenetic relationship with E2C.

The highly conserved E2C helix-1 is the main DNA recognition element of the domain (Figure 1) [16-19]. Site-directed mutagenesis shows that the DNA binding helix accounts for most of the free energy of binding [20] and stabilizes the transition state for the two-state binding route [21]. The DNA binding helix belongs to the first β-α-β element and spans residues 293 to 305 according to the numbering of HPV16 E2C. Although this rough classification is correct in overall terms, a closer look into the dihedral angles in the high resolution structures belonging to HPV31 [22], BPV1 [23], HPV18 [24], HPV6 [25] and HPV16 [13,14] reveals that the region of the DNA binding helix can be reclassified as a mixed α- and 3_{10}-helix (Figure 1). The length of the α-helix varies from 10 residues (HPV6) to 14 residues (HPV18 monomer A and BPV1). It is even variable between the two monomers of HPV18. Conserved residues contacting directly the DNA bases (N294, K297, C298, Y301 and R302) [11] are all in the first 10 residues from the N-cap, which is the minimum length for the α-helix. There are also non-specific contacts between N294, T295, K297, R300, Y301 and R302 and the DNA backbone [21]. A second region involving the C-

terminal half of the DNA binding helix, variable in length but present in all strains, contains 3_{10} helix as suggested by molecular simulations [26]. This stretch shows a series of positively charged residues, of which K304 and K305 make electrostatic interactions with the phosphate groups of the DNA [21].

The E2C DNA binding helix is anchored to the rest of the domain by three residues forming extensive hydrophobic interactions with the outer face of the β-barrel, L296, L299 and F303 [10,13,14]. The DNA binding helix presents a well defined conformation in all crystal structures available to date [13,14,19,22-25]. However, it was noticed that amide hydrogens of the HPV31 E2C DNA binding helix exchange with solvent unexpectedly fast, suggesting that the helix is dynamically fluctuating [22]. On the other hand, the HPV16 E2C DNA binding helix shows a deviation from a canonical α-helix in the C-terminus [13,14]. Molecular dynamics studies suggest that the HPV16 protein presents a flexible DNA binding helix and a rigid β-barrel, whereas the BPV1 protein shows the opposite conformational behavior [26]. More recently, experimental evidence for the backbone mobility of some residues belonging to the HPV16 E2C DNA binding helix was presented [27]. This flexible region undergoes subtle changes when E2C binds its target DNA [14,27] and modulates binding of anti-E2C antibodies [28].

Upon addition of urea or guanidinium chloride, E2C unfolds in a two-state manner without populated intermediates [18,29]. There is residual structure in the unfolded state in the region of the DNA binding helix (Figure 1) [30], which may play a role in the initiation of folding [31,32]. E2C can also form amyloid fibrils under mild solvent conditions [33]. Unfolding of the DNA binding helix is necessary for this transition to take place, suggesting that its unusual dynamic properties may help triggering an amyloid route in a quasi-spontaneous equilibrium [33,34].

The E2C DNA binding helix seems to be the key to folding and function of the HPV16 E2 protein, the master regulator of a high-risk papillomavirus type [35]. In this work we investigate the conformational preferences and DNA binding capacity of the isolated DNA binding helix using a synthetic 19mer peptide. We then compare these properties with those of the DNA binding site in the context of the whole domain and found a strong correlation with folding nuclei [30] and a conformational switch. We discuss the implications for α-helix-nucleic acid binding in general.

RESULTS

Polyproline type II and α-helix conformations in α1-E2C from circular dichroism.

The peptide studied in this work (α1-E2C) corresponds to the DNA binding helix of the HPV16 E2C domain, plus flanking residues (Figure 1). We have used far-UV circular dichroism (CD) to determine the conformational state of the peptide in solution. In 10 mM sodium phosphate, pH 7.0 at 25 °C, α1-E2C shows a spectrum of an apparently disordered conformation, in particular compared to that of the E2C domain (Figure 2A). Careful examination of the spectra shows negative bands at 200 and 228 nm and a positive band at 218 nm, compatible with a mixture of random coil and polyproline type II (PII) conformations [36]. The ratio of the molar ellipticities at 222 and 208 nm is approximately 0.4, far from the value of 1 expected for a stable α-helix and further supporting the presence of PII conformation [37].

PII conformations are sensitive to temperature, with higher temperatures decreasing the intensity of the peak at 218 nm [38,39]. This is indeed evident from the temperature dependence of the CD spectrum of α1-E2C (Figure 2A). The characteristic broad positive band at 218 nm can be clearly observed in the difference spectra between 5 and 85 °C. The presence of an isodichroic point in the temperature transition suggests a two-state equilibrium between PII and a more unstructured state. The minimum at 199 nm at low temperature is shifted to 202 nm at the highest temperature, further supporting the presence of temperature-sensitive PII structure.

PII conformations can be stabilized by addition of denaturants such as guanidinium chloride (GdmCl) [39,40]. Addition of up to 6M GdmCl induces the formation of PII structure in α1-E2C (Figure 2B), as evidenced by the increase in the positive band around 218 nm. The inset shows a fit of the observed molar ellipticity at 5 °C and 218 nm to a two-state coil-PII model (see materials and methods, [39]). The calculated change in free energy for the coil-PII equilibrium in 0 M GdmCl is 1.3 ± 0.3 kcal/mol, which corresponds to 8.6 ± 4.5 % of PII. A similar transition was observed at 25 °C (change in free energy 2.0 ± 0.6 kcal/mol) or if urea is employed instead of GdmCl (data not shown). Although the two-state model is most likely only a crude approximation for the conformational equilibria of the peptide, these results clearly confirm that for α1-E2C in solution there is an equilibrium between coil and PII conformations.

The ellipticity of α1-E2C between pH 5 and 9 shows a defined transition around pH 7 (Figure 2C and inset). There is an increase in PII conformation at higher pH values, as deduced from the increase of the ellipticity at 200 nm and the shift in the minimum from 200 to 202 nm. We assign this transition to histidine 306, the only ionizable group of the peptide in this range (Figure 1).

The central residues of α1-E2C are in an α-helical conformation in the full-length E2C domain. We used 1,1,1-trifluoroethanol (TFE) to detect and quantify the residual helical structure present in the peptide [41,42]. The addition of increasing concentrations of TFE to α1-E2C induces the appearance of two negative bands at 222 and 208 nm in the CD spectrum, typical of an α-helix [41,42] (Figure 2D). The inset shows a fit of the observed molar ellipticity at 222 nm to a two-state coil-helix model (see materials and methods). The approximation of the helix-coil equilibrium to a two-state model is supported by the presence of an isodichroic point at around 203 nm and aims only at an estimation of the helical content in the absence of cosolvent. The fitted change in free energy for the coil-helix equilibrium in the absence of TFE is 1.7 ± 0.2 kcal/mol, which corresponds to 5.2 ± 1.7 % of helix. The molar ellipticity at high concentrations of TFE is –17830 deg·cm^2/dmol. When compared to the standard molar ellipticity of a residue in a helical conformation [43] (see methods), our result corresponds to 10 of the 19 amino acids of the peptide (53%) being on average in α-helix structure. This figure is in excellent agreement with the 11 residues of canonical helix in full-length E2C (Figure 1) [14,19].

Determination of α-helix and PII structures in α1-E2C by NMR

We investigated the solution structure of α1-E2C by NMR spectroscopy, which provides information about the conformation of individual residues in the sequence. Similar to CD measurements, we analyzed the conformation of the peptide in aqueous solution, in a water-TFE mixture and in an 8 M urea solution. Figure 3 shows the overlapping of the region of crosspeaks in the [^{13}C, ^{1}H]-HMQC spectrum for α1-E2C in the three solvent conditions (see also Tables I, II and III of supplementary material for chemical shift assignments of α1-E2C). The analysis was carried out using coupling constants and chemical shifts in addition to the usual NOE observations, to minimize biased conclusions that can be obtained in highly flexible systems.

The analysis of the secondary chemical shifts allows determining the elements of secondary structure stabilized in the peptide in the different environments here studied. Figure 4A shows the chemical shift index (CSI) [44] of the H_α, C_α, and C_β resonances of the α1-E2C peptide in the aqueous buffer and in the water-TFE and water-urea solutions. CSI values of -1, +1 and -1 for H_α, C_α, and C_β, respectively, are indicative of α-helical conformations (depicted as upward red arrows in the figure), while the opposite values are observed for β-strands (downward blue arrows in the figure). In contrast, 0 indexes (gray bars in Figure 4) suggest that no secondary chemical shift is observed for the given nucleus.

In the TFE solution the α-helical conformation spans from A293 to H306 as evidenced from the CSI analysis (Figure 4A). This result is consistent with the NOESY pattern, where α-helix characteristic crosspeaks are observed for this region (data not shown). $^1J_{C\alpha H\alpha}$ values larger than random coil values [45] and $^3J_{HNH\alpha}$ smaller than random coil values, in general smaller than 5 Hz, are typical for α-helices [46]. The values of $^1J_{C\alpha H\alpha}$ and $^3J_{HNH\alpha}$ coupling constants for this tract are indeed indicative of an α-helix (Figure 4B, black, and Tables IV and V of supplementary material).

In order to investigate regions involving PII conformation in α1-E2C, we studied its behavior in a solution containing 8 M urea. As expected for high denaturant concentrations, the chemical shifts appear mainly clustered according to amino acid type (Figure 3). Correspondingly, most of the CSI observed are 0, except for the C-terminal region of the peptide (Figure 4A). This region displays some C_α, and C_β indexes compatible with a β-strand conformation, while the H_α indexes are zero. Overall, from the CSI analysis, in the urea-water mixture the peptide seems to adopt a disordered structure with the stabilization of an extended conformation in the C-terminal region. The PII is an extended left-handed helix, often lacking characteristic proton or carbon chemical shift deviations from random coil values as found in α-helix and β-sheet conformations [47,48]. This absence of chemical shift dispersion leads to a strong overlap of signals in the NOESY spectrum and precludes an accurate determination of PII structure solely from chemical shifts and from the pattern of NOESY crosspeaks. As a consequence, PII is often misclassified as random coil conformation by NMR techniques [48,49]. Although the C-terminal region of the peptide displays some CSI coherent with a β-strand conformation, inspection of the $^1J_{C\alpha H\alpha}$ values

shows values larger than those observed in a random coil conformation for this tract (Figure 4B, red), while β-strands display the $^1J_{C\alpha H\alpha}$ smaller than random coil [45]. Likewise, the difference between measured $^1J_{C\alpha H\alpha}$ and random coil values characteristic of PII helices is reported to be larger than 1.1 Hz [48]. In the case of α1-E2C in urea, $^1J_{C\alpha H\alpha}$ larger than random coil values are observed systematically for residues from L299 to K305 (Figure 4B and Table IV of supplementary material). The $^1J_{C\alpha H\alpha}$ values of these residues are larger in the urea solution at 5°C than at 25°C (see Table IV of supplementary material) [38]. Since the large $^1J_{C\alpha H\alpha}$ values increase at 5°C and are incompatible with a β-strand, we conclude that residues L299 to K305 adopt PII conformation. The location of PII explains the influence of the protonation state of nearby histidine 306 on the coil-PII equilibrium deduced from the pH titration of α1-E2C (Figure 2C).

Two distinct structural regions for α1-E2C in aqueous solution.

From CD experiments, we know that both PII and helix conformations are partially stabilized in aqueous solution (Figure 2). Complementing NMR experiments can assign conformations to specific residues of the peptide. From A293 to L299, the crosspeak pattern in the NOESY spectrum and the CSI show a tendency of the peptide to adopt α-helical structure (Figure 4A). This tendency is weaker in water than in TFE, as judged from the magnitude of the chemical shifts and the very low intensity of the NH-NH (i, i+1) crosspeaks characteristic of α-helices (data not shown). C_β indexes are 0 for this region in water, in contrast to -1 observed in TFE. In addition, the mean difference with random coil values is -0.18 and 0.9 ppm for H_α and C_α nuclei, respectively, for residues 293-299 in water, versus -0.25 and 3.4 ppm for the same nuclei in the TFE solution. From residue R300 to the C-terminal end of the peptide, the chemical shifts are more similar to random coil values. Interestingly, these residues have $^1J_{C\alpha H\alpha}$ constants larger than those observed for random coil conformation (Figure 4B, blue and Table IV of supplementary material). The $^1J_{C\alpha H\alpha}$ coupling constants for this tract are even larger in the presence of urea (Figure 4B), a clear evidence that the PII conformation spans these residues.

In aqueous solution, α1-E2C shows a dynamic behavior, displaying a conformational equilibrium between disordered conformations and a nascent PII and disordered conformations and an α-helix structure in two well-differentiated regions. The PII region that includes R300 to

K305 corresponds very well with the PII detected in the presence of urea. The α-helix, however, only spans from A293 to L299 compared to the 293-305 region detected in the presence of TFE.

Conformation of the DNA binding helix in α1-E2C versus in the context the full-length E2C domain

We examined whether the conformational space explored by α1-E2C resembles that of the DNA binding helix in the context of the full-length domain. The determined three dimensional structure of HPV16 E2C defined residues 292-305 as the DNA binding helix [19], matching the region of the peptide that forms an α-helix in the presence of TFE. In order to correlate the results in solution presented here, together with NMR data indicating non-canonical structures in solution, we carefully reexamined the DNA binding helix of the crystal structure. The observed dihedral angles are compatible with a α-helix only for the tract 293-303, whereas the region 304-309 displays characteristics of 3_{10} helix (Figure 1). Figure 4C shows the CSI for H_α, C_α, C_β and CO nuclei for residues 289-307 in the protein [38]. Chemical shifts, the NOE pattern (not shown), and the $^3J_{HNH}$ (Table V supplementary material) observed for this region of the protein are compatible with a very stable α-helix limited to the region A293-R302 [13]. Thus, in the N-terminal region, the tertiary structure of the domain can consolidate the local conformational tendencies of α1-E2C. The DNA binding helix kinks at residue F303, which presents unusual chemical shift values, and residues 303-305 show deviations from canonical α-helix [13,14,26] (Figure 4C). In the C-terminal region, the tertiary contacts do not completely override the mixed helix/PII local tendencies of the sequence, leading to a distorted and kinked DNA binding helix after F303. This confirms the reclassification of the DNA binding helix of the crystal structure as pure α-helix for residues 293-303, and 3_{10} helix for residues 304-309.

Conformational plasticity of the DNA binding helix in the context of the full-length E2C domain

The protein contains five histidine residues per monomer, and the assignment of the ^1H and ^{15}N nuclei of the imidazolic ring for all histidines were already presented [13]. The correlations between the $H^{\varepsilon 1}$ and $H^{\delta 2}$ with the two nitrogens ($N^{\varepsilon 2}$ and $N^{\delta 1}$) were sufficiently well resolved in the long-range ^1H-^{15}N HSQC for the five histidine residues and assigned using NOE contacts [13].

By following their chemical shift as a function of pH we could estimate the apparent pK_a for each histidine. We have further explored the conformational plasticity of the C-terminal region of the DNA binding helix in full-length E2C by measuring the influence of H306 protonation on its conformation. The observed pK_a for H306 following the chemical shift of $H^{\varepsilon 1}$ is 5.65 ± 0.05 (Figure 5A) (a complete structural and dynamic analysis of the pK_a of all histidine side chains will be published elsewhere). The resulting pK_a is about 0.6 units lower than that expected for a solvent exposed histidine, indicating that the protein tertiary structure stabilizes the neutral form of H306. The observed chemical shift for the two nitrogens in the non-protonated state (211.4 and 201.9 ppm for $N^{\delta 1}$ and $N^{\varepsilon 2}$, respectively), indicate that this histidine exist as a mixture of ~6:4 of the $N^{\varepsilon 2}$-H and the $N^{\delta 1}$-H tautomers. Taking into account that for a solvent exposed histidine the expected population of the two tautomers is 4:1 [50], the interaction of H306 with both the DNA binding helix and the hydrophobic core of the monomer stabilizes the $N^{\delta 1}$-H tautomer. The interconversion between the two tautomers is fast, leading to a rapid movement of the hydrogen between the two nitrogens.

In order to estimate the influence of the protonation of H306 on the chemical shift of amide nitrogens a series of ^{1}H-^{15}N HSQC were recorded in the same range of pH (5.0-9.0). In this way, it was possible to detect those residues that titrate with the same apparent pK_a as H306 and the associated differences in ^{15}N chemical shift between the conformers with protonated and non protonated H306 (see methods). As an example, Figure 5B shows the titration curve for the amide nitrogen of F303. It displays an apparent pKa of 5.75 ± 0.08, in excellent agreement with the value found following the $H^{\varepsilon 1}$ signal of H306. Overall, eleven residues were found to be perturbed by H306 protonation (Figure 5C). These residues include the C-terminus of the helix (R300-R302), F303, the entire 3_{10} helix (K304-L309) and three residues close in space (T316, Y335 and Q347).

Mutation of "hinge" residue F303 to leucine stabilizes the E2C domain but decreases the affinity for its target DNA

The atypical C-terminal region of the E2C DNA binding helix is anchored to the tertiary structure of the domain by phenylalanine 303 [13]. E2C F303 shows unusual values for the H_α and C_α chemical shifts (Figure 4) and is strongly perturbed by titration of residue H306 (Figure 5).

We hypothesized that F303 is crucial for fixing the highly plastic 300-307 sequence in a biologically active conformation. We tested this hypothesis by mutating F303 to leucine and measuring the changes on specific DNA binding and E2C stability against irreversible denaturation.

We previously reported that 17 mutations of residues forming direct contacts with the DNA have small effects on the stability of E2C, decreasing the temperature denaturation mid-temperature (T_m) by 0.6 ± 1.6 °C [20]. In contrast, the F303L mutant is clearly more stable than wild-type E2C, having a higher T_m (Figure 6A). A fit of the data to a one-step irreversible denaturation model for a transition between a native dimer and an irreversibly denatured monomer [51,52] yields a T_m of 50.5 ± 0.2 °C (E_A 29.9 ± 0.9 kcal/mol) for the wild-type protein and 57.5 ± 0.2 °C (E_A 30.3 ± 0.8 kcal/mol) for the F303L mutant. A possible explanation is that a leucine at position 303 of E2C is better at locking the 300-307 sequence in a folded conformation than phenylalanine.

We have also measured binding of F303L E2C to a cognate target site-35 from the HPV16 genome (Figure 6B). The dissociation constant for this variant is 11.4±0.1 nM, as compared to 1.2±0.1 nM in the case of the wild type domain [20]. This corresponds to a loss in binding free energy of 1.3 ± 0.1 kcal/mol for the F303L mutation, a rather large change for a conservative replacement in a side chain that does not form direct contacts with the DNA. We interpret that a leucine at position 303 of E2C locks the 300-307 sequence in a conformation less favorable for binding than a phenylalanine. Altogether, our data confirm that the residue at position 303 is important for allowing the flexible C-terminal region of the E2C DNA binding helix to adopt a conformation that is capable of binding DNA, at the expense of protein stability.

Binding of α1-E2C to a specific E2C DNA site

We tested whether α1-E2C can bind DNA in a tight and specific manner in the absence of the rest of the domain. Since the full-length domain targets a pseudopalindromic sequence using symmetrical helices, we tested binding of α1-E2C to a target half-site, hemisite35 (see methods). Preliminary fluorescence spectroscopy experiments indicated that the affinity of α1-E2C for hemisite35 is too low to be measured using fluorescein-labeled DNA unlike full-length E2C [16]. Moreover, there is no tryptophan residue in the peptide to follow the inverse titration in excess of unlabeled DNA [16]. We carried out electrophoretic mobility shift assays (EMSA) at 25 °C in 10

mM phosphate, pH 7.0 to obtain an estimate for the affinity of α1-E2C towards DNA (Figure 7) (see materials and methods). The free DNA and the peptide-DNA complexes migrated in opposite directions due to the high isoelectric point of α1-E2C (~11) (data not shown). The fluorescence corresponding to the free DNA was used to estimate the dissociation constant (K_D). α1-E2C binds fluorescein-labeled hemisite35 with a K_D of ~7 μM (Figure 7A). The same K_D was obtained in the experiments performed with fixed concentrations of DNA and peptide (Figure 7A). Our results are in reasonable agreement with the K_D of ~1 μM measured for a similar peptide in different buffer conditions [53]. The dissociation constant for the complex between α1-E2C and a randomized oligonucleotide, iset-35, is approximately 50 μM (Figure 7A). Thus, α1-E2C shows a small degree of sequence specificity, binding its target DNA about 7-fold stronger than a nonspecific sequence. Full-length E2C binds the same DNA much more strongly (K_D in the low nanomolar range [16,18,20]) and with more than 1000-fold specificity [16,18].

α1-E2C shows nascent native helical structure in solution, strongly resembling the conformation of the same sequence within the full-length E2C domain, both free and in complex with DNA (Figure 4) [13,14]. We have used CD spectroscopy to test whether the peptide folds into a helical conformation upon specific DNA binding. We co-incubated α1-E2C and hemisite35 at 25 °C in 10 mM phosphate, pH 7.0. The concentration of both species was 25 μM, enough for binding saturation according to the estimated K_D of ~7 μM. An analogous experiment using full-length E2C was able to detect binding-induced changes in the DNA [16]. Figure 7B shows that in both the far-UV and the near-UV regions the sum of the individual spectra was indistinguishable from the spectrum of the mixture. Thus, formation of the complex between α1-E2C and DNA does not induce alpha helix formation in α1-E2C and does not induce conformational changes in the DNA. The lack of structural stabilization by binding agrees with the low affinity non-specific binding we observe.

Discussion

The re-definition of the DNA binding helix of the HPV16 E2C domain in solution as a strained helix, with a bona fide α-helix at its N-terminal half, and a distorted C-terminal half containing 3_{10} helix, finds a remarkable correlation with our results for the α1-E2C peptide (Figure 1, bottom). Overall, the peptide shows a significant tendency to adopt a helical conformation since a long and stable α-helix starting from residue A292 to H306 is clearly formed in TFE. At high concentrations of urea, there is an increase in PPII elongated conformations in the C-terminal end of the peptide. In the absence of TFE or urea we found evidence for the incipient formation of both α-helix (A292-C298) and PPII conformations (L299-K305). The large number of positively charged residues present in the second half of the peptide appears as a logical explanation for the formation of an extended PPII conformer.

The particular dimeric β-barrel topology of E2C holds the DNA binding helix in a "ready for binding" conformation [10,13,14,17,21]. In line with this, the F303L substitution shows that anchoring the C-terminal half of the DNA binding helix to the outer face of the β-barrel not only modulates the stability of the protein, but also affinity for DNA as much as mutation of many residues contacting the DNA (Figure 6) [20]. This "ready for binding" conformation is strained and malleable. Molecular dynamics simulations suggested that the C-terminal region of the DNA binding helix spends a significant amount of time in a polyproline type II conformation [26,27], and the two monomers in the crystal structure of HPV18 E2C differ mostly in the conformation of this part, leading to a different structural alignment for monomer A (Figure 1). The tertiary structure of this region heavily stabilizes the neutral form of H306, leading to a diminished pKa value with respect to a solvent-exposed histidine. Upon protonation of H306, several amide groups of residues in this region show chemical shift perturbation. Although we cannot exclude that this behaviour is simply the consequence of being close to the histidine ring that changes protonation state, it is conceivable that a protonated histidine will cause a conformational transition in view of the altered pKa and tautomer populations observed for H306 at physiological pH. Although further experiments will be necessary to understand how the domain stabilizes a 3_{10} helix in this region, we have now established experimentally that the helix to polyproline II transition is caused by the conformational preferences of this sequence stretch and that these are of local nature (Figures 4 and 5). It has been suggested that polyproline type II

conformations can act as nucleation sites for the formation of oligomeric β-sheet structure [54,55]. Since the DNA binding helix unfolds in the initial step of the E2C amyloid route [33], we propose that nascent PII structure in its C-terminal region may act as a nucleation site for the amyloid route of the domain.

NMR measurements in mild denaturing conditions identified residual structure in the unfolded state of E2C in the regions corresponding to the DNA binding helix and the β2 strand in the native state of the domain [30]. The residual structure consists mainly of α-helix in residues 292-297 and of beta sheet and turns in residues 300-304 [30], in good agreement with the nascent secondary structure of α1-E2C (Figure 1). This result shows that residual structure in the E2C DNA binding helix in the unfolded state is of local nature and is independent of residual structure in the adjacent β2-strand. The coincidence between the residual structure in the unfolded state and the regions of α and 3_{10} helix in the folded domain is remarkable and underlines the importance of the local conformational tendencies of this sequence during folding of the entire domain.

Short natural sequences from intrinsically disordered domains [6,7], or proteins with few long-range contacts [8,9] have high intrinsic helix propensity and are able to fold into α-helices upon binding of their cognate DNA [3-5]. Despite having high intrinsic helix propensity, binding of the α1-E2C peptide to hemisite35 is extremely weak and does not induce formation of alpha helix or any other observable structure (Figure 7). Interestingly, non-cognate peptide-nucleic acid interactions from in vitro selection adopt from structured helices to disordered but foldable structures upon binding [56]. These results suggest that evolution finely tunes the conformational preferences of DNA binding helices beyond solely helical propensity to ensure tight and specific DNA binding and regulatory functions modulated by conformational switches. This tuning may be exerted through local interactions [6-9] or, as for α1-E2C, in the context of this particular dimeric β-barrel domain. The case of E2 shows a strained "non-canonical" DNA binding helix, highly conserved among all known papillomaviruses, that contains two of the three folding nuclei [30], indicating an evolution driven strong interplay between DNA recognition, folding initiation and conformational modulation. Perhaps these traits are reflected in the unusually high

global dynamic nature yet highly stable characteristics of this domain, with a strong coupling between folding and dimerization, i.e., tertiary and quaternary structure [22,34].

Materials and methods

Peptide Synthesis and Purification. The α1-E2C peptide contains residues 289-307 of HPV16 E2 (LKGDANTLKCLRYRFKKHA). The last residue in the sequence was mutated from Cys to Ala to avoid redox side reactions. The peptide was synthesized at the Yale University Keck facility, with N-terminus acetylated and C-terminus amidated, and purified by reverse phase HPLC. We confirmed its molecular mass by mass spectrometry. Quantification of peptide concentration was carried out by absorbance at 274 nm and 220 nm.

DNA Synthesis. Fluorescein-labeled single-stranded oligonucleotides containing the complete target site 35 of the HPV16 genome (5' GTAACCGAAATCGGTTGA 3'), a hemisite (5' GTAGCACAAATCGGTTGA 3') and a randomized site (5' ACATGGACCTGTCAAAGTA 3') were purchased, HPLC purified, from Integrated DNA technologies (Coralville, IA). Double-stranded oligonucleotides were annealed by mixing equal amounts of each strand in 10 mM Bis-Tris-HCl buffer, pH 7 and 50 mM NaCl, incubating 5 minutes at 95 °C, and slowly cooling to 25 °C for 16 hours. Completeness was checked by native polyacrylamide gel electrophoresis.

Circular Dichroism. CD measurements were carried out on a Jasco J-810 spectropolarimeter using a Peltier temperature-controlled sample holder and a 0.1 cm path length cell. The measured molar ellipticity of a peptide containing α-helix, polyproline type II (PII) and coil conformations is the sum of the individual contributions [57]:

$$[\theta] = f_{\alpha}[\theta]_{\alpha} + f_{PII}[\theta]_{PII} + f_{U}[\theta]_{U} \qquad [1]$$

where f_i are the fractions of each conformation and $[\theta]_i$ the mean residue ellipticities of these conformations. The relationship between the number of residues in α-helical conformation and the ellipticity at 222 nm follows an empirical equation [43]:

$$[\theta]_{222nm} = -39500 \cdot \left(1 - \frac{2.57}{n}\right) \qquad [2]$$

There are not established values for $[\theta]_{PII}$ and $[\theta]_{U}$.

We used increasing amounts of 1,1,1-trifluoroethanol (TFE) and guanidinium chloride (GdmCl) to stabilize α-helix and PII conformations, respectively [40-42]. If we assume that the populations of PII (for the TFE titration) and α-helix (for the GdmCl titration) do not change significantly during the titration, we can fit the observed ellipticity to a two-state equilibrium model [39,58].

$$Coil \leftrightarrow \alpha - helix; Coil \leftrightarrow PII \quad [3]$$

We considered the free energy for α-helix formation to depend linearly on the TFE/water molar ratio and the PII structures to depend on the molar concentration of GdmCl, with a proportionality constant m:

$$\begin{aligned} \Delta G_{coil\leftrightarrow\alpha-helix}([TFE]/[water]) &= \Delta G^{0\%TFE}_{coil\leftrightarrow\alpha-helix} - m_{TFE}\cdot[TFE]/[water]; \\ \Delta G_{coil\leftrightarrow PII}([GdmCl]) &= \Delta G^{0MGdmCl}_{coil\leftrightarrow PII} - m_{PII}\cdot[GdmCl] \end{aligned} \quad [4]$$

The molar ellipticity at a fixed wavelength (222 nm for α-helix, 218 nm for PII) during the titration can be fitted to the following equation to extract values for DG and m:

$$[\theta] = \frac{[\theta]^{TFE,GdmCl} + [\theta]^{water}\cdot exp\left(-\Delta G\left(TFE,GdmCl\right)/RT\right)}{1+exp\left(-\Delta G\left(TFE,GdmCl\right)/RT\right)} \quad [5]$$

where $[\theta]^{water}$ and $[\theta]^{TFE,GdmCl}$ are the mean residue ellipticities in water and at high cosolvent concentration, R is the gas constant and T is the temperature.

NMR Experiments. α1-E2C (3 mM) was dissolved in aqueous buffer (20 mM phosphate buffer pH 6.5, 2 mM DTT, 0.01% NaN_3 and 5% D2O), TFE/water mixture (40% TFE-d_3, 60% 20 mM phosphate pH 6.5, 2 mM DTT and 0.01% NaN_3) and urea/water mixture (8M urea, 20 mM phosphate pH 6.5, 2 mM DTT, 0.01% NaN_3 and 5% D2O). We used urea instead of GdmCl for NMR experiments because it was not possible to tune the probe in presence of high concentrations of GdmCl. The NMR experiments were performed at 25 °C or 5 °C using a Bruker Avance 700 and Avance 400 spectrometers equipped with triple resonance probes incorporating self-shielded gradient coils. All the heteronuclear correlation experiments were carried out at natural abundance. 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 [59] or the echo-antiecho method. The spectra were processed on Silicon Graphics workstations by the NMRPipe software [60] and analyzed using NMR View [61]. Assignment of backbone resonances of α1-E2C was performed using a combination of the following 2D NMR experiments: TOCSY (mixing times of 50 and 80 ms), [^{1}H-^{13}C] and [^{1}H-^{15}N] HMQC [62,63], and [^{1}H-^{13}C] HMQC-TOCSY (mixing times of 40 and 80 ms) [64], the latter mentioned being very helpful due to the overlap of the signals. The sequential connectivity across the peptide was established using the NOESY and ROESY spectra (mixing times of 0.15 s and 0.20 s) [65], according to the sequential assignment method [66]. The $^{1}J_{C\alpha H\alpha}$ coupling constants were measured from the in-phase splitting patterns of the C_α-H_α cross-peaks in the HMQC spectra without carbon decoupling. The $^{3}J_{HNH\alpha}$ values were measured from the in-phase splitting patterns of the NH-C cross peaks in the [^{1}H-^{13}C] HSQC-TOCSY spectrum collected with 8k complex data points in F2.

Uniformly ^{15}N-labeled HPV-16 E2 DNA-binding domain was expressed and purified as previously described [31]. The protein concentration was 0.4 mM in 50 mM sodium phosphate, 5 mM DTT, pH 6.5. NMR experiments of protein samples were performed at 30 °C on a Bruker Avance400. A series of two-dimensional ^{1}H-^{15}N HSQC spectra were recorded at distinct pH values, namely pH 5.0, 5.5, 6.0, 6.5, 7.0, 7.5, 8.0, 8.5, 9.0. At pH higher than 9.0 and lower than 5.0, irreversible protein precipitation occurred. A total of 4096 complex points in t_2 with 256 t_1 increments were acquired. The spectral widths were set to 6410 Hz in F_2 and 1120 Hz in F_1, placing the carrier frequency at 4.70 ppm in F_2 and at 118.6 ppm in F_1. At every pH a two-dimensional ^{1}H-^{15}N long range HSQC (LR-HSQC) spectrum was acquired to observe two-bond correlations between nitrogen and proton resonances belonging the histidine side-chain spin system [67]. The delay for the INEPT-type magnetization transfer in the LR-HSQC was set to 11 ms. A total of 2048 complex points in t_2 with 64 t_1 increments were acquired. Spectral width was 6410 Hz in F_2 and 4258 Hz in F_1, placing the carrier frequency at 4.70 ppm in the proton dimension and at 208.4 ppm in the nitrogen dimension.

The pK_a values of titration curves were determined by analyzing the pH dependencies of amide N and H^N resonances in the HSQC spectra for backbone signals and of $N^{\delta 1}/N^{\delta 2}$ and $H^{\delta 2}/H^{\delta 1}$ resonances in the LR-HSQC spectra for histidine side chains. The pH titration curves were fitted to a modified Henderson-Hasselbalch equation by nonlinear least-squares analysis:

$$\delta_{obs} = \frac{\delta_{pr} + \delta_{nonpr} \cdot 10^{(pH - pKa)}}{1 + 10^{(pH - pKa)}} \quad [6]$$

in which δ_{obs} is the chemical shift observed at each pH value and δ_{pr} and δ_{nonpr} are the chemical shifts for the protonated and deprotonated histidines, respectively. Curve fits were performed using the KaleidaGraph3.5 Software (Synergy Software).

E2C mutant F303L. Site-directed mutagenesis, recombinant expression and purification were performed as described [16,20]. Equilibrium binding to the full target DNA site35 was followed by fluorescence spectroscopy in 20 mM bis-Tris/HCl pH 7.0, 0.2 M NaCl, 1 mM DTT, at 25 °C. A 10 nM solution of the fluorescein-labeled DNA was titrated with increasing amounts of F303L E2C [16,20].

Irreversible thermal denaturation was followed by the change in molar ellipticity at 225 nm in 20 mM bis-Tris/HCl pH 7.0, 0.2 M NaCl, 1 mM DTT, 1 M GdmCl, scanning at 2 K/min. The data were fitted to a one-step irreversible denaturation model for a transition between a native dimer and an irreversibly denatured monomer [51,52].

$$[\theta] = \left(exp\left[-exp\left(\frac{E_A(T - T_m)}{RT_m^2} \right) \right] \right) \cdot [\theta]_N + \left(1 - \left(exp\left[-exp\left(\frac{E_A(T - T_m)}{RT_m^2} \right) \right] \right) \right) \cdot [\theta]_U \quad [7]$$

Where $[\theta]_N$ and $[\theta]_U$ are the molar ellipticities of the native and irreversibly denatured states, R is the gas constant, T is temperature, T_m is the temperature at which the protein is 50% denatured and E_A is the activation enthalpy at T_m.

Electrophoretic Mobility Shift Assay (EMSA). Reaction mixtures contained 1 µM fluorescein-labeled hemisite35 (or fluorescein-labeled randomized site) and different amounts of α1-E2C or 10µM α1-E2C and increasing amounts of fluorescein-labeled hemisite35. Mixtures were incubated 60 min at room temperature in a final volume of 10 µL in 10 mM phosphate pH 7.0, 1 mM DTT. Mixtures were loaded continuously into running non-denaturing 10% polyacrylamide gels containing TBE 0.5x (0.1 M Tris-HCl pH 8.0 0.15 M sodium borate, 4 mM EDTA). The gels were resolved at 4 V/cm, 10 °C for 2 h. Fluorescein bands corresponding to free oligonucleotide were detected by UV and used to extract the dissociation constant as described

[16,20]. We used a gel with increasing concentrations of fluorescein-labeled site without peptide to check that the fluorescence signal was proportional to oligonucleotide concentration (Figure 7B, left).

Acknowledgements

This work was supported by Welcome Trust [GR077355AYA]; Agencia Nacional de Promoción Científica y Tecnológica [PICT 2000 01-08959]; Consejo Nacional de Investigaciones Científicas y Técnicas [doctoral fellowship to A.D.N. and D.U.F., G.P.G. and D.W. are Career Investigators]; and Agencia Española de Cooperación Internacional [postdoctoral fellowship to I.E.S.]. We would like to acknowledge Mariano Dellarole for discussion.

Figure legends

Figure 1. Secondary structure in the DNA binding helix region for E2C domains from HPV31 [22], BPV1 [23], HPV18 [24], HPV6 [25] and HPV16 [13,14], in the unfolded state of HPV16 E2C [30] and in the α1-E2C peptide under different conditions (this work). β-strands are shown as arrows, α-helices as red cylinders, 3_{10} helices as green rectangles and folding nuclei as blue cylinders. Numbering corresponds to HPV16 E2C. The first residues in the HPV18 E2C structure are HM due to the cloning procedure, but NT in nature. The two monomers of this domain present slightly different structures in the crystal and are shown separately.

Figure 2. Structural characterization of α1-E2C by far-UV circular dichroism. Buffer is 10 mM sodium phosphate, pH 7.0 and peptide concentration is 40 μM, temperature is 25 °C unless stated otherwise. (A) Spectra of α1-E2C from 5 to 85 °C. Also shown is the spectrum of full-length E2C for comparison (dashed line). Inset: Difference spectrum between 5 and 85 °C. (B) Spectra of α1-E2C at 5 °C from 0 to 6M GdmCl. Inset: Molar ellipticity at 218 nm as a function of GdmCl concentration and fit to a two-state coil-PII equilibrium (3). (C) Spectra of α1-E2C in 10 mM sodium citrate-phosphate buffer at pH from 3.4 to 8.4. Inset: Molar ellipticity at 200 nm as a function of pH. (D) Spectra of α1-E2C from 0 to 60% (v/v) TFE. Inset: Molar ellipticity at 222 nm as a function of TFE concentration and fit to a two-state coil-helix equilibrium.

Figure 3. [^{13}C,^{1}H]-HSQC NMR spectra of α1-E2C in aqueous buffer (blue), 40% TFE (black) and 8M urea (red). All experiments were performed in 20 mM phosphate, pH 6.5 at 25 °C and natural isotope abundance.

Figure 4. Structural characterization of α1-E2C by NMR. (A) Chemical shift indexes of $^{1}H_{\alpha}$, $^{13}C_{\alpha}$ and $^{13}C_{\beta}$ for α1-E2C in 40% TFE, in aqueous buffer and in 8 M urea in the top, middle and lower panels, respectively. The chemical shift indexes were calculated using NMRView [44]. Index values depicted as downward blue arrows are indicative of β-strands, upward red arrows of α-helices, and gray bars are used for 0 index values. (B) $^{1}J_{C\alpha H\alpha}$ coupling constants relative to a random coil conformation [56]. The value of 1.1 Hz is signaled in the plot. Black triangles: TFE solution. Red squares: urea solution. Blue circles: aqueous solution. (C) Chemical shift indexes

of $^{1}H_{\alpha}$,$^{13}C_{\alpha}$, $^{13}C_{\beta}$, and CO for the DNA recognition helix in the full-length E2C ([44], BioMagResBank database accession number 5952).

Figure 5. Effect of pH on the E2C DNA binding helix. (A) Titration of H306, followed by ^{1}H chemical shift. The line is a fit to a modified Henderson-Hasselbalch equation (see methods). (B) Titration of F303, followed by ^{15}N chemical shift. The line is a fit to a modified Henderson-Hasselbalch equation (see methods). (C) Residues affected by titration of H306. The observed pK_a-values for these residues range from 5.4 to 5.8.

Figure 6. Mutation of hinge residue F303 of E2C to leucine. (A) Stability of wild type (empty circles) and F303L (black circles) E2C against irreversible thermal denaturation, followed by CD at 225 nm. The line is a fit to a one-step irreversible denaturation model for a transition between a native dimer and an irreversibly denatured monomer [51,52] (see materials and methods). (B) Binding of F303L E2C to fluorescein-labeled specific target site35, followed by fluorescence. The line is a fit to a quadratic binding equation [20] (see materials and methods).

Figure 7. DNA binding by α1-E2C. (A) Electrophoretic Mobility Shift Assay. Top panels show the polyacrylamide gel and bottom panels the complex concentration as a function of α1-E2C or DNA. All lines are fits to a quadratic binding equation [20]. Reaction mixtures in the left panels contain 1 μM fluorescein-labeled hemisite35 (filled circles) or randomized iset35 (empty circles) and increasing amounts of α1-E2C. Reaction mixtures in the right panels contain increasing amounts of fluorescein-labeled hemisite35 and no peptide or 10μM α1-E2C (filled circles). The lanes in the absence of peptide show that the fluorescence signal is proportional to the concentration of DNA. (B) CD spectra of a stoichiometric (25 μM) mixture of α1-E2C and hemisite35 in 10 mM phosphate buffer, pH 7.0 (circles) and sum of spectra for the DNA and the peptide (triangles). The bottom panel shows the difference between the spectrum of the mixture and the sum of spectra.

References

1. Huffman, J. L.; Brennan, R. G. Curr Opin Struct Biol 2002, 12, 98-106.
2. Warren, A. J. Curr Opin Struct Biol 2002, 12, 107-114.
3. Wright, P. E.; Dyson, H. J. J Mol Biol 1999, 293, 321-331.
4. Dyson, H. J.; Wright, P. E. Curr Opin Struct Biol 2002, 12, 54-60.
5. Reedstrom, R. J.; Royer, C. A. J Mol Biol 1995, 253, 266-276.
6. Vila, R.; Ponte, I.; Collado, M.; Arrondo, J. L.; Suau, P. J Biol Chem 2001, 276, 30898-30903.
7. Vila, R.; Ponte, I.; Collado, M.; Arrondo, J. L.; Jimenez, M. A.; Rico, M.; Suau, P. J Biol Chem 2001, 276, 46429-46435.
8. Rao, H.; Mohr, S. C.; Fairhead, H.; Setlow, P. FEBS Lett 1992, 305, 115-120.
9. Lee, K. S.; Bumbaca, D.; Kosman, J.; Setlow, P.; Jedrzejas, M. J. Proc Natl Acad Sci U S A 2008, 105, 2806-2811.
10. Hegde, R. S. Annu Rev Biophys Biomol Struct 2002, 31, 343-360.
11. Sanchez, I. E.; Dellarole, M.; Gaston, K.; de Prat Gay, G. Nucleic Acids Res 2008, 36, 756-769.
12. Morgan, I. M.; Donaldson, M. M. In Papillomavirus Research: From Natural History to Vaccines and Beyond; Saveria Campo, M., Ed.; Caisreir Academic Press: Wymondham, 2006, p 73-82.
13. Nadra, A. D.; Eliseo, T.; Mok, Y. K.; Almeida, C. L.; Bycroft, M.; Paci, M.; Prat Gay, G. d.; Cicero, D. O. J Biomol NMR 2004, 30, 211-214.
14. Cicero, D. O.; Nadra, A. D.; Eliseo, T.; Dellarole, M.; Paci, M.; Prat Gay, G. d. Biochemistry 2006, 45, 6551-6560.
15. Bochkarev, A.; Barwell, J. A.; Pfuetzner, R. A.; Bochkareva, E.; Frappier, L.; Edwards, A. M. Cell 1996, 84, 791-800.
16. Ferreiro, D. U.; Lima, L. M.; Nadra, A. D.; Alonso, L. G.; Goldbaum, F. A.; Prat Gay, G. d. Biochemistry 2000, 39, 14692-14701.
17. Ferreiro, D. U.; Prat Gay, G. d. J Mol Biol 2003, 331, 89-99.
18. Dellarole, M.; Sanchez, I. E.; Freire, E.; Prat Gay, G. d. Biochemistry 2007, 46, 12441-12450.
19. Hegde, R. S.; Androphy, E. J. J Mol Biol 1998, 284, 1479-1489.
20. Ferreiro, D. U.; Dellarole, M.; Nadra, A. D.; Prat Gay, G. d. J Biol Chem 2005, 280, 32480-32484.
21. Ferreiro, D. U.; Sanchez, I. E.; de Prat Gay, G. Proc Natl Acad Sci U S A 2008, 105, 10797-10802.
22. Liang, H.; Petros, A. M.; Meadows, R. P.; Yoon, H. S.; Egan, D. A.; Walter, K.; Holzman, T. F.; Robins, T.; Fesik, S. W. Biochemistry 1996, 35, 2095-2103.
23. Hegde, R. S.; Grossman, S. R.; Laimins, L. A.; Sigler, P. B. Nature 1992, 359, 505-512.
24. Kim, S. S.; Tam, J. K.; Wang, A. F.; Hegde, R. S. J Biol Chem 2000, 275, 31245-31254.
25. Hooley, E.; Fairweather, V.; Clarke, A. R.; Gaston, K.; Brady, R. L. Nucleic Acids Res 2006, 34, 3897-3908.
26. Falconi, M.; Santolamazza, A.; Eliseo, T.; Prat Gay, G. d.; Cicero, D. O.; Desideri, A. Febs J 2007, 274, 2385-2395.
27. Falconi, M.; Oteri, F.; Eliseo, T.; Cicero, D. O.; Desideri, A. Biophys J 2008, 95, 1108-1117.

28. Cerutti, M. L.; Ferreiro, D. U.; Sanguineti, S.; Goldbaum, F. A.; de Prat-Gay, G. Biochemistry 2006, 45, 15520-15528.
29. Mok, Y. K.; Prat Gay, G. d.; Butler, P. J.; Bycroft, M. Protein Sci 1996, 5, 310-319.
30. Mok, Y. K.; Alonso, L. G.; Lima, L. M.; Bycroft, M.; Prat Gay, G. d. Protein Sci 2000, 9, 799-811.
31. Mok, Y. K.; Bycroft, M.; Prat Gay, G. d. Nat Struct Biol 1996, 3, 711-717.
32. Prat Gay, G. d.; Nadra, A. D.; Corrales-Izquierdo, F. J.; Alonso, L. G.; Ferreiro, D. U.; Mok, Y. K. J Mol Biol 2005, 351, 672-682.
33. Wetzler, D. E.; Castano, E. M.; Prat Gay, G. d. Protein Sci 2007, 16, 744-754.
34. Prat Gay, G. d.; Gaston, K.; Cicero, D. O. Front Biosci 2008, 13, 6006-6021.
35. Munoz, N.; Bosch, F. X.; de Sanjose, S.; Herrero, R.; Castellsague, X.; Shah, K. V.; Snijders, P. J.; Meijer, C. J. N Engl J Med 2003, 348, 518-527.
36. Shi, Z.; Olson, C. A.; Rose, G. D.; Baldwin, R. L.; Kallenbach, N. R. Proc Natl Acad Sci U S A 2002, 99, 9190-9195.
37. Lakshminarayanan, R.; Fan, D.; Du, C.; Moradian-Oldak, J. Biophys J 2007, 93, 3664-3674.
38. Tiffany, M. L.; Krimm, S. Biopolymers 1972, 11, 2309-2316.
39. Garcia-Alai, M. M.; Gallo, M.; Salame, M.; Wetzler, D. E.; McBride, A. A.; Paci, M.; Cicero, D. O.; Prat Gay, G. d. Structure 2006, 14, 309-319.
40. Tiffany, M. L.; Krimm, S. Biopolymers 1973, 12, 575-587.
41. Buck, M. Q Rev Biophys 1998, 31, 297-355.
42. Nelson, J. W.; Kallenbach, N. R. Proteins 1986, 1, 211-217.
43. Chen, Y. H.; Yang, J. T.; Chau, K. H. Biochemistry 1974, 13, 3350-3359.
44. Wishart, D. S.; Sykes, B. D. Methods Enzymol 1994, 239, 363-392.
45. Vuister, G. W.; Delaglio, F.; Bax, A. J Am Chem Soc 1992, 114, 9674-9675.
46. Smith, L. J.; Bolin, K. A.; Schwalbe, H.; MacArthur, M. W.; Thornton, J. M.; Dobson, C. M. J Mol Biol 1996, 255, 494-506.
47. Shi, Z.; Woody, R. W.; Kallenbach, N. R. Adv Protein Chem 2002, 62, 163-240.
48. Lam, S. L.; Hsu, V. L. Biopolymers 2003, 69, 270-281.
49. Siligardi, G.; Drake, A. F. Biopolymers 1995, 37, 281-292.
50. Bachovchin, W. W. Biochemistry 1986, 25, 7751-7759.
51. Sanchez-Ruiz, J. M. Biophysical Journal 1992, 61, 921-935.
52. Rodriguez-Larrea, D.; Minning, S.; Borchert, T. V.; Sanchez-Ruiz, J. M. J Mol Biol 2006, 360, 715-724.
53. Faber-Barata, J.; Mohana-Borges, R.; Lima, L. M. FEBS Lett 2006, 580, 1919-1924.
54. Blanch, E. W.; Morozova-Roche, L. A.; Cochran, D. A.; Doig, A. J.; Hecht, L.; Barron, L. D. J Mol Biol 2000, 301, 553-563.
55. JiJi, R. D.; Balakrishnan, G.; Hu, Y.; Spiro, T. G. Biochemistry 2006, 45, 34-41.
56. Smith, C. A.; Calabro, V.; Frankel, A. D. Mol Cell 2000, 6, 1067-1076.
57. Park, S. H.; Shalongo, W.; Stellwagen, E. Protein Sci 1997, 6, 1694-1700.
58. Pace, C. N. Methods Enzymol 1986, 131, 266-280.
59. Marion, D.; Wuthrich, K. Biochem Biophys Res Commun 1983, 113, 967-974.
60. Delaglio, F.; Grzesiek, S.; Vuister, G. W.; Zhu, G.; Pfeifer, J.; Bax, A. J Biomol NMR 1995, 6, 277-293.
61. Johnson, b. A.; Blevins, R. A. J Biomol NMR 1994, 4, 603-614.
62. Bax, A.; Griffey, R. H.; Hawkins, B. L. J Magn Resonance 1983, 55, 301-315.
63. Grzesiek, S.; Bax, A. J Am chem Soc 1993, 115, 12593-12594.

64. Braunschweiler, L.; Ernst, R. R. J Magn Resonance 1983, 55, 521-528.
65. Bax, A.; Davis, D. G. J Magn Resonance 1985, 65, 355-360.
66. Wuthrich, K. NMR of proteins and nucleic acids.; Wiley-Interscience Publication: New York, 1986.
67. Pelton, J. G.; Torchia, D. A.; Meadow, N. D.; Roseman, S. Protein Sci 1993, 2, 543-558.

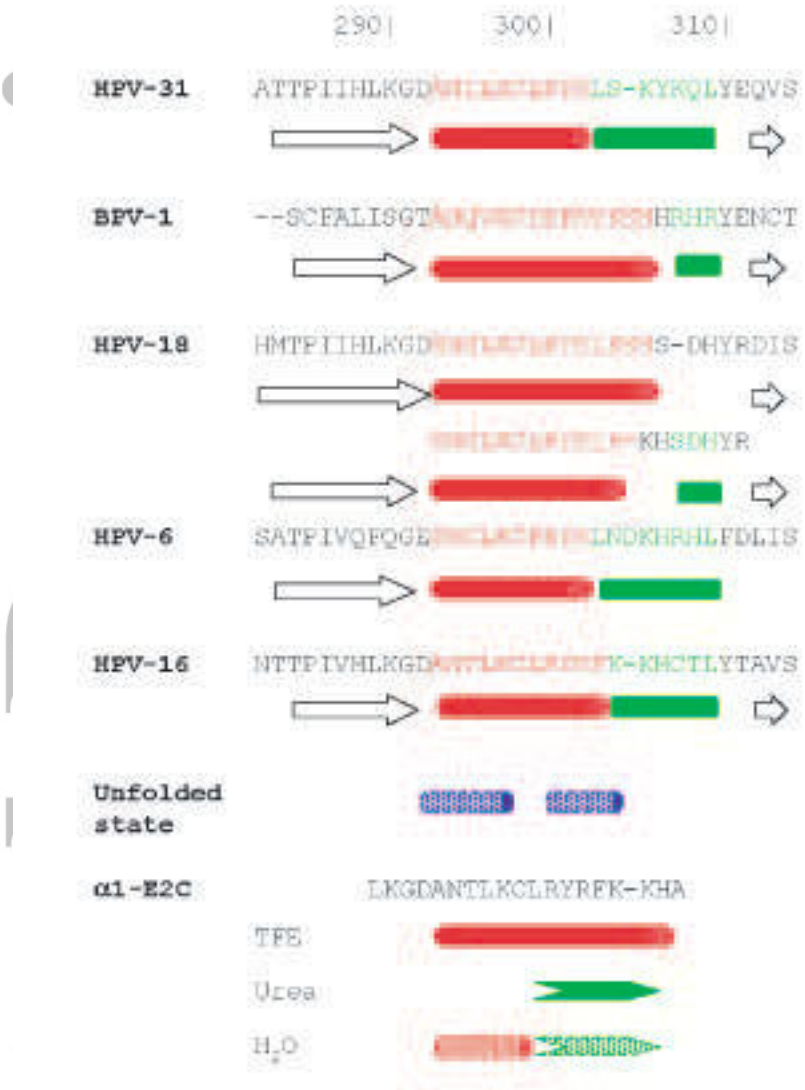


Figure 1

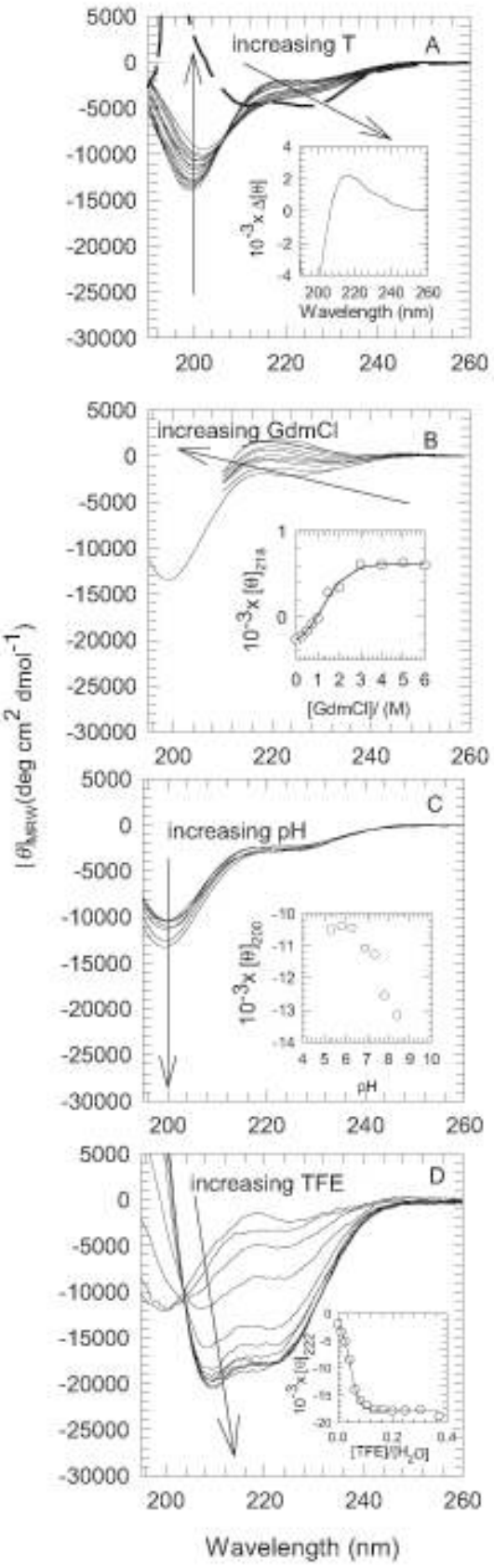

increasing T
A
increasing GdmCl
B
increasing pH
C
increasing TFE
D
Wavelength (nm)
[GdmCl] (M)
pH
[TFE]/[H2O]

Figure 2

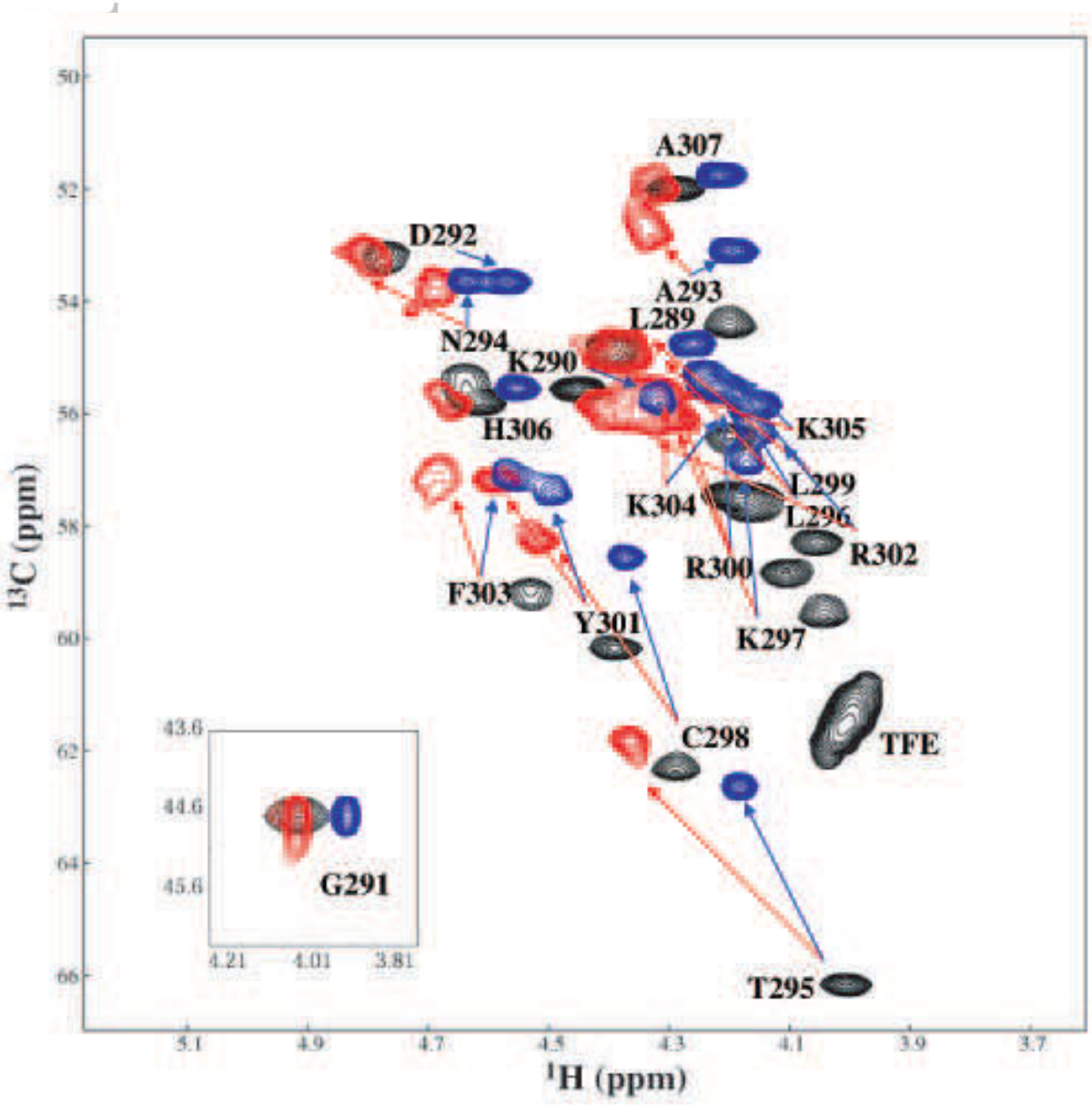


Figure 3

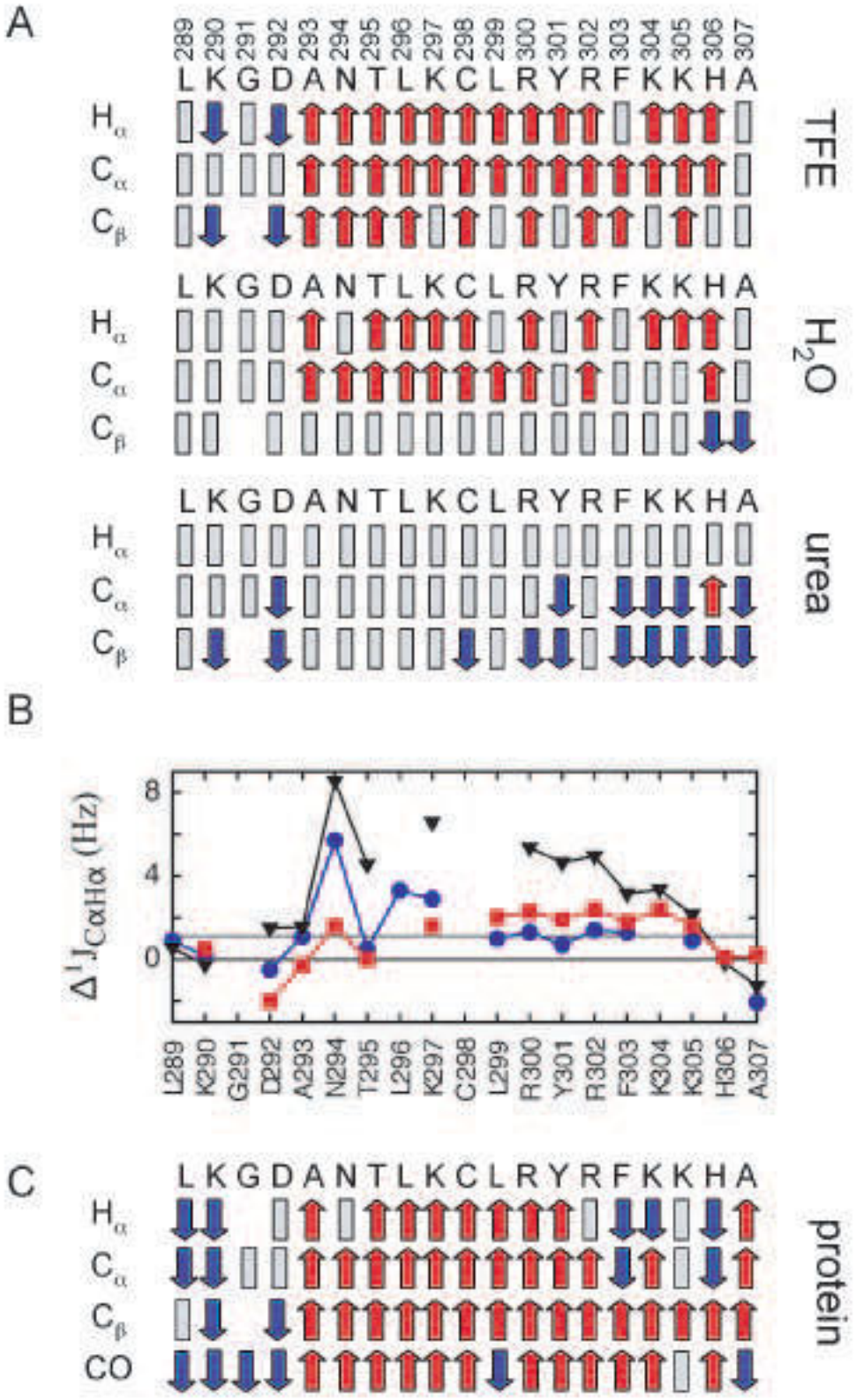

A
289 290 291 292 293 294 295 296 297 298 299 300 301 302 303 304 305 306 307
L K G D A N T L K C L R Y R F K K H A
Hα
Cα
Cβ
TFE
H2O
urea
B
Δ1JCαHα (Hz)
8
4
0
L289 K290 G291 D292 A293 N294 T295 L296 K297 C298 L299 R300 Y301 R302 F303 K304 K305 H306 A307
C
CO
protein

Figure 4

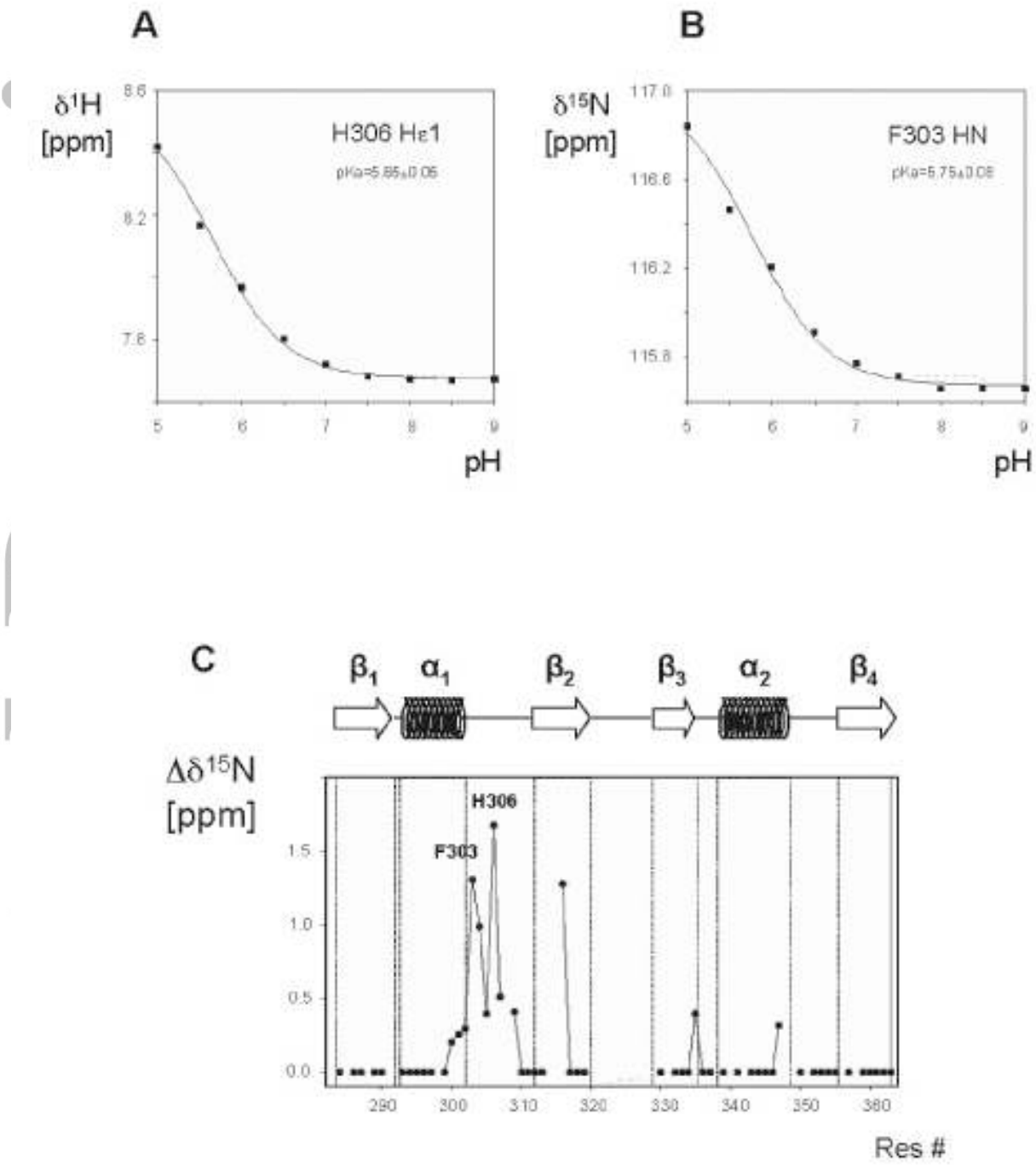
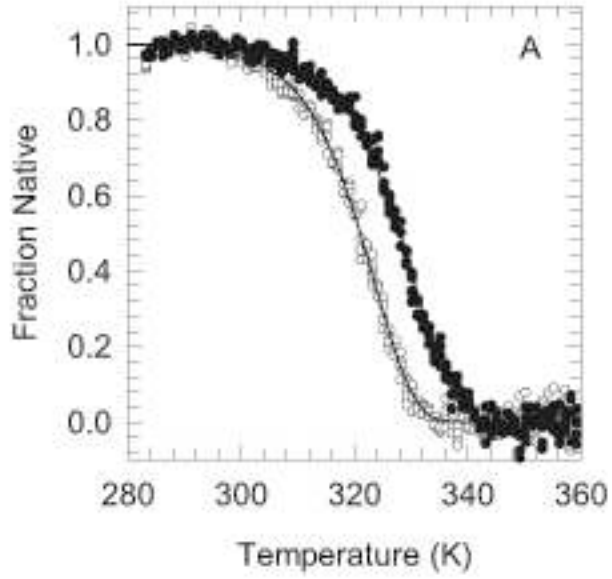

A
B
δ¹H
[ppm]
H306 Hε1
pKa=5.85±0.05
δ¹⁵N
[ppm]
F303 HN
pKa=5.75±0.08
pH
pH
C
β1 α1 β2 β3 α2 β4
ΔδN
[ppm]
H306
F303
Res #

Figure 5


A
Fraction Native
1.0
0.8
0.6
0.4
0.2
0.0
280
300
320
340
360
Temperature (K)

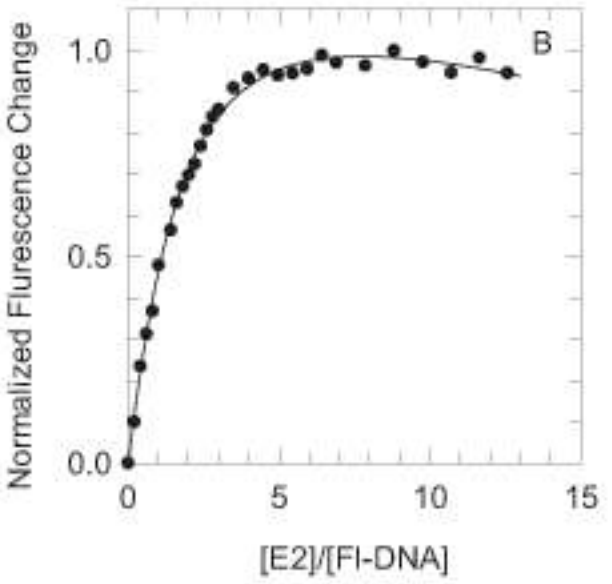

B
Normalized Flurescence Change
1.0
0.5
0.0
0
5
10
15
[E2]/[FI-DNA]

Figure 6

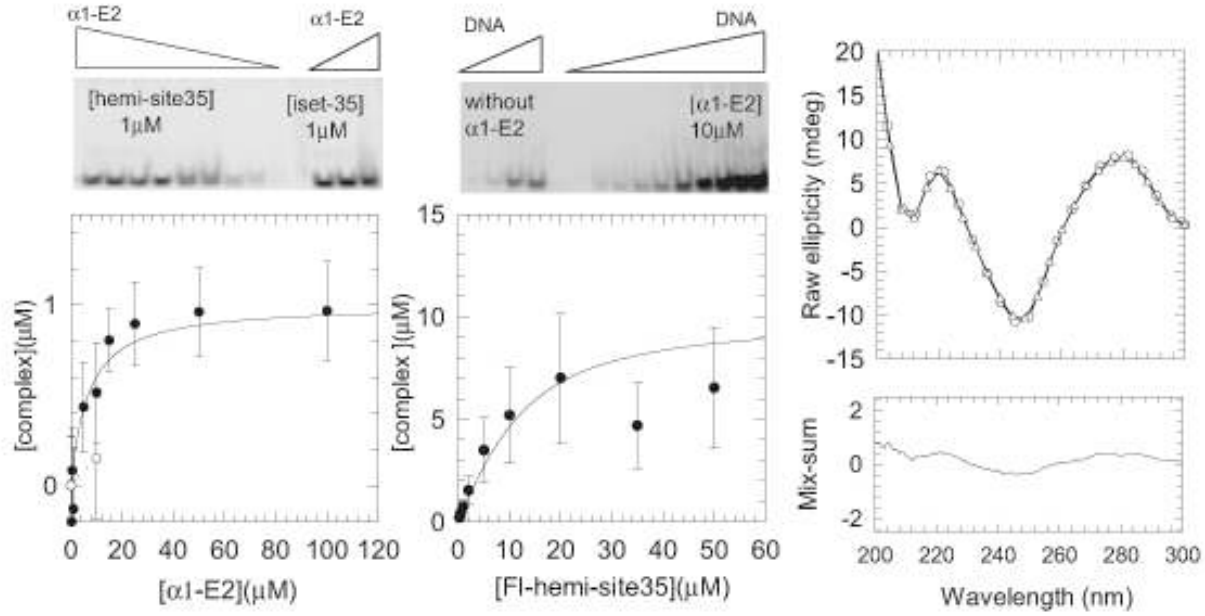


Figure 7