A quasi-spontaneous amyloid route in a DNA binding gene regulatory domain: The papillomavirus HPV16 E2 protein

Diana E. Wetzler, Eduardo M. Castaño and Gonzalo de Prat-Gay

Protein Sci. 2007 16: 744-754
Access the most recent version at doi:10.1110/ps.062594007

A quasi-spontaneous amyloid route in a DNA binding gene regulatory domain: The papillomavirus HPV16 E2 protein

DIANA E. WETZLER, EDUARDO M. CASTAÑO, AND GONZALO DE PRAT-GAY
Instituto Leloir and CONICET, Buenos Aires 1405, Argentina

(RECEIVED October 3, 2006; FINAL REVISION November 29, 2006; ACCEPTED December 22, 2006)

Abstract

The DNA binding domain of papillomavirus E2 proteins is at the center of the regulation of gene transcription and replication of the virus. Its unique fold consists of a β-barrel domain that combines an eight-stranded dimeric β-barrel core interface with two symmetrical DNA binding α-helices and other two helices, packed against the central barrel. Treatment with low amounts of trifluoroethanol readily leads to a mostly β-sheet oligomeric species, with a loss of near-UV circular dichroism signal and increase in its ANS binding capacity, indicating that buried hydrophobic surfaces become accessible to the solvent. This species subsequently undergoes a slow transition into amyloid aggregates as determined by light scattering and Congo red and thioflavin T binding. Electron microscopy shows short amyloid fibers with a curly aspect as the end product. The amyloid route is completely prevented by addition of stoichiometrical amounts of specific DNA, strongly suggesting that unfolding of the DNA binding α-helix is required for the formation of the intermediate. The slow nature of this expanded β-oligomeric species and the availability of several different conformational probes make it an excellent model for investigating amyloid mechanisms. The mild perturbation required for entering an amyloid route is indicative of a preexisting equilibrium. Oligomerization processes are required for the assembly of transcription initiation and DNA replication machineries, where proteins from different viruses must come together with host cell proteins. The E2 protein is a virus-encoded multifunctional master regulator that may exert one of its multiple functions through its ability to oligomerize.

Keywords: DNA binding domains; protein structure/folding; circular dichroism; fluorescence; kinetics; amyloid; β-oligomers; E2 protein; HPV

Supplemental material: see www.proteinscience.org

It is now widely accepted that most proteins may form amyloid-type structures given the appropriate conditions, which include modifications in the chemical and physical parameters of the environment (Chiti et al. 1999). However, some proteins show a stronger tendency to form these structures, based on amino acid sequence and composition (Chiti et al. 2000; Wright et al. 2005). The process of amyloid formation is related to an extended number of human pathologies; more than 20 soluble proteins are involved in known amyloidosis, and it is believed that there are many misfolding-related diseases that have not yet been identified as such (Dobson 2005). Amyloid fibrils have been formed in vitro from disease-associated as well as from disease-unrelated proteins and peptides, and despite having different folding topologies and characteristics, their aggregated forms have many

Reprint requests to: Gonzalo De Prat-Gay, Instituto Leloir and CONICET, Patricias Argentinas 435, Buenos Aires 1405, Argentina; e-mail: gpratgay@leloir.org.ar; fax: 11-54-48652246.

Article and publication are at http://www.proteinscience.org/cgi/doi/10.1110/ps.062594007.

properties in common (Uversky and Fink 2004; Dobson 2005). The structural features of amyloid fibers largely rely on main chain interactions, leading to the characteristic repetitive β-sheet pattern, and are thus generic to a large extent. Investigating the overall mechanistic events that lead to amyloid formation, thus, can benefit from the use of a number of models representing different ''starting points,'' irrespective of a linked pathology.

Human papillomaviruses are worldwide distributed pathogens that can cause epithelial lesions and cellular proliferation that in many cases lead to malignant transformation. Such is the case of the HPV16 strain, responsible for >60% of HPV-related cervical cancers, a major cause of death among women (Bosch et al. 2006). Papillomaviruses have a very small DNA genome (8 kbp) and must therefore rely on cellular machinery for most processes (Howley 1996). E2 is a key protein responsible for the positive and negative control of all viral genes, depending on the cellular environment and differentiation stage of the epithelium (Kalantari and Bernard 2006). In addition, it plays essential roles in DNA replication and episomal DNA mitotic migration (Chow and Broker 2006) and generates a strong immune response in patients (Lin et al. 1993; Davidson et al. 2001) and mice (Cerutti et al. 2001). It is being explored for its use in vaccines (Corona Gutierrez et al. 2004) and is an obvious candidate for antiviral drug design (Greenfield and Cuthill 2001).

E2 protein is a ca 400 amino acid polypeptide with conserved N-terminal transactivation and DNA binding domains separated by a variable region (McBride et al. 1991). The structure of the DNA binding C-terminal domain (E2C) was solved initially for the bovine-infecting virus and uncovered a novel folding topology, the dimeric β-barrel, consisting of a central eight-stranded barrel composed of four strands from each symmetric monomer with two helices packed against each monomer, one being the DNA binding domain (Hegde et al. 1992). Since then, several structures have been solved for different types (Hegde 2002) and they have shown an almost identical topology, only shared so far by the DNA binding domain of the Epstein Barr nuclear antigen 1, the origin binding protein responsible for the virus replication (Bochkarev et al. 1996), which shares no sequence homology, DNA sequence specificity, or phylogenetic relationship with papillomavirus E2.

We have been involved in the characterization of the folding and DNA binding mechanisms of HPV16 E2C. The dimer undergoes a concerted dissociation and unfolding with no observable intermediates at equilibrium (Mok et al. 1996b; Foguel et al. 1998), leading to a denatured state with residual local structure (Mok et al. 2000). A nonnative monomeric intermediate is observed in the kinetic folding pathway (Mok et al. 1996a) but not on the unfolding direction (Prat-Gay et al. 2005). The recently determined solution structure of HPV16 E2C reveals a highly dynamic but otherwise formed DNA binding helix (Nadra et al. 2004) in agreement with previous results on the HPV31 domain (Liang et al. 1996). The equilibrium and kinetic binding mechanism of HPV16 E2C to its target DNA was described (Ferreiro et al. 2000; Ferreiro and de Prat-Gay 2003; Ferreiro et al. 2005), and we recently determined the solution structure of the domain bound to DNA (Cicero et al. 2006).

In the present work we found that the marginal stability of HPV16 E2C at pH 5.6 allows low concentrations of the cosolvent 2, 2, 2-trifluoroethanol (TFE) to accelerate the entrance of the protein into an amyloidogenic pathway at room temperature, ultimately leading to the assembly of insoluble amyloid fibers. We observed and characterized a slow forming expanded β-oligomeric intermediate in controlled conditions, at the expense of the local unfolding of the major DNA binding helix. This unique topology, the dimeric β-barrel, can be used as a model to study fundamental mechanisms behind the formation of amyloid intermediates and fibers, and the readiness of the reaction suggests a possible occurrence in vivo.

Results

Effect of TFE on the conformation of HPV16 E2C

The conformational changes of E2C caused by solvent modifications were monitored by far- and near-UV circular dichroism spectroscopy (CD). The far-UV CD spectrum of E2 shows two characteristic negative bands at 212 and 225 nm (Mok et al. 1996b) in agreement with its topology, which combines an eight-stranded dimeric β-barrel core interface with two symmetrical DNA binding α-helices and other two helices at the opposite side, all packed against the central barrel. The near-UV CD spectrum shows a maximum ~275 nm due to its aromatic amino acids that sense the asymmetrical environment, indicative of the compact tertiary structure of the protein (Fig. 1A,B).

No significant changes are observable either in far- or near-UV spectra upon addition of little amounts of TFE, up to 7.5% (v/v) of TFE. Spectral changes indicative of modifications in secondary and tertiary structure are observed between 7.5% and 30% (v/v) of TFE (Fig. 1C). The spectra of a sample with 12.5% TFE are shown in Figure 1; the far-UV spectrum uncovers a minimum ~216 nm, typical of β-sheet conformation (Fig. 1A, triangles), while the intensity of the near-UV maximum decreases, indicating the partial loss of the tertiary structure (Fig. 1B, triangles). At 50% TFE, the far-UV spectrum presents two negative bands at 208 and 222 nm that indicate the stabilization of extensive α-helical conformation, clearly of nonnative nature (Fig. 1A, squares).

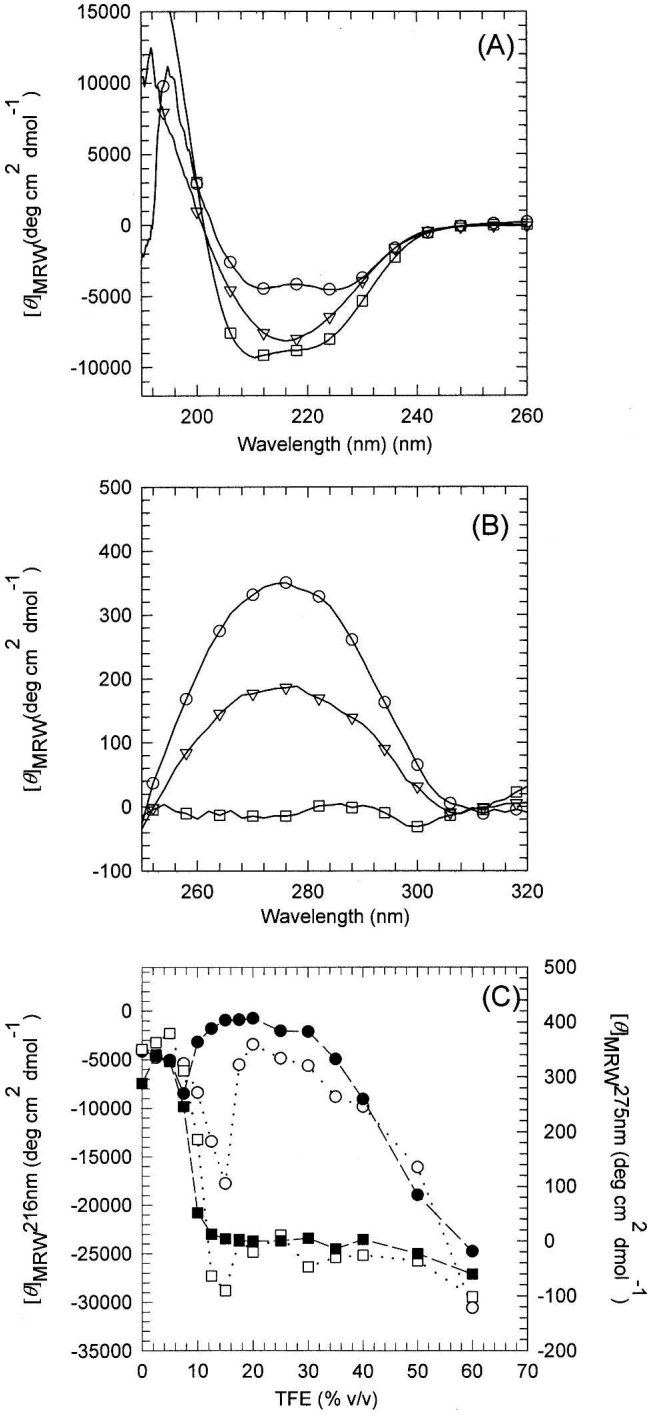

Figure 1. CD Spectra of E2C (10 μM) in acetate buffer (25 mM, pH 5.6) at different TFE concentration. (*A*) far-UV CD spectra and (*B*) near- UV CD spectra. Circles, triangles, and squares correspond to 0%, 12.5%, and 50.0% (v/v) of TFE, respectively. (*C*) Molar ellipticity at 216 nm (circles) and at 275 nm (squares) as a function of TFE concentration. Open symbols correspond to measurements after 2 h of incubation and filled symbols to overnight-incubated samples.

At these cosolvent concentrations, the near-UV aromatic band completely disappears, indicating that the aromatic chromophores are differently solvated, which in turn suggests a loss in tertiary compactness.

A detailed investigation of the TFE transition is shown in Figure 1C. After a 2-h incubation period there is a biphasic change where the maximum formation of β-sheet takes place at 15% TFE. From 10% to 30% there is a loss of signal that is caused by aggregation up to a certain value (~35%) at which the cosolvent starts stabilizing α-helix in regions that are flexible or do not have propensity for that conformation (Fig. 1A). Overnight incubation leads to aggregation of the samples within the TFE range where the β-sheet content is increased as the complete loss of signal shows (Fig. 1C), also evident from visual inspection. The compact tertiary structure is lost within two hours and the signal is lost overnight in samples containing higher amounts of TFE, once again, because of the aggregation observed (Fig. 1C, squares).

Kinetic of formation of β-sheet species

The slow change in CD spectra together with the even slower appearance of an aggregate prompted a kinetic analysis of the β-sheet species as the next step. Formation of β-sheet species also takes place at pH 7.0 in MES buffer, but the half-life time is threefold slower (spectra not shown). For practical reasons we carried out the experiments at pH 5.6 and fixed the concentration of TFE to 12.5%. The formation of the β-sheet species was followed by far- and near-UV CD spectroscopy at 216 and 275 nm, respectively, after treatment with TFE. Both traces are superimposable, strongly suggesting that the increase in β-sheet and the loss of the tertiary structure occur simultaneously in this sample (Fig. 2A).

The rate of β-sheet formation is dramatically decreased with the increase in protein concentration (Figs. 2B, 5B, see below). This observation indicates that an oligomeric β-sheet-containing intermediate is being formed. However, analysis of the lifetimes indicates that the reaction is independent of the concentration above 50 μM, which strongly suggests that a first-order folding or rearrangement reaction becomes the rate-limiting step (Fig. 2B, inset). At this point we can conclude that the intermediate species are of oligomeric nature.

The slow aggregation of the β-sheet oligomeric (E2C-βOs) intermediates was monitored by light scattering absorbance at 350 nm under constant agitation in a spectrophotometer. The aggregation shows a lag phase not unusual for these types of processes (Fig. 3, inset), and after ~10 days the curve shows a plateau, indicating that the process reaches a steady state (Fig. 3). No changes in light scattering were observed in the samples containing either 0% or 50% of TFE, indicating absence

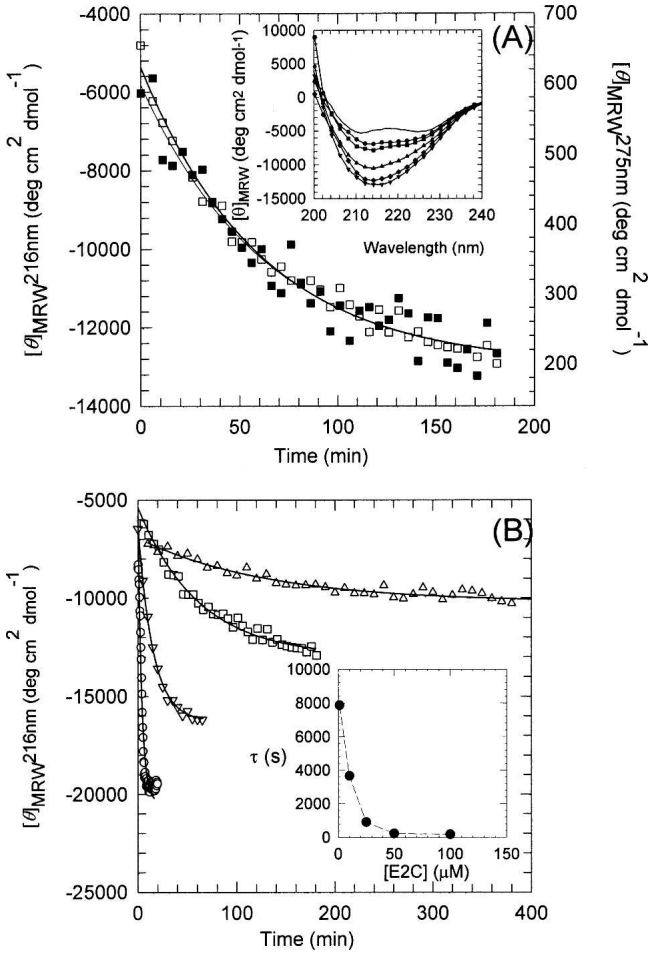

Figure 2. Kinetic of β-oligomer formation followed by CD. (*A*) The same sample followed at a fixed wavelength. Open squares correspond to molar ellipticity at 216 nm and filled squares to 275 nm. (*Inset*) CD spectra of E2C 10 μM in acetate buffer (25 mM, pH 5.6). The spectra correspond to the sample before the addition of TFE (continuous line), immediately (circles), and after 10 (squares), 50 (up triangles), 120 (diamonds), and 180 min (down triangles) after the addition of 12.5% of TFE. (*B*) Measurements at different E2C concentrations: up triangles, 1 μM; squares, 10 μM; down triangles, 25 μM; and circles, 50 μM. Lines correspond to exponential decay fits. (*Inset*) Dependence of the lifetime of β-sheet intermediates formation with E2C concentration.

of aggregation (Fig. 3, inset). We can conclude that there is no aggregation in samples where the β-sheet oligomeric intermediates are not observable by CD. In addition, the decrease of soluble material after 10 d indicates that the aggregates are formed at the expense of the β-sheet oligomers, stressing the transient nature of the latter.

Amyloid nature of β-sheet oligomeric intermediates and subsequent aggregates

In order to characterize the transient oligomeric species, and taking advantage of its slow nature, we tested them for binding of amyloid dyes, thioflavin T (ThT), and Congo red (CR), indicative of repetitive β-sheet structures. ThT does not bind to the E2C dimer in solution as no spectral change is observed (Fig. 4A), but a fluorescence increase is observed for the β-sheet oligomeric species, which indicates binding of the dye. Samples that were incubated for 30 d show a maximum binding of ThT (Fig. 4A, upward triangles). A similar situation is observed for CR binding; the spectrum of dimeric untreated E2C is identical to the CR spectrum without protein (Fig. 4B), and both E2C-βOs and the final aggregates show a marked spectral shift, typical of repetitive β-sheet structures found in amyloids (Fig. 4B; Klunk et al. 1999; LeVine III 1999).

As an additional probe, we used the hydrophobic surface binding dye ANS, that was shown to bind to dimeric E2C (Mok et al. 1996a), as it is often found in nucleic acid binding proteins. Binding of ANS to E2C is expressed by the characteristic blue shift and the increase in ANS fluorescence emission. The soluble β-sheet oligomers show a substantially increased fluorescence, indicative of the existence of hydrophobic surface areas exposed to the solvent. The 30-d incubated sample shows a substantially increased binding of ANS based on intensity, but the maximum wavelength is similar to that of the soluble β-sheet oligomers (Fig. 4C) (Ferrao-Gonzales et al. 2005).

Physical characterization of E2C-βOs and fibrillar species

In order to evaluate the sizes of the species involved in the process of amyloid formation we made use of dynamic

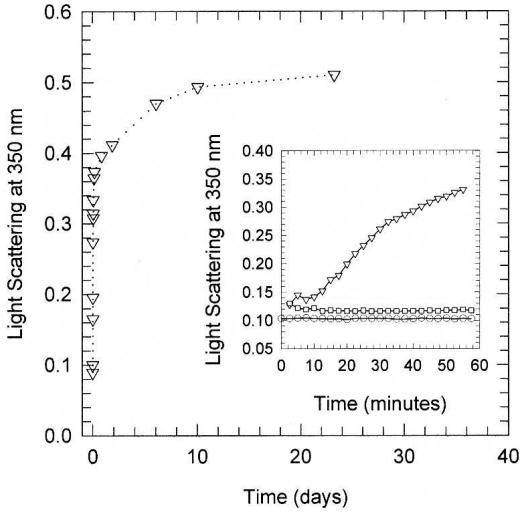

Figure 3. Kinetic of aggregation followed by light scattering. E2C 100 μM in acetate buffer (25 mM, pH 5.6). Circles, triangles, and squares correspond to 0%, 12.5%, and 50.0% (v/v) of TFE, respectively. (*Inset*) Short time measurements.

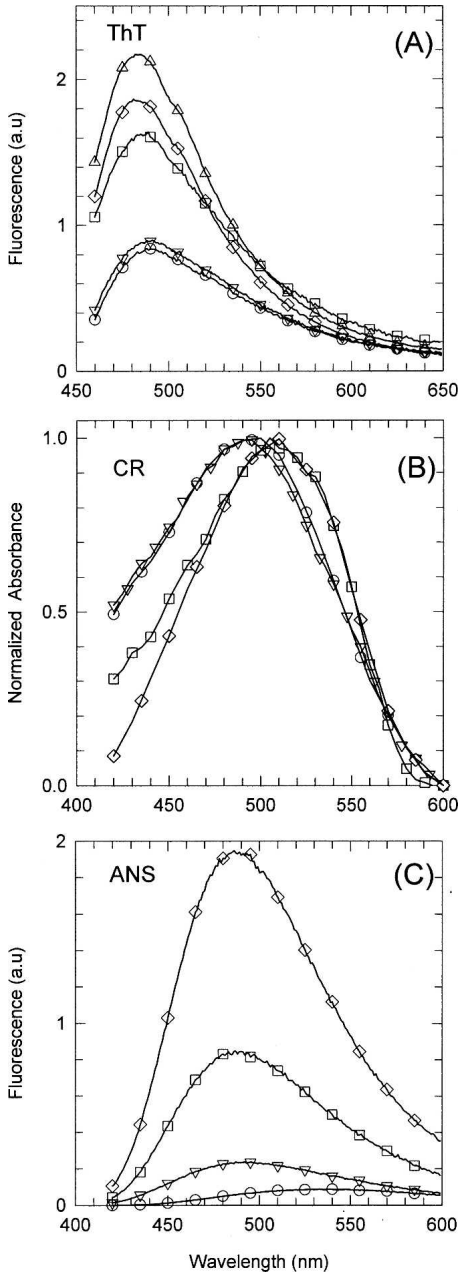

Figure 4. Thioflavin, Congo red, and ANS binding assays. (*A*) ThT fluorescence spectra, (*B*) CR absorbance spectra, (*C*) ANS fluorescence spectra. Curves correspond to the dye without protein (circles), E2C (down triangles), β-sheet intermediates (squares), E2C samples after 10 d of incubation with 12.5% (v/v) of TFE (diamonds), and E2C samples after 30 d of incubation with 12.5% (v/v) of TFE (*up* triangles).

light scattering (DLS). The DLS of a solution containing 25 μM E2C was measured before and at different times after addition of 12.5% TFE. Before addition, particles of ~5 nm were observed corresponding to the E2C dimer, in agreement with its expected size (Hegde et al. 1992). After one lifetime elapsed, particles ~80 nm could be observed. The size distribution measured 1 h after the TFE addition was also ~75–80 nm but it was narrower than the previous one (Fig. 5A). The existence of smaller intermediate species cannot be discarded, as their low concentration and their short lifetime might be below the detection limit of the equipment, or not analyzable because of the time interval required to averaged measurements. Figure 5B shows the time at which the sizes were measured, indicated in the kinetics of β-sheet formation in identical condition.

Electron microscopy examination of negatively stained samples revealed that, after 10 d of incubation in 12.5% TFE, E2C adopted a heterogeneous morphology, showing

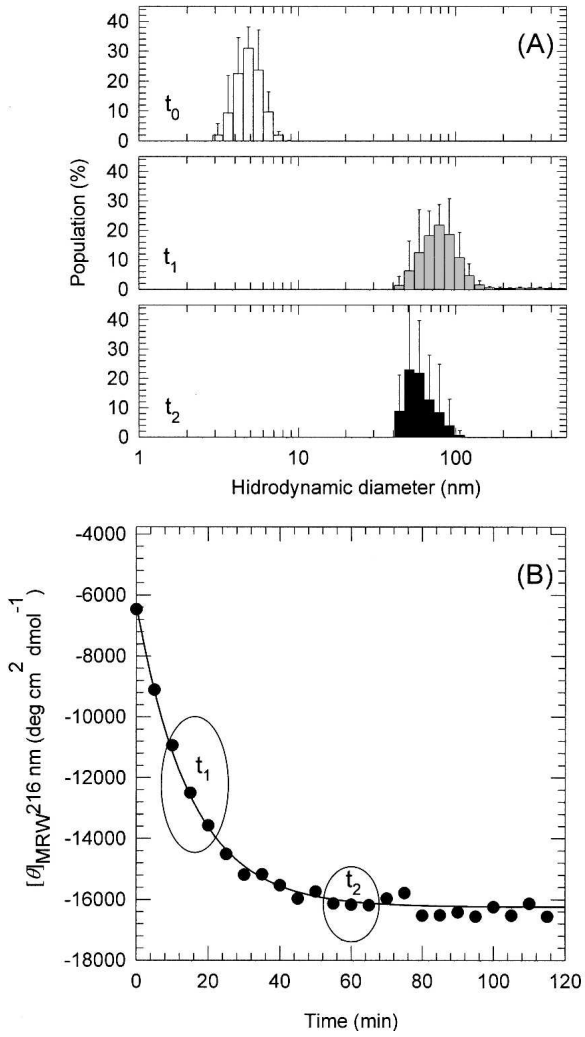

Figure 5. Size distribution measured by DLS. (*A*) E2C 25 μM in acetate buffer (25 mM, pH 5.6). The *upper* panel shows the size distribution profile of the sample before the addition of TFE. The other two panels correspond to the sample 15 and 60 min after the TFE addition (12.5% v/v). (*B*) Decay profile of the same sample followed by CD at 216 nm. The points marked as t_1 an t_2 correspond to the time ranges where the DLS measurements were performed.

granular structures (white arrows) and small annuli (arrowheads) with diameters of ~5 nm and 10 nm, respectively (Fig. 6A). In addition, short fibers of 7–8 nm in width and 30 nm in length (black arrows) were also observed at the shortest incubation time. After a 30-d incubation, straight fibrils up to ~300 nm either free or arranged in bundles could be observed (Fig. 6B).

Specific DNA blocks the amyloid route

The effect of specific DNA on the amyloid route of E2C was investigated by CD spectroscopy. Spectra of E2C samples without DNA and with substoichiometrical, stoichiometrical, and with an excess of specific DNA were recorded one minute and four hours after the addition of 12.5% TFE (Fig. 7). In the absence of DNA, the expected increase in negative ellipticity in the β-sheet region of the CD spectrum is observed (Fig. 7A). A small change is observed with the addition of 0.5 molar equivalents, but no β-sheet distinguishable due to the DNA contribution (Fig. 7B). No change in the CD spectra was observed in samples containing stoichiometrical (Fig. 7C) or in the presence of an excess of specific DNA (not shown). In order to confirm the absence of repetitive β-sheet in the oligomers and late aggregates, we used CR binding assays; no shift in CR absorbance spectra was observed in the samples containing stoichiometrical (Fig. 7F) and an excess of specific DNA against the shift that was observed in the absence of it (Fig. 7D). Thus, the formation of the oligomeric intermediates and subsequent fibrillar aggregates is completely prevented by addition of stoichiometrical amounts of a DNA duplex containing the specific E2 binding site.

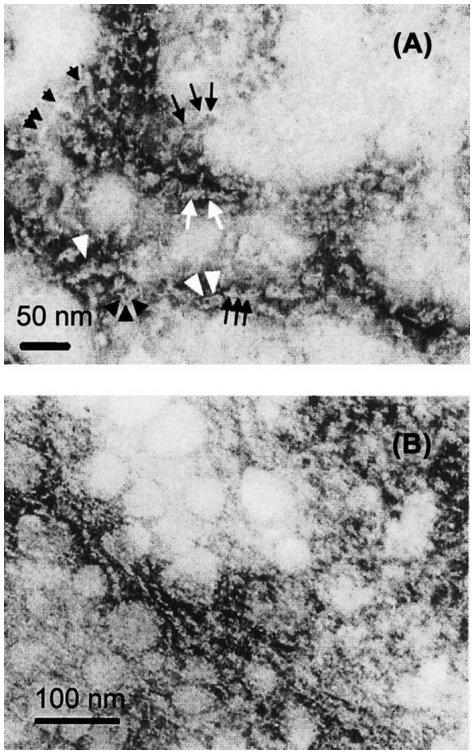

Figure 6. Electron micrograph of the samples. Both micrographs correspond to E2C (100 μM) sample in acetate buffer (25 mM, pH 5.6) and 12.5% (v/v) of TFE, incubated for different periods of time, (*A*) 10 d, (*B*) 30 d. E2C adopted a heterogenous morphology, showing granular structures (white arrows) and small annuli (arrowheads) with diameters of ~5 nm and 10 nm, respectively. In addition, short fibers of 7–8 nm in width and 30 nm in length (black arrows).

Discussion

The DNA binding domain of the HPV16 E2 transcriptional regulator bears an unusual folding topology, the dimeric β-barrel, where dimer interface is formed by two half-barrels contributed by each monomer. The domain is stable but fairly dynamic; in particular, the DNA recognition helix exchanges its amidic protons very fast (Liang et al. 1996; Nadra 2005). Besides an evident role of domain plasticity in DNA recognition, this feature implies a conformational equilibrium of the helix in conditions where the core, comprised by the dimer interface, remains stable. We have previously investigated the folding transition of this domain (Mok et al. 1996a,b; Prat-Gay et al. 2005), and in this work we show that mild perturbation of this conformational equilibrium leads to an amyloid route. The ability to form amyloids is thought to be a generic feature in polypeptides (Chiti et al. 1999), and the E2C domain provides a valuable model for investigating essential mechanisms. In addition, it is a unique example of a DNA binding and transcription regulatory protein with almost spontaneous amyloid route, at least in solution.

The addition of small amounts of TFE as cosolvent, at physiologically compatible pH and temperatures and relatively low protein concentrations (10 μM), leads to the formation of a soluble and oligomeric intermediate species with a substantially increased β-sheet content. This species is further slowly converted into insoluble amyloid fibers, and the overall nature of the process and probes available anticipates the possibility of an extensive dissection of the fundamental reaction mechanism. Similar results can be found in the literature for proteins of different topologies, being these disease-associated or not. Two examples of β-barrel but otherwise monomeric proteins, the acidic fibroblast growth factor (Srisailam et al. 2003) and α-chymotrypsin (Pallares et al. 2004), present a similar behavior after TFE addition. The addition of little or moderate amounts of TFE stabilizes the formation of β-sheet conformers, and the addition of higher amounts of alcohol stabilizes an extensive α-helical conformation. Other examples with proteins of

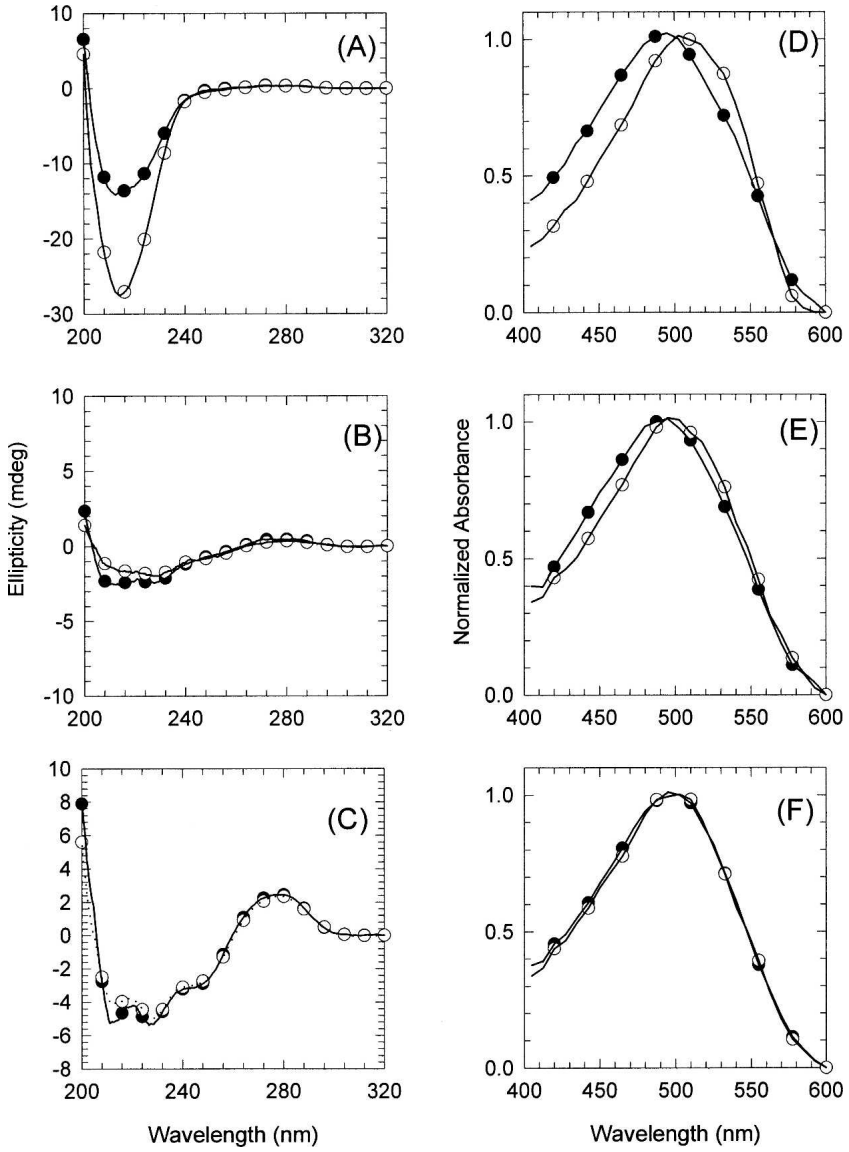

Figure 7. Effect of specific DNA. (*A–C*) UV-CD spectra of E2C 10 μM in acetate buffer (25 mM, pH 5.6) and 12.5% (v/v) of TFE with different concentration of specific DNA: 0, 5, and 10 μM, respectively. (*D–F*) CR absorbance spectra of the same samples. Filled circles correspond to measurements 1 min after the addition of TFE and open circles to 4 h after the addition.

different topologies are the cases of Transthyretin (TTR) (predominately β-sheet structure) (MacPhee and Dobson 2000) and the peptide LYS(11–36) (α+β structure) (Liu et al. 2004); in the second case the addition of a higher concentration of TFE is needed to stabilize the β-sheet conformers but the overall behavior is similar (MacPhee and Dobson 2000).

The concentration dependence of the β-sheet formation clearly indicates oligomeric species. However, a loss in the concentration dependence >50 μM suggests that a first-order reaction becomes rate-limiting, possibly due to a post-oligomerization conformational rearrangement. The E2C-βOs display a loss of the aromatic signal in the near-UV CD, which is suggestive of a partial loss in tertiary structure. This agrees with an increase in ANS binding, which, together with a substantial secondary structure content, indicates ordered molten globule-like properties where the solvent infiltrates the interior of the protein, but with a substantial amount of structure retained. The E2C-βOs thus appear to lose compactness and expose hydrophobic regions to the solvent that leads to subsequent aggregation processes. Binding of CR and ThT by the E2C-βOs indicates they contain repetitive β-sheet structures, which is reasonable, considering that the end point of the reaction are amyloid fibers. Clearly, the formation of the repetitive amyloid-like β-sheet already

takes place at least in part in the soluble oligomeric intermediate, and it is not parallel to the ordered aggregation into fibers.

Formation of prefibrillar soluble oligomers precedes amyloid fiber formation in the Alzheimer Aβ peptides, where the precursor proteolytic fragments are disordered but require some residual structure to form the oligomers (Chen and Glabe 2006). On the side of full-length folded proteins, TTR constitutes an example of an amyloidogenic protein found in inherited pathologies (Kelly 1996). Soluble prefibrillar oligomers were found when TTR was taken to an acid A-state of molten globule-like characteristics, where the β-sheet content was increased as a result of the self-assembly process (Lindgren et al. 2005). These and other examples of amyloids either from disordered fragments and peptides or from folded proteins that were perturbed to enter an amyloid route have in common the prefibrillar soluble oligomers with amyloid-type properties, like binding ThT and CR dyes and its secondary structure. These oligomers appear in both general routes (from peptides or folded proteins) making the entrance into fiber formation converge, although they can be bypassed in cases such as the Aβ peptides (Bitan et al. 2003).

Our results with the HPV16 E2C domain show a similar molten globule-like nature of the β-sheet oligomeric intermediates as in the case of proteins of different topologies like TTR, α-synuclein, acylphosphatase (Lansbury Jr. 1999), all of them sharing a structured prefibrillar oligomeric intermediate formation in the fibrillization process. It remains to be established whether unfolded or partially folded monomers are the direct precursors of the amyloid prefibrillar intermediates. In any case, the nature of the fold does not allow a folded monomer of E2C to be stable in solution for long enough to allow full characterization (Hegde et al. 1992; Mok et al. 1996b; Prat-Gay et al. 2005). What is clear from the CD experiments is that no global unfolding is required for entering the amyloid route. The large increase in β-sheet signal in the far-UV CD indicates not only the increase in β-sheet formation but that there are changes in the nature of those structures, plus the disappearance of helix. We cannot discriminate at this point whether all the barrel β-sheets remain in exactly their native conformation. We hypothesize an intermediate situation: Some native structure is retained and some nonnative β-sheet structure is formed.

The reaction leading to amyloid fibrillar aggregates displays a characteristic lag phase, indicative of a nucleation process (Dobson 2005). The appearance of the fibrillar aggregates at the expense of the disappearance of the soluble E2C-βOs indicates the intermediate nature of the latter. No amyloid formation occurs in the absence or with 50% TFE, conditions where the β-sheet oligomerization transition is not observed. Electron microscopy (EM) shows short fibers at 10 d of incubation and larger and curly fibers at 30 d of incubation. The growth in the size correlates with the increase of CD and ThT binding in that time interval, which must necessarily be indicative of more repetitive β-sheet structures having polymerized. Size discrepancy between EM and DLS measurements may be due to the detection limit of the second technique. Particle scattering intensity is proportional to molecular size and, in spite of that, the initial protein concentration is over the detection limit (0.5 mg/mL for this protein size). Low concentrations of small oligomers that can appear at the first steps of the kinetics might be below this detection limit. The larger size oligomers are easily detected and can mask the presence of small oligomers that are likely to exist at earlier steps. On the other hand, when EM is used, these larger oligomers could be disassembled into smaller subunits due to the sample dilution required to perform these measurements. Nevertheless, considering the results of both techniques it can be concluded that soluble oligomers ~30–80 nm are formed prior to fibril deposition.

Several lines of evidence made us consider the local unfolding of the DNA binding helix as the conformational trigger leading into the amyloid route. Large amounts (≥35%) of TFE do not allow the formation of the β-sheet oligomeric species that ultimately lead to amyloid fibers, and a substantial increase in α-helix, most likely of nonnative nature, is observed. TFE will largely stabilize existing α-helix structures; in particular, it is known that the DNA binding helix, although fully formed and observed in NMR and X-ray structures, is dynamic and exchanges its backbone amide protons rapidly with the solvent. In addition, the inhibition of the formation of the E2C-βSOs by phosphate agrees with previous observations on a large stabilization caused by phosphate (Lima and Prat-Gay 1997), which appears to be based on surface electrostatic interactions at the many positive side chain residues in the DNA binding helix. Additionally, it has been determined that an important flexible epitope of E2C comprises at least half of the DNA recognition helix (Cerutti et al. 2006).

The fact that DNA prevents β-sheet formation is clear from the lack of change in CD and the disappearance of CR binding properties. Thus, DNA prevents oligomerization, β-sheet formation, and, as a consequence, the ultimate formation of the amyloid fibers. We conclude that the native state has a conformational equilibrium at the DNA binding helix in which a partly unfolded state is scarcely populated, but this equilibrium is shifted upon the addition of small amounts of TFE. The DNA binding helix is stabilized by large amounts of TFE, phosphate, higher pH, and, as expected, by interaction with its specific DNA target. The β-sheet content is increased at

the expense of the disappearance of the DNA binding helix, leading to the formation of soluble oligomers (E2C-βOs), which slowly yield short amyloid fibers. The binding of DNA may prevent entrance into the amyloid route by a double effect: the depopulation of the partially unfolded precursor and the stabilization of the native conformation once the DNA–protein complex is formed. Similar results which employ ligands to avoid fibril formation can be found in the literature, such as the reduction of the amyloidogenicity of acylphosphatase by specific binding of ligands to the native conformation (Chiti et al. 2001) and protein stabilization via interaction with thyroxine as in the case of stabilization of TTR tetramer (Miroy et al. 1996).

It is noticeable that the DNA binding helix is present and largely populated in most E2C domains studied to date, and that it can undergo a local folding transition that leads to β-sheet structure, at least in the HPV16 strain. Thus, the dynamic properties of this helix have evolved to display a certain tension or plasticity that is important for DNA recognition but allow the existence of other equilibria in solution that may lead to an amyloid route. The nature of the helix of HPV16 E2C is being investigated in depth in our laboratory, in particular in comparison with more rigid, structural α-helices. It would be interesting to know how a peptide spanning the amino acid sequence of the DNA binding helix of HPV16 E2C behaves in the absence of the protein core, and what the transitions that allow an α-helix to β-sheet transition might be. However, we cannot rule out that the DNA binding α-helix stabilizes the protein globally. What is certain is that the destabilization of this helix is what triggers the process. We shall wait for a complete evaluation of the local versus global stabilization in the domain by NMR methods and how this affects the amyloid route. In any case, our main hypothesis at this stage is that the local destabilization of the helix is essential for attaining a yet undefined intermediate structure that leads to an amyloid route via the β-sheet oligomers, acting as a trigger, and this may have consequences for other systems.

The possibility of triggering an amyloid process in a small model protein with a mild perturbation and the diverse number of analytical probes that emerge from this work will allow a deeper insight into the basic mechanisms of amyloid formation. The readiness of formation and the eventual manipulation of the conditions to accumulate the E2C-βOs will allow us to characterize these species as models of prefibrillar oligomeric amyloid intermediates, which are currently believed to be the main responsible for cell toxicity in human diseases (Hardy and Selkoe 2002; Dobson 2005). Moreover, prefibrillar oligomers of different unrelated proteins were shown to display toxicity levels comparable to those of the Alzheimer's Aβ-peptide (Dobson 2005).

The recent report of a functional amyloid formation supports the hypothesis that ''amyloid'' is an evolved quaternary protein fold used by organisms from bacteria to humans not exclusive of misfolding diseases or pathological situations (Fowler et al. 2006). Although there is no evidence of biological relevance for the HPV life cycle, a functional significance or effect, either direct or as a side reaction, cannot be discarded due to the existing equilibrium between the native-like form and the β-sheet oligomers and the mild conditions required to perturb it. In a recent study (Dao et al. 2006), E2 was shown to be in substantial amounts associated along the mitotic spindle fibers throughout the different phases of mitosis, which suggest some degree of polimerization by the protein. Moreover the region mapped for this association was a β-strand of the C-terminal domain, our object of study in this work. Finally, given the key role of E2 in the control of viral gene expression and genome replication, information on the conformational equilibria of its DNA binding domain will render useful information for antiviral drug design.

Materials and Methods

Protein expression and purification

The recombinant HPV16 E2C (E2C) was expressed in *Escherichia coli* BL21(DE3) and purified as described previously (Mok et al. 1996a).

DNA synthesis

Double-stranded 18-bp oligonucleotides containing E2 recognition sequence (site 35 in HPV16 genome) was prepared as follows: Single-stranded oligonucleotides were purchased, HPLC-purified, from Integrated DNA technologies. Annealing was performed by mixing equal amounts of oligos in 10 mM Bis-Tris-HCl buffer, pH 7.0, and 50 mM NaCl, incubating 5 min at 95°C, and slowly cooling to 25°C for 16 h.

Circular dichroism and absorbance spectroscopy

CD measurements were carried out on a Jasco J-810 spectropolarimeter. CD spectra in the far- and near-UV region were collected using a Peltier temperature-controlled sample holder at 25°C in a 0.1-cm path length cell in 25 mM buffer solution. Protein was assayed at 1, 10, 25, 50, and 100 μM in different TFE concentrations. The aggregation was followed by absorption spectroscopy. E2C 100 μM acetate buffer (25 mM, pH 5.6) sample aggregation with different TFE concentrations was followed by monitoring the UV absorbance increase at 350 nm as a function of time with continuous stirring.

Congo red binding

A 450–650-nm spectrum of the buffer (50 mM phosphate, pH 7.0) plus a protein aliquot solution (with a final protein

concentration 10 μM) was subtracted from the spectrum of the buffer plus protein solution plus CR (5 μM). The same procedure was employed using fresh E2C, β-sheet intermediates, and fibers formed in 25 mM acetate buffer at pH 5.6 with 12.5% (v/v) of TFE.

Thioflavin fluorescence

Thioflavin T assays were carried out using aliquots of preincubated protein in 25 mM acetate buffer at pH 5.6 with 12.5% (v/v) of TFE. Measurements were performed in a final protein concentration of 10 μM in 50 mM glycine buffer at pH 9.0 and Thioflavin T at 20 μM. Emission spectra were recorded on an Aminco-Bowman spectrofluorimeter using 435-nm excitation wavelength. The same procedure was assayed with fresh E2C, β-sheet intermediates, and fibers.

ANS fluorescence emission

Fluorescence emission spectra for ANS binding were carried out using 350-nm excitation wavelength in the Aminco-Bowman spectrofluorimeter. Protein concentration was kept at 10 μM and ANS concentration was 100 μM. The same procedure was applied to fresh E2C, β-sheet intermediates, and fibers formed in 25 mM acetate buffer at pH 5.6 with 12.5% (v/v) of TFE. Kinetic measurements were performed with aliquots of the sample incubated with TFE at different time intervals, following ANS fluorescence at a fixed wavelength of 490 nm.

Dynamic light scattering

Measurements were performed in 25 mM acetate buffer at pH 5.6 before and after the addition of 12.5% (v/v) of TFE, using a Zetasizer Nano S DLS device from Malvern Instruments (Malvern). Protein concentration was kept at 25 μM in order to be over the instrument detection limit (minimum concentration recommended 0.5 mg/L for particle size ≤10 nm) and to control the oligomerization rate. Temperature was maintained at 25°C by a Peltier control system. Results were processed employing the software package included in the equipment.

Electron microscopy

Fifteen-microliter water-diluted aliquots of E2C (100 μM) sample in acetate buffer (25 mM, pH 5.6) and 12.5% (v/v) of TFE, incubated for different periods of time, were absorbed for one minute onto a Formvar-coated nickel gird (Electron Microscopy Science) and then stained with 2% uranyl acetate aqueous solution.

Acknowledgments

This work was supported by a Wellcome Trust Grant (GR077355AIA) and by a grant from ANPCyT (PICT 2000 01-08959). D.E.W. holds a post-doctoral fellowship from CONICET. G.P.G. and E.M.C. are Career Investigators from CONICET, Argentine National Science Council.

References

Bitan, G., Kirkitadze, M.D., Lomakin, A., Vollers, S.S., Benedek, G.B., and Teplow, D.B. 2003. Amyloid β-protein (Aβ) assembly: Aβ 40 and Aβ 42 oligomerize through distinct pathways. *Proc. Natl. Acad. Sci.* **100:** 330–335.

Bochkarev, A., Barwell, J.A., Pfuetzner, R.A., Bochkareva, E., Frappier, L., and 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.

Bosch, F.X., Sanjosé, S., Castellsagué, X., Moreno, V., and Muñoz, N. 2006. Epidemiology of human papillomavirus infections and associations with cervical cancer: New opportunities for prevention. In *Papillomavirus research: From natural history to vaccines and beyond* (ed. M. Saveria Campo), pp. 19–40. Caister Academic Press, Wymondham, U.K.

Cerutti, M.L., Centeno, J.M., Goldbaum, F.A., and de Prat-Gay, G. 2001. Generation of sequence-specific, high affinity anti-DNA antibodies. *J. Biol. Chem.* **276:** 12769–12773.

Cerutti, M.L., Ferreiro, D.U., Sanguineti, S., Goldbaum, F.A., and Prat-Gay, G. 2006. Antibody recognition of a flexible epitope at the DNA binding site of the human papillomavirus transciptional regulator e2. *Biochemistry* **45:** 15520–15528.

Chen, Y.R. and Glabe, C.G. 2006. Distinct early folding and aggregation properties of Alzheimer amyloid-β peptide Aβ40 and Aβ42: Stable trimer or tetramer formation by Aβ42. *J. Biol. Chem.* **34:** 24414–24422.

Chiti, F., Webster, P., Taddei, N., Clark, A., Stefani, M., Ramponi, G., and Dobson, C.M. 1999. Designing conditions for in vitro formation of amyloid protofilaments and fibrils. *Proc. Natl. Acad. Sci.* **96:** 3590–3594.

Chiti, F., Taddei, N., Bucciantini, M., White, P., Ramponi, G., and Dobson, C.M. 2000. Mutational analysis of the propensity for amyloid formation by a globular protein. *EMBO J.* **19:** 1441–1449.

Chiti, F., Taddei, N., Stefani, M., Dobson, C.M., and Ramponi, G. 2001. Reduction of the amyloidogenicity of a protein by specific binding of ligands to the native conformation. *Protein Sci.* **10:** 879–886.

Chow, L.T. and Broker, T.R. 2006. Mechanisms and regulation of papillomavirus DNA replication. In *Papillomavirus research: From natural history to vaccines and beyond* (ed. M. Saveria Campo), pp. 53–71. Caister Academic Press, Wymondham, U.K.

Cicero, D.O., Nadra, A.D., Eliseo, T., Dellarole, M., Paci, M., and 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.

Corona Gutierrez, C.M., Tinoco, A., Navarro, T., Contreras, M.L., Cortes, R.R., Calzado, P., Reyes, L., Posternak, R., Morosoli, G., Verde, M.L., et al. 2004. Therapeutic vaccination with MVA E2 can eliminate precancerous lesions (CIN 1, CIN 2, and CIN 3) associated with infection by oncogenic human papillomavirus. *Hum. Gene Ther.* **15:** 421–431.

Dao, L.D., Duffy, A., Van Tine, B.A., Wu, S.Y., Chiang, C.M., Broker, T.R., and Chow, L.T. 2006. Dynamic localization of the human papillomavirus type 11 origin binding protein E2 through mitosis while in association with the spindle apparatus. *J. Virol.* **80:** 4792–4800.

Davidson, E.J., Brown, M.D., Burt, D.J., Parish, J.L., Gaston, K., Kitchener, H.C., Stacey, S.N., and Stern, P.L. 2001. Human T cell responses to HPV 16 E2 generated with monocyte-derived dendritic cells. *Int. J. Cancer* **94:** 807–812.

Dobson, C.M. 2005. An overview of protein misfolding diseases. In *Protein folding handbook* (eds. J. Buchner and T. Kiefhaber), pp. 1093–1113. Wiley-VCH Verlag GMBH & Co. KgaA, Weinheim, Germany.

Ferrao-Gonzales, A.D., Robbs, B.K., Moreau, V.H., Ferreira, A., Juliano, L., Valente, A.P., Almeida, F.C., Silva, J.L., and Foguel, D. 2005. Controlling β-amyloid oligomerization by the use of naphthalene sulfonates: Trapping low molecular weight oligomeric species. *J. Biol. Chem.* **280:** 34747–34754.

Ferreiro, D.U. and 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.

Ferreiro, D.U., Lima, L.M., Nadra, A.D., Alonso, L.G., Goldbaum, F.A., and 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.

Ferreiro, D.U., Dellarole, M., Nadra, A.D., and de Prat-Gay, G. 2005. Free energy contributions to direct readout of a DNA sequence. *J. Biol. Chem.* **280:** 32480–32484.

Foguel, D., Silva, J.L., and de Prat-Gay, G. 1998. Characterization of a partially folded monomer of the DNA-binding domain of human papillomavirus E2 protein obtained at high pressure. *J. Biol. Chem.* **273:** 9050–9057.

Fowler, D.M., Koulov, A.V., Alory-Jost, C., Marks, M.S., Balch, W.E., and Kelly, J.W. 2006. Functional amyloid formation within mammalian tissue. *PLoS Biol.* **4:** 100–107.

Greenfield, I. and Cuthill, S. 2001. Antivirals. In *Human papillomaviruses. Clinical and scientific advances* (eds. J.C. Sterling and S.K. Tyring), pp. 120–130. Arnold Publisher, London.

Hardy, J. and Selkoe, D.J. 2002. The amyloid hypothesis of Alzheimer's disease: Progress and problems on the road to therapeutics. *Science* **297:** 353–356.

Hegde, R.S. 2002. The papillomavirus E2 proteins: Structure, function, and biology. *Annu. Rev. Biophys. Biomol. Struct.* **31:** 343–360.

Hegde, R.S., Grossman, S.R., Laimins, L.A., and 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.

Howley, P.M. 1996. Papillomaviridae: The viruses and their replication. In *Fields Virology* (eds. B.N. Fields et al.), Vol. 2, pp. 2045–2076. Lippincott-Raven, Philadelphia.

Kalantari, M. and Bernard, H.-U. 2006. Gene expression of papillomaviruses. In *Papillomavirus research: From natural history to vaccines and beyond* (ed. M. Saveria Campo), pp. 41–52. Caister Academic Press, Wymondham, U.K.

Kelly, J.W. 1996. Alternative conformations of amyloidogenic proteins govern their behavior. *Curr. Opin. Struct. Biol.* **6:** 11–17.

Klunk, W.E., Jacob, R.F., and Mason, R.P. 1999. Quantifying amyloid by congo red spectral shift assay. *Methods Enzymol.* **309:** 285–305.

Lansbury Jr., P.T. 1999. Evolution of amyloid: What normal protein folding may tell us about fibrillogenesis and disease. *Proc. Natl. Acad. Sci.* **96:** 3342–3344.

LeVine III, H. 1999. Quantification of β-sheet amyloid fibril structures with thioflavin T. *Methods Enzymol.* **309:** 274–284.

Liang, H., Petros, A.M., Meadows, R.P., Yoon, H.S., Egan, D.A., Walter, K., Holzman, T.F., Robins, T., and Fesik, S.W. 1996. Solution structure of the DNA-binding domain of a human papillomavirus E2 protein: Evidence for flexible DNA-binding regions. *Biochemistry* **35:** 2095–2103.

Lima, L.M. and 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.

Lin, Y.L., Borenstein, L.A., Selvakumar, R., Ahmed, R., and Wettstein, F.O. 1993. Progression from papilloma to carcinoma is accompanied by changes in antibody response to papillomavirus proteins. *J. Virol.* **67:** 382–389.

Lindgren, M., Sorgjerd, K., and Hammarstrom, P. 2005. Detection and characterization of aggregates, prefibrillar amyloidogenic oligomers, and protofibrils using fluorescence spectroscopy. *Biophys. J.* **88:** 4200–4212.

Liu, W., Prausnitz, J.M., and Blanch, H.W. 2004. Amyloid fibril formation by peptide LYS (11-36) in aqueous trifluoroethanol. *Biomacromolecules* **5:** 1818–1823.

MacPhee, C.E. and Dobson, C.M. 2000. Chemical dissection and reassembly of amyloid fibrils formed by a peptide fragment of transthyretin. *J. Mol. Biol.* **297:** 1203–1215.

McBride, A.A., Romanczuk, H., and Howley, P.M. 1991. The papillomavirus E2 regulatory proteins. *J. Biol. Chem.* **266:** 18411–18414.

Miroy, G.J., Lai, Z., Lashuel, H.A., Peterson, S.A., Strang, C., and Kelly, J.W. 1996. Inhibiting transthyretin amyloid fibril formation via protein stabilization. *Proc. Natl. Acad. Sci.* **93:** 15051–15056.

Mok, Y.K., Bycroft, M., and de Prat-Gay, G. 1996a. 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.

Mok, Y.K., de Prat-Gay, G., Butler, P.J., and Bycroft, M. 1996b. Equilibrium dissociation and unfolding of the dimeric human papillomavirus strain-16 E2 DNA-binding domain. *Protein Sci.* **5:** 310–319.

Mok, Y.K., Alonso, L.G., Lima, L.M., Bycroft, M., and de Prat-Gay, G. 2000. Folding of a dimeric β-barrel: Residual structure in the urea denatured state of the human papillomavirus E2 DNA binding domain. *Protein Sci.* **9:** 799–811.

Nadra, A.D. 2005. "Estudios estructurales en solución del factor de trascripción de E2C de HPV-16 y su interacción con ADN." Ph.D. thesis, Universidad de Buenos Aires, Buenos Aires.

Nadra, A.D., Eliseo, T., Mok, Y.K., Almeida, C.L., Bycroft, M., Paci, M., de Prat-Gay, G., and Cicero, D.O. 2004. Solution structure of the HPV-16 E2 DNA binding domain, a transcriptional regulator with a dimeric β-barrel fold. *J. Biomol. NMR* **30:** 211–214.

Pallares, I., Vendrell, J., Aviles, F.X., and Ventura, S. 2004. Amyloid fibril formation by a partially structured intermediate state of α-chymotrypsin. *J. Mol. Biol.* **342:** 321–331.

Prat-Gay, G., Nadra, A.D., Corrales-Izquierdo, F.J., Alonso, L.G., Ferreiro, D.U., and Mok, Y.K. 2005. The folding mechanism of a dimeric β-barrel domain. *J. Mol. Biol.* **351:** 672–682.

Srisailam, S., Kumar, T.K., Rajalingam, D., Kathir, K.M., Sheu, H.S., Jan, F.J., Chao, P.C., and Yu, C. 2003. Amyloid-like fibril formation in an all β-barrel protein. Partially structured intermediate state(s) is a precursor for fibril formation. *J. Biol. Chem.* **278:** 17701–17709.

Uversky, V.N. and Fink, A.L. 2004. Conformational constraints for amyloid fibrillation: The importance of being unfolded. *Biochim. Biophys. Acta* **1698:** 131–153.

Wright, C.F., Teichmann, S.A., Clarke, J., and Dobson, C.M. 2005. The importance of sequence diversity in the aggregation and evolution of proteins. *Nature* **438:** 878–881.