the FEBS Journal

Molecular dynamics of the DNA-binding domain of the papillomavirus E2 transcriptional regulator uncover differential properties for DNA target accommodation

M. Falconi[1], A. Santolamazza[1], T. Eliseo[2], G. de Prat-Gay[3], D. O. Cicero[2] and A. Desideri[1]

1 Department of Biology and CIBB (Centro Interdipartimentale di Biostatistica e Bioinformatica), University of Rome 'Tor Vergata', Italy
2 Department of Science and Chemical Technologies, University of Rome 'Tor Vergata', Italy
3 Instituto de Investigaciones Bioquimicas Fundación Leloir, Facultad de Ciencias Exactas y Naturales and CONICET, Universidad de Buenos Aires, Argentina

Keywords
molecular dynamics simulation; papillomavirus; protein–DNA recognition; protein flexibility; transcription factor

Correspondence
A. Desideri, Department of Biology, University of Rome 'Tor Vergata', Via della Ricerca Scientifica, 00133 Rome, Italy
Fax: +39 06 202279
Tel: +39 06 72594376
E-mail: desideri@uniroma2.it

(Received 5 December 2006, revised 5 March 2007, accepted 7 March 2007)

doi:10.1111/j.1742-4658.2007.05773.x

Papillomaviruses are small DNA tumor viruses that infect mammalian hosts, with consequences from benign to cancerous lesions. The Early protein 2 is the master regulator for the virus life cycle, participating in gene transcription, DNA replication, and viral episome migration. All of these functions rely on primary target recognition by its dimeric DNA-binding domain. In this work, we performed molecular dynamics simulations in order to gain insights into the structural dynamics of the DNA-binding domains of two prototypic strains, human papillomavirus strain 16 and the bovine papillomavirus strain 1. The simulations underline different dynamic features in the two proteins. The human papillomavirus strain 16 domain displays a higher flexibility of the β2–β3 connecting loop in comparison with the bovine papillomavirus strain 1 domain, with a consequent effect on the DNA-binding helices, and thus on the modulation of DNA recognition. A compact β-barrel is found in human papillomavirus strain 16, whereas the bovine papillomavirus strain 1 protein is characterized by a loose β-barrel with a large number of cavities filled by water, which provides great flexibility. The rigidity of the human papillomavirus strain 16 β-barrel prevents protein deformation, and, as a consequence, deformable spacers are the preferred targets in complex formation. In contrast, in bovine papillomavirus strain 1, a more deformable β-barrel confers greater adaptability to the protein, allowing the binding of less flexible DNA regions. The flexibility data are confirmed by the experimental NMR S^2 values, which are reproduced well by calculation. This feature may provide the protein with an ability to discriminate between spacer sequences. Clearly, the deformability required for the formation of the Early protein 2 C-terminal DNA-binding domain–DNA complexes of various types is based not only on the rigidity of the base sequences in the DNA spacers, but also on the intrinsic deformability properties of each domain.

The mechanisms by which DNA sequences are recognized by proteins have been intensively investigated in the past few decades. Although these studies describe intricate hydrogen-bonding networks between amino acid side chains and DNA bases, a simple code for protein–DNA recognition based on noncovalent

Abbreviations
BPV-1, bovine papillomavirus strain 1; DBD, DNA-binding domain; E2, Early protein 2; HPV, human papillomavirus; MD, molecular dynamics; rmsf, root-mean-square fluctuations.

chemistry has failed to emerge. Instead, it appears that the specificity of protein–DNA reactions derives from a balance of several factors [1]. In addition to base–amino acid contacts, these include contacts among the protein and the phosphodiester backbone of the DNA, solvent-mediated interactions, and the structural adaptability of the reactants.

An added layer of complexity is evident in physiologic environments where proteins have to select among multiple, similar binding site sequences and interact with short DNA sequences in the presence of a vast excess of nonspecific DNA [1]. In such situations, it is likely that small differences in reaction affinity, or kinetics, can deeply influence regulatory events. The evolution of protein–DNA interactions could be viewed as an ongoing process of tailoring the balance between the stereochemical constraints outlined above to the required biological function [1].

The papillomaviruses represent a good model system for the investigation of such issues, because there are many viral strains that have coevolved with their vertebrate hosts for over a 100 million years, providing a database for the study of molecular, structural and functional coevolution. The papillomaviruses are a group of small DNA tumor viruses that induce warts in mammals. The Early 2 proteins (E2s) regulate expression of all viral genes [2,3] and viral replication through association with the Early 1 protein helicase [4–6].

E2 consists of three domains: the well-conserved N-terminal transactivation domain, a variable intermediate hinge region, and a C-terminal DNA-binding/dimerization domain [7]. Crystal and NMR structures of the bovine papillomavirus type 1 (BPV-1) E2 DNA-binding domain (DBD) (BPV-1 E2-DBD), alone [8,9] and in complex with an oligonucleotide [9,10], have been solved. Also in the case of human papillomavirus (HPV) type 16, the structure of the E2 DBD (HPV-16 E2-DBD) alone [11,12] and in complex with DNA is available [13]. These structures revealed that the protein forms a dimeric β-barrel with surface 'recognition' α-helices (Fig. 1). The dimeric β-barrel domain is an unusual topology, shared only by the Epstein–Barr EBNA1 DBD [14]. In this topology, secondary, tertiary and quaternary structure are coupled, and the dimerization interface is composed of two four-stranded half-β-barrels [15].

Although the tertiary structures of all characterized E2-DBDs are similar, there is an interesting variation in the relative orientation of the two subunits [1]. On this basis, the E2s can be divided into two distinct classes, one including HPV-16 and HPV-31, and the other BPV-1 and HPV-18 [1]. These differences in quaternary structure are likely to induce a different DNA deformation upon E2 binding.

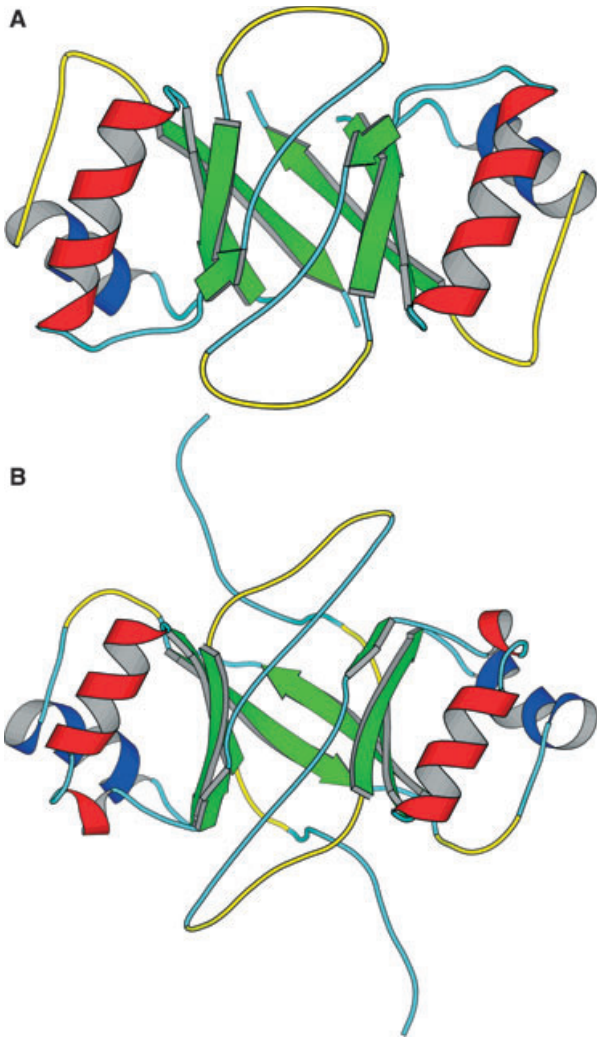

Fig. 1. DNA interaction side view of the HPV-16 DBD E2 structure (A) and BPV-1 DBD E2 structure (B). The α-helices involved in DNA recognition are shown as red spiral ribbons, and other α-helices are represented by blue spiral ribbons. β-Strands are indicated by green arrows. The yellow wire represents high-flexibility regions, and the cyan wire indicates the remaining random-coil structure and the turns. This picture was produced using the program MOLSCRIPT [45].

The transcriptional regulation, growth inhibition and replication functions of E2 are mediated through its interaction with a palindromic consensus sequence ACCgN4cGGT, where N4 indicates the 'spacer' nucleotides and small letters represent preferred but not totally conserved nucleotides. Multiple E2 binding sites that differ in the sequences of the central N4 'spacer' nucleotides are present in the viral genomes (17 in BPV-1 and 4 in HPV-16). Whereas BPV-1 E2 shows only two- to eight-fold differences in affinity towards

E2 binding sites with different spacers, HPV-16 E2 displays a nearly 300-fold enhanced affinity for the E2 binding sites containing AA(A/T)N spacers [16,17]. The structure of the spacer region, which is not contacted by the protein, is critical for the formation of the high-affinity sequence-specific protein–DNA complex, and the differential binding affinity has been proposed to be regulated by the intrinsic structure and deformability encoded in the base sequence of the DNA target [18].

The two proteins also display differential affinity towards binding sites possessing nicked or gapped spacers, indicating distinct differences in their sensitivity to DNA structure and/or flexibility [17]. Despite these differences, the residues involved in direct base interactions are identical [17], indicating that this is not the mechanism responsible for discriminating the DNA-binding site sequence. Previous molecular dynamics (MD) studies have mainly investigated the structural behavior of BPV-1 DNA target sequences [19,20], and have suggested that the structures of both free and bound DNA half-sites are very close to each other, but have not discussed the protein behavior in detail.

In this work, we have investigated, through MD simulation, the structural–dynamic properties of the DBD of human and bovine papillomavirus E2s. The results show that the domains from different species, although having the same secondary and tertiary structures, show a different distribution of molecular flexibility. The mechanical properties that characterize the two proteins, together with the different structural and conformational features of the spacer regions in the DNA target sequences, indicate diverse mechanisms for the recognition of the DNA.

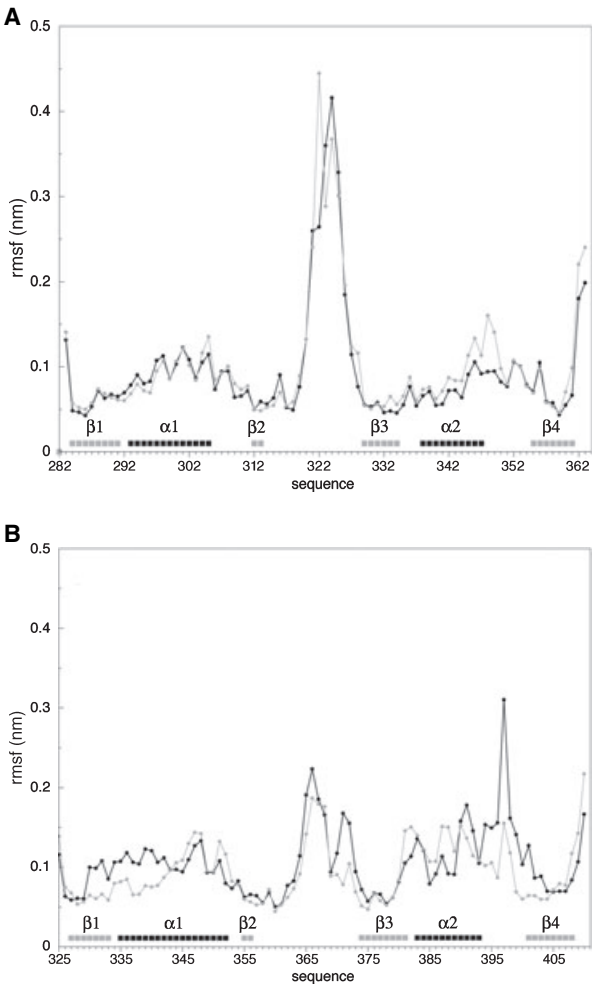

Fig. 2. rmsf averaged over each residue for each subunit of HPV-16 E2 (A) and BPV-1 E2 (B). The residues of the first subunit are indicated by black filled circles, and the residues of the second subunit are indicated by gray filled circles. The residues that in the NMR starting structure are in the α-helix and β-strand are indicated by black and gray squares, respectively.

Results

Analysis of root-mean-square fluctuations

The main chain root-mean-square fluctuations (rmsf), calculated over the trajectories and averaged over each residue, for the HPV-16 and BPV-1 E2s are shown in Fig. 2A,B. In both proteins, the α-helices show a relatively high rmsf value when compared with the β-segments. The largest fluctuations are observed in HPV-16 (Fig. 2A), and in particular in the large loop region connecting strand β2 and β3 (Gly321–Ser328), where the rmsf reaches a value higher than 0.45 nm. In the same region of BPV-1 (Fig. 2B), the protein fluctuation is smaller, the corresponding rmsf value being about 0.2 nm. In BPV-1, the largest fluctuation is observed at the level of helix α2 and of the loop connecting this helix with β-strand 4 (Pro383–Asn400).

In Fig. 2B, the first 14 amino acids have been removed because their rmsf values are out of scale. These residues belong to the last part of the linker region, between the N-terminal and the C-terminal domains, known to be extremely flexible and not structured. The addition of a variable number of amino acids to this region enhances the stability of BPV-1 to urea denaturation [8].

This indicates that the extension of the BPV1 E2 N-terminal DBD has a role in domain stability and DNA binding [21].

Experimental data show that the HPV-16 loop (Gly321–Ser328) is exposed to the solvent and remains flexible even after forming a complex with DNA [13]. Moreover, the presence of two lysines and two histidines in this region complement well the negatively charged phosphate backbone of the nucleic acid [13], and mutations of residues located in this region lead to changes in the DNA recognition kinetics and in the stability of the complex [22], suggesting that there is an involvement of this loop in HPV-16 E2–DNA recognition.

Another interesting aspect of HPV-16 is represented by the presence of Lys349 in the loop connecting helix α2 and strand β4. This residue may have a role in DNA binding, as a single mutation of Lys349 to alanine weakens the DNA binding of the HPV-16 E2 C-terminal domain by 1.0 kcal·mol^{-1} [13].

Secondary structure analysis

The secondary structure analysis was carried out on both proteins for all the simulation times. The few differences that emerged on comparing the results concern the α-helices, and in particular the DNA recognition helix α1 [1], represented in red in Fig. 1. Figure 3A,B shows the conformational evolution, as a function of time, of the residues that start the simulation in the α-helix. In the HPV-16 protein, some residues inside helix α1 lost their regular structure and adopted an alternative 'turn' or '3–10 helix' conformation, suggesting a 'conformational adaptability' to better fit the DNA major groove recognition site. These alterations are probably necessary to permit suitable plasticity of helix α1 when it interacts with the DNA major groove. Structural changes involving the central part of helix α1 were not observed in the simulation of BPV-1, where some residues in the C-terminal part of the α-helix switched their secondary structure to a 'turn' conformation (Fig. 3A,B).

S^2 analysis for the NH atoms

Nuclear magnetic relaxation spectroscopy is one of the few experimental sources providing spatially resolved information on subnanosecond dynamics of biomolecules in solution. Dipolar relaxation data of heteronuclear spin pairs, such as ^{13}C–H and ^{15}N–H, are often interpreted using the Lipari–Szabo model [23], in which the motion of the involved internuclear vector is characterized by an internal time scale τ_e, an overall time scale τ_c, and an order parameter S^2. The order parameter S^2 [23–25] was calculated from MD simulation for the NH atoms of the E2 chains of both HPV-16 and BPV-1 (Fig. 4A,B), and in the case of HPV-16 was compared with the corresponding S^2 experimental values measured by NMR spectroscopy. S^2 values close to 0 indicate high flexibility, and S^2 values close to 1 indicate low flexibility.

In HPV-16 (Fig. 4A), the trends of the experimental and calculated S^2 values were similar. In fact, both NMR and MD identified the large loop region connecting strands β2 and β3 (Gly321–Ser328) as the region characterized by the highest flexibility. Moreover, both NMR and MD also identified a relatively large degree of flexibility in the region including the loop between helix α2 and strand β4, and the initial part of strand β4 (Cys350–Val356).

In the BPV-1 protein (Fig. 4B), the lowest S^2 values, and therefore the greatest flexibility, were observed on the N-terminal tail close to strand β1 (Gly324–Phe328) (see also Fig. 1B), and on the loop connecting helix α2 and strand β4 (Pro396–Asn400). Also, in this case the protein showed great flexibility at the level of the loop connecting strands β2 and β3 (Asn366–Ala374), even though these values are lower than those observed in HPV-16.

Analysis of cavities

The presence of cavities that occur in the internal part and on the surface of the two proteins has been evaluated by applying the program SURFNET [26] (Fig. 5A,B). In this program, gap regions are defined by filling the region between the two molecules with gap-spheres and then computing a three-dimensional density map that defines the surface of the gap region [26].

In the interior of the HPV-16 β-barrel, at the interface of each dimer, only one small cavity, present for a short percentage of the simulation time, was found, whereas several crevices were present on the protein surface (Fig. 5A). The absence of internal cavities in the HPV-16 protein indicates that the barrel is hardly accessible to solvent, as confirmed by the high S^2 values, indicative of rigidity, found in the regions close to strands β1 and β4 (Fig. 4A). Several crevices are also present on the protein surface of BPV-1, but in this protein, several cavities inside the β-barrel, at the dimer interface, were also observed (Fig. 5B). This result agrees with the low S^2 values for residues surrounding the cavities. Both parameters are, in fact, indicative of relatively high flexibility.

This behavior underlines a greater degree of compactness of the β-barrel in the HPV-16 domain than in the BPV-1 domain.

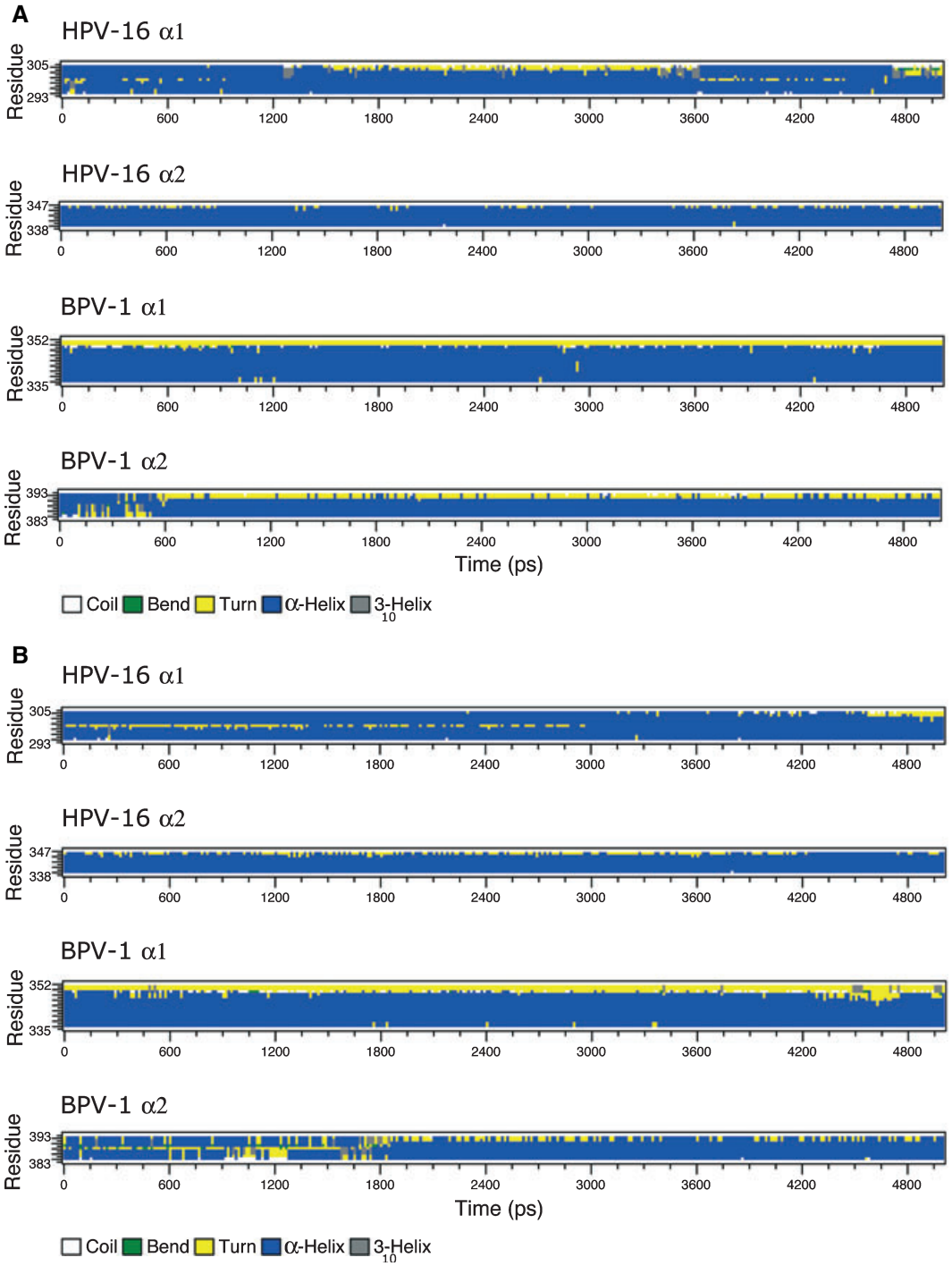

Fig. 3. Secondary structure evolution, as a function of time, for the protein segments that start the simulations as α-helices. (A) first subunit and (B) second subunit of HPV-16 E2 and BPV-1 E2. A color code identifying the secondary structure is shown.

Principal component analysis

Principal component analysis has been applied to both the HPV-16 and BPV-1 E2 trajectories to identify the main 3N directions along which the majority of the protein motion is defined [27,28]. The analysis is based on the diagonalization of the covariance matrix built from the atomic fluctuations after the removal of the

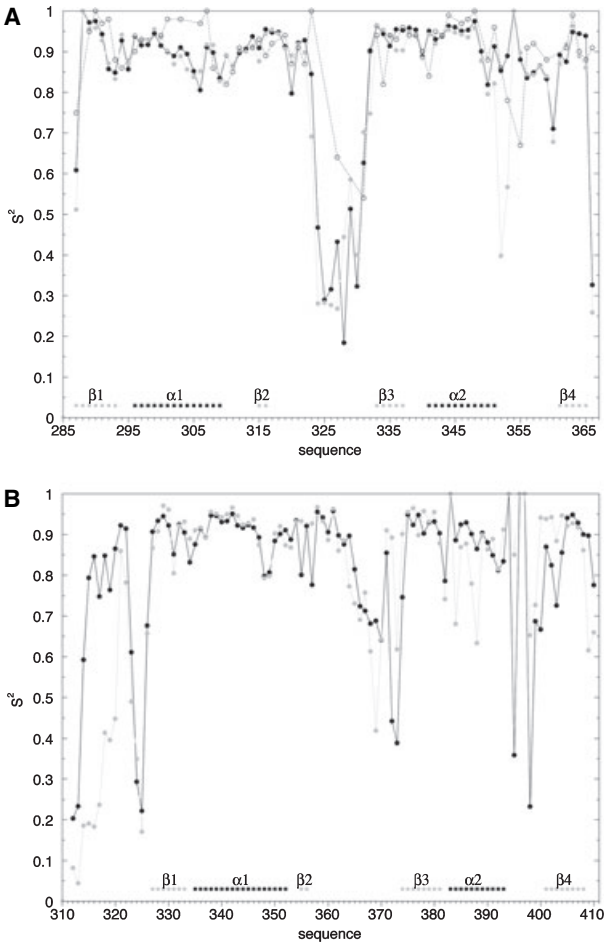

Fig. 4. S^2 order parameters evaluated in the simulation for the NH groups of HPV-16 E2 (A) and BPV-1 E2 (B). In (A), simulation values are compared with the corresponding parameters measured by NMR spectroscopy. An arbitrary value of one has been given to the proline residues that cannot be assigned because of the absence of the NH group. In (A), the black dotted line represents the S^2 NMR values; in (A) and (B), the black and the gray point-dashed lines show the values calculated in MD for chain A and chain B, respectively. The residues that in the NMR starting structure are in the α-helix and β-strand conformations are indicated by the black and gray squares, respectively.

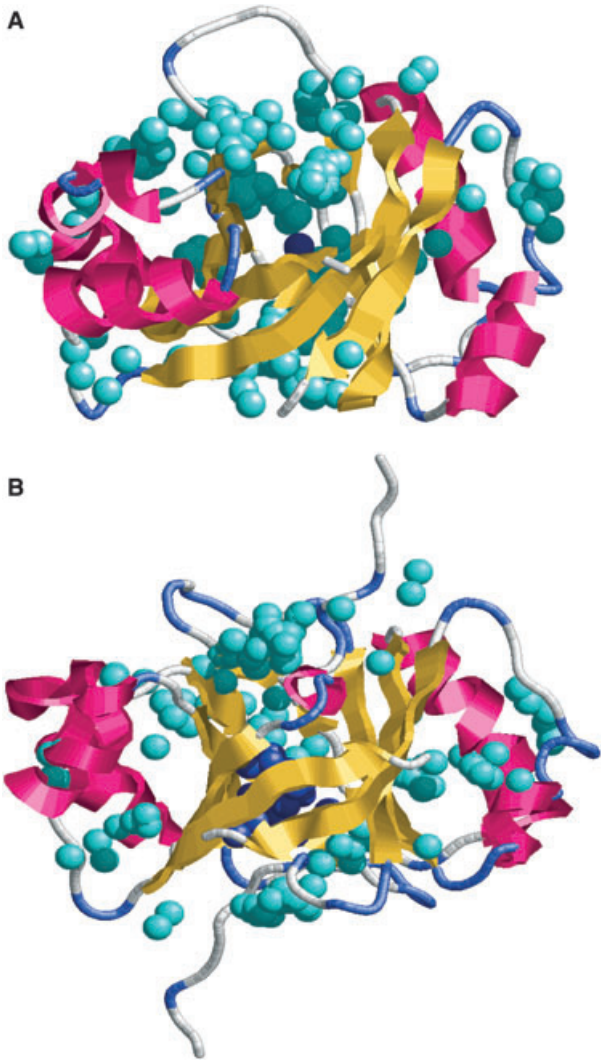

Fig. 5. Cavities detected during the MD simulation inside and outside HPV-16 E2 (A) and BPV-1 E2 (B). The α-helices are shown as red spiral ribbons, and the β-strands are indicated by yellow arrows. The white and blue wires represent the loops and the turns, respectively. Cyan spheres represent the geometric centers of the cavities that are located over the protein surface, and the blue spheres indicate the geometric centers of the cavities that are located inside the β-barrel. The picture was produced using the program RASMOL [46].

translational and rotational movement, and was carried out on the Cα atoms of the proteins.

Large displacements occurred for both proteins along the first eigenvector, characterized by the largest eigenvalue (data not shown). The motion along the first eigenvector for the HPV-16 and BPV-1 E2s can be well appreciated by looking at the Cα projections shown in Fig. 6A,B. Ten projections of the motion were extracted and plotted to illustrate the different dynamic behavior of the two proteins. The HPV-16 protein (Fig. 6A) showed a rigid β-barrel, a highly fluctuating β2–β3 loop, and a partial deformation in the center of the recognition helices. The BPV-1 protein (Fig. 6B) showed a fluctuating β-barrel, a relatively rigid β2–β3 loop, no deformation of the recognition α-helices, and substantial fluctuation of the long N-terminal tail.

Discussion

The results obtained in these simulations highlight a difference in the structural behavior of the two

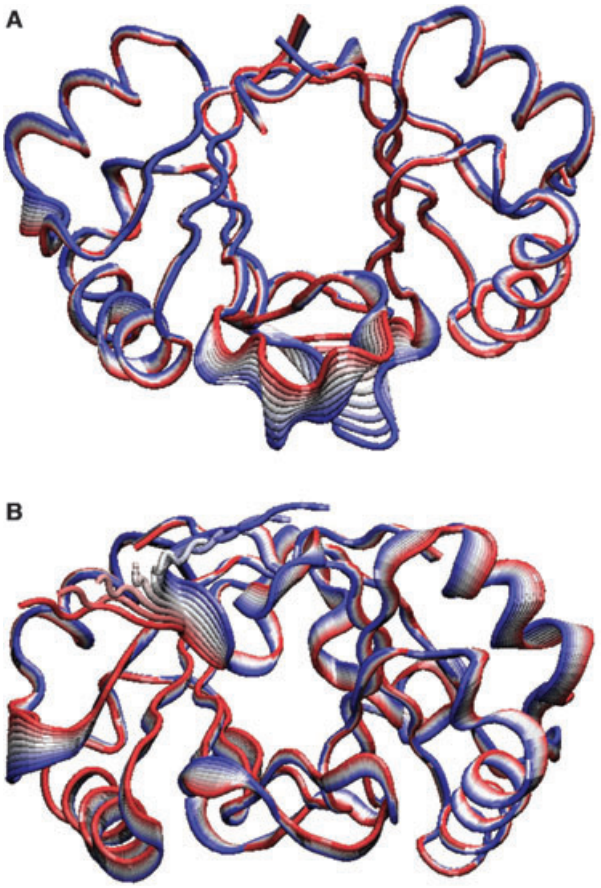

Fig. 6. Representation of 10 projections of the motion along the first eigenvector for HPV-16 E2 (A) and BPV-1 E2 (B). The picture was produced using the program VMD [47].

proteins that is probably correlated with different ways of recognizing the DNA. The HPV-16 protein shows high flexibility in the loop connecting strands β2 and β3, but displays at the same time a more rigid β-barrel (see Fig. 6A), as monitored by the fast fluctuations sampled by MD simulation. On the other hand, after several hours, the solvent accessibility of NH groups in the core of the protein becomes high [12,29], indicating that the protein is subjected to low-frequency motions.

In the HPV-16 protein, the high-frequency flexibility of the large loop, connecting strands β2 and β3 (see Fig. 6A), may balance the rigidity of the barrel, facilitating DNA binding. In this protein, a certain degree of plasticity is also shown by the DNA-binding helices α1, which, in their central part, partially lose secondary structure, as indicated by the DSSP analysis in the simulation (Fig. 3A,B). This feature provides an adaptability to the DNA interaction sites that compensates for the lower mobility of the barrel. The relatively high plasticity of helix α1 is in line with the fast solvent exchange observed for most of the amide groups of the recognition helix α1 in the homologous HPV-31 E2 [12,29]. Moreover, chemical shift values of the C-terminal part of the helix display deviations from those expected for a regular helix, in particular at the level of the Phe303 residue [13]. Interestingly, these deviations are maintained also in the DNA-bound form [13].

In the HPV-16 protein, the β2–β3 loop is characterized by a large number of positive charges that contribute to the DNA binding [13]. In fact, mutations to alanine of the residues Lys325 and Lys327, located in the β2–β3 loop, produce a decrease in the DNA binding [22] that is restored upon back mutation into arginine, indicating the importance of the positive charges for the occurrence of nonspecific contacts with DNA [22].

The BPV-1 protein shows low flexibility of the loop connecting the β2 and β3 strands (Fig. 6B). The low flexibility is counterbalanced by a larger barrel flexibility (Fig. 6B), as indicated by the large number of cavities present in its interior (Fig. 5B), which allow a broad range of movements for the structured helices α1 that are necessary for nonspecific DNA target recognition.

From these data, it is possible to propose a significant role in DNA binding for the loop connecting strands β2 and β3 in HPV-16, and for the regions close to strands β1 and β4 in BPV-1. For HPV-16, the flexibility of the loop may alter the adaptability of the DNA-binding helices, thus modulating the discrimination of specific versus nonspecific DNA sequences. For both the E2s, we suggest that the 'indirect readout' [30] plays a significant role in DNA sequence recognition, although a systematic rational perturbation of the DNA-binding interface showed that, in the case of HPV-16, most of the binding energy comes from a 'direct readout' recognition mode [22].

In fact, recognition of DNA by proteins, in addition to direct interactions, relies also on indirect effects that reflect several energetic contributions to the response of DNA sequences to twisting and bending distortions induced by proteins. In the light of what has been observed in this work, we can suggest that the indirect effects are not only attributable to the DNA molecular structure, but are finely tuned by the mechanical and dynamic properties of the specific protein structure involved in the interaction. These properties may modulate the indirect effects that, by stabilizing the complex, can lead to the selection of an alternative binding site.

Experimental procedures

MD simulations and analysis

The HPV-16 [12] and BPV-1 [8] E2 coordinates were obtained by NMR and stored in the Protein Data Bank (http://www.rcsb.org/pdb; PDB codes 1R8P and 1DBD, respectively). Two simulations of 5.08 ns were carried out on the HPV-16 and BPV-1 proteins. The system topologies were obtained with the AMBER LEAP module [31], and modeled with the all-atoms AMBER95 force field [32,33]. The proteins were immersed in rectangular boxes filled with TIP3P water molecules [34] (Table 1), imposing a minimal distance between the solute and the box walls of 10.0 Å. The two systems were neutralized with the AMBER LEAP module, adding the necessary amount of chloride ions (Table 2) in electrostatically favorable positions. Optimization and relaxation of solvent and ions were initially performed by means of three energy minimizations and two MD simulations (Table 2), keeping the solute atoms constrained to their initial positions with decreasing force constants of 500 and 25 kcal·(mol Å)$^{-1}$. Thereafter, the system was minimized without any constraint, and simulated for 60.0 ps at constant temperature of 300 K using Berendsen's method [35] and at a constant pressure of 1 bar with a 2.0 fs time step. Each system was thermalized for about 200 ps before the trajectory acquisition (see time column in Table 2).

Table 1. Size of the simulated systems.

Protein type	BPV-1	HPV-16
Total atoms	31 750	23 468
Protein atoms	3134	2664
Amino acids	200	162
Water molecules	9534	6930
Chloride ions	14	14
Simulation box *X* side (Å)	77	56
Simulation box *Y* side (Å)	56	73
Simulation box *Z* side (Å)	83	64
Saved configurations	10 160	10 160

Pressure and temperature coupling constants were 0.4 ps. The atomic positions were saved every 250 steps (0.5 ps) for the analysis. The two systems were simulated under periodic boundary conditions, using a cut-off radius of 9.0 Å for the nonbonded interactions, and updating the neighbor pair list every 10 steps. The electrostatic interactions were calculated with the particle mesh Ewald method [36,37]. The SHAKE algorithm [38] was used to constrain all bond lengths involving hydrogen atoms. The systems were simulated at CASPUR research center of Rome, Italy (Inter Universities Consortium for Supercomputing Applications) on Power 4 IBM parallel computers by using an 8 CPU cluster.

The systems were simulated for 5.0 ns, a time sufficient to evaluate protein loop, α-helix, and β-barrel fluctuations, and to identify differences in the dynamics of these two proteins. The rmsd from the starting structures of the two proteins (supplementary Fig. S1) showed, in fact, good stability over all the simulation times. Also, the S^2 values (supplementary Figs S2 and S3) and rmsf (supplementary Figs S4 and S5) calculated by splitting the trajectories into three segments give relatively similar results.

The analyses of trajectories for both systems were carried out over 5 ns using the GROMACS MD package version 3.2.1 program [39] and codes written in-house. The atomic rmsf values were computed using the following definition implemented in the GROMACS utility g_rmsf [39]:

$$RMSF_i = \sqrt{\sum_{\alpha=1}^{3} \langle (r_{i,\alpha}^{\min}(t) - \bar{r}_{i,\alpha})^2 \rangle_{MD}} \quad (1)$$

where the averages were computed over the equilibrated MD trajectory.

The calculation of order parameter S^2 for the backbone N–H bonds followed approaches published previously by other groups [23–25]. In short, the order parameters S^2 of a bond vector $\vec{\mu}(t)$ is computed as:

$$S^2 = 3/2[<x^2>^2 + <y^2>^2 + <z^2>^2 + 2<xy>^2 + 2<xz>^2 + 2<yz>^2] - 1/2 \quad (2)$$

in which x, y and z are the components of the unit bond vector $\vec{\mu}(t)$ along three Cartesian axes. Here, the braces

Table 2. System thermalization phases. EM, energy minimization.

Execution time (ps)	Thermalization phases	Number of steps	Position restraint value [kcal·(mol Å)]
0	EM1 with position restraints	10 000	500
0	EM2 with position restraints	20 000	500
12.5	MD1 with position restraints	25 000 of 0.5 fs	500
0	EM3 with position restraints	15 000	25
25.0	MD2 with position restraints	25 000 of 1.0 fs	25
0	EM4	10 000	–
20.0	MD3	10 000 of 2.0 fs	–
40.0	MD4	20 000 of 2.0 fs	–
100.0	MD5	50 000 of 2.0 fs	–

stand for the ensemble average. Prior to the calculation of S^2, overall translation and rotation of the protein molecule were removed.

The volume of the cavities and external crevices was measured using the program SURFNET [26]. The time evolution of the secondary structures was calculated using the GROMACS utility do_dssp [39], which iteratively uses the program DSSP [35] to evaluate the secondary structures.

Principal component analysis [27,28] was done using the GROMACS MD package version 3.1.4 [39].

Relaxation data and backbone dynamics analysis

^{15}N-Labeled E2 was prepared as previously described [12]. ^{15}N relaxation measurements were performed on a sample containing 50 mM sodium phosphate, 5 mM dithiothreitol (pH 6.5), and a protein concentration of 0.9 mM. Measurements of ^{15}N T_1, T_2 and ^{1}H–^{15}N NOE were performed at a ^{15}N frequency of 70.94 MHz, using standard pulse schemes [41,42]. ^{15}N relaxation data were analyzed in terms of model-free formalism, making use of the program DASHA [43].

Relaxation experiments were carried out at 30 °C on a Bruker Avance700 spectrometer (Rheinstetten, Germany) at a ^{15}N resonance frequency of 70.9 MHz. Measurements of ^{15}N T_1, T_2 and ^{1}H–^{15}N NOE were made by performing established ^{1}H-detected pulse schemes [41,44] in an interleaved manner to collect six points with delays of 14, 210, 420, 700, 1191, 1542 ms for T_1, and six points with delays of 8.2, 24.5, 40.8, 57.1, 73.4, 97.9 ms for T_2. Integrated crosspeak volumes of nonoverlapped resonances were fitted to two-parameter monoexponential decays. The uncertainties of peak intensities were evaluated as the SD of the spectral noise measured in a region free of crosspeaks. The heteronuclear NOE values were determined from the ratio of peak volumes of spectra recorded with and without ^{1}H saturation, employing a net relaxation delay of 5 s for each scan in both experiments.

References

1 Hegde RS (2002) The papillomavirus E2 proteins: structure, function, and biology. *Annu Rev Biophys Biomol Struct* **31**, 343–360.

2 Androphy EJ, Lowy DR & Schiller JT (1987) Bovine papillomavirus E2 trans-acting gene product binds to specific sites in papillomavirus DNA. *Nature* **325**, 70–73.

3 Romanczuk H, Thierry F & Howley PM (1990) Mutational analysis of cis-elements involved in E2 modulation of human papillomavirus type 16 P97 and type 18 P105 promoters. *J Virol* **64**, 2849–2859.

4 Chiang CM, Ustav M, Stenlund A, Ho TF, Broker TR & Chow LT (1992) Viral E1 and E2 proteins support replication of homologous and heterologous papillomavirus origins. *Proc Natl Acad Sci USA* **89**, 5799–5803.

5 Del Vecchio AM, Romaczuk H, Howley PM & Baker CC (1992) Transient replication of human papillomavirus DNAs. *J Virol* **66**, 5949–5958.

6 Ustav M & Stenlund A (1991) Transient replication of BPV-1 requires two viral polypeptides encoded by the E1 and E2 open reading frames. *EMBO J* **10**, 449–457.

7 Giri I & Yaniv M (1988) Structural and mutational analysis of E2 trans-activating proteins of papillomaviruses reveals three distinct functional domains. *EMBO J* **7**, 2823–2829.

8 Veeraraghavan S, Mello C, Androphy E & Baleja JD (1999) Structural correlates for enhanced stability in the E2 DNA-binding domain from bovine papillomavirus. *Biochemistry* **38**, 16115–16124.

9 Hegde RS, Wang AF, Kim SS & Schapira M (1998) Subunit rearrangement accompanies sequence-specific DNA-binding by the bovine papillomavirus-1 E2 protein. *J Mol Biol* **276**, 797–808.

10 Hegde RS, Grossman SR, Laimins LA & Sigler PB (1992) Crystal structure at 1.7 Å of the bovine papillomavirus-1 E2 DNA-binding domain bound to its DNA target. *Nature* **359**, 505–512.

11 Hegde RS & Androphy EJ (1998) Crystal structure of the E2 DNA-binding domain from human papillomavirus type 16: implications for its DNA binding-site selection mechanism. *J Mol Biol* **284**, 1479–1489.

12 Nadra AD, Eliseo T, Mok YK, Almeida CL, Bycroft M, Paci M, de Prat-Gay G & Cicero DO (2004) Solution structure of the HPV-16 E2 DNA binding domain, a transcriptional regulator with a dimeric beta-barrel fold. *J Biomol NMR* **30**, 211–214.

13 Cicero DO, Nadra AD, Eliseo T, Dellarole M, Paci M & de Prat-Gay G (2006) Structural and thermodynamic basis for the enhanced transcriptional control by the human papillomavirus strain-16 E2 protein. *Biochemistry* **45**, 6551–6560.

14 Bochkarev A, Barwell JA, Pfuetzner RA, Furey W Jr, Edwards AM & Frappier L (1995) Crystal structure of the DNA-binding domain of the Epstein–Barr virus origin-binding protein EBNA 1. *Cell* **83**, 39–46.

15 de Prat-Gay G, Nadra AD, Corrales-Izquierdo FJ, Alonso LG, Ferreiro DU & Mok YK (2005) The folding mechanism of a dimeric beta-barrel domain. *J Mol Biol* **351**, 672–682.

16 Bedrosian CL & Bastia D (1990) The DNA-binding domain of HPV-16 E2 protein interaction with the viral enhancer: protein-induced DNA bending and role of the non-conserved core sequence in binding site affinity. *Virology* **174**, 557–575.

17 Hines CS, Meghoo C, Shetty S, Biburger M, Brenowitz M & Hegde RS (1998) DNA structure and flexibility in the sequence-specific binding of papillomavirus E2 proteins. *J Mol Biol* **276**, 809–818.

18 Hizver J, Rozenberg H, Frolow F, Rabinovich D & Shakked Z (2001) DNA bending by an adenine-thymine

tract and its role in gene regulation. *Proc Natl Acad Sci USA* **98**, 8490–8495.
19 Djuranovic D, Oguey C & Hartmann B (2004) The role of DNA structure and dynamics in the recognition of bovine papillomavirus E2 protein target sequences. *J Mol Biol* **339**, 785–796.
20 Djuranovic D & Hartmann B (2005) Molecular dynamics studies on free and bound targets of the bovine papillomavirus type I e2 protein: the protein binding effect on DNA and the recognition mechanism. *Biophys J* **89**, 2542–2551.
21 Pepinsky RB, Prakash SS, Corina K, Grossel MJ, Barsoum J & Androphy EJ (1997) Sequences flanking the core DNA-binding domain of bovine papillomavirus type 1 E2 contribute to DNA-binding function. *J Virol* **71**, 828–831.
22 Ferreiro D, Dellarole M, Nadra AD & De Prat-Gay G (2005) Free energy contributions to direct readout of a DNA sequence. *J Biol Chem* **280**, 32480–32484.
23 Lipari G & Szabo A (1982) Model-free approach to the interpretation of nuclear magnetic resonance relaxation in macromolecules. 1. Theory and range of validity. *J Am Chem Soc* **104**, 4546–4559.
24 Henry ER & Szabo A (1985) Influence of vibrational motion on solid state line shapes and NMR relaxation. *J Chem Phys* **82**, 4753–4761.
25 Hu H, Clarkson MW, Hermans J & Lee AL (2003) Increased rigidity of eglin c at acidic pH: evidence from NMR spin relaxation and MD simulations. *Biochemistry* **42**, 13856–13868.
26 Laskowski RA (1995) SURFNET: a program for visualizing molecular surfaces, cavities, and intermolecular interactions. *J Mol Graph* **13**, 323–330.
27 Garcia AE (1992) Large-amplitude nonlinear motions in proteins. *Phys Rev Lett* **68**, 2696–2699.
28 Amadei A, Linssen AB & Berendsen HJ (1993) Essential dynamics of proteins. *Proteins* **17**, 412–425.
29 Liang H, Petros AM, Meadows RP, Yoon HS, Egan DA, Walter K, Holzman TF, Robins T & Fesik SW (1996) *Biochemistry* **35**, 2095–2103.
30 Gromiha MM, Siebers JG, Selvaraj S, Kono H & Sarai A (2004) Intermolecular and intramolecular readout mechanisms in protein–DNA recognition. *J Mol Biol* **337**, 285–294.
31 Case DA, Cheatham TE III, Darden T, Gohlke H, Luo R, Merz KM, Onufriev A Jr, Simmerling C, Wang B & Woods R (2005) The Amber biomolecular simulation programs. *J Comput Chem* **26**, 1668–1688.
32 Cornell WD, Cieplak P, Bayly CI, Gould IR, Kenneth M, Merz J, Ferguson DM, Spellmeyer DC, Fox T, Caldwell JW *et al*. (1995) A second generation force field for the simulations of proteins, nucleic acids and organic molecules. *J Am Chem Soc* **117**, 5179–5197.
33 Ponder JW & Case DA (2003) Force fields for protein simulations. *Adv Prot Chem* **66**, 27–85.
34 Jorgensen WL, Chandrasekhar J, Madura JD, Impey RW & Klein ML (1983) Comparison of simple potential functions for simulating liquid water. *J Chem Phys* **79**, 926–935.
35 Berendsen HJC, Postma JPM, van Gusteren WF, Di Nola A & Haak JR (1984) Molecular dynamics with coupling to an external bath. *J Comput Phys* **81**, 3684–3690.
36 Darden T, York D & Pedersen L (1993) Particle mesh Ewald and N.log (n) method for Ewald sums in large systems. *J Chem Phys* **98**, 10089–10092.
37 Cheatham TE, Miller JL, Fox T, Darden TA & Kolman PA (1995) Molecular dynamics simulation on solvated biomolecular systems: the particle mesh Ewald method leads to stable trajectories of DNA, RNA and proteins. *J Am Chem Soc* **117**, 4193–4194.
38 Ryckaert JP, Ciccotti G & Berendsen HJC (1977) Numerical integration of the Cartesian equations of motion of a system with constraints: molecular dynamics of n-alkanes. *J Comput Phys* **23**, 327–341.
39 Berendsen HJC, van der Spool D & van Drunen R (1995) GROMACS: a message-passing parallel molecular dynamics implementation. *Comp Phys Commun* **95**, 43–56.
40 Kabsch W & Sander C (1983) Dictionary of protein secondary structure: pattern recognition of hydrogen-bonded and geometrical features. *Biopolymers* **22**, 2577–2637.
41 Kay LE, Torchia DA & Bax A (1989) Backbone dynamics of proteins as studied by 15N inverse detected heteronuclear NMR spectroscopy: application to staphylococcal nuclease. *Biochemistry* **28**, 8972–8979.
42 Barbato G, Ikura M, Kay LE, Pastor RW & Bax A (1992) Backbone dynamics of calmodulin studied by 15N relaxation using inverse detected two-dimensional NMR spectroscopy: the central helix is flexible. *Biochemistry* **31**, 5269–5278.
43 Orekhov VY, Nolde DE, Golovanov AP, Korzhnev PM & Arseniev AS (1995) Processing of heteronuclear NMR relaxation data with the new software DASHA. *Appl Magn Reson* **9**, 581–588.
44 Stone MJ, Fairbrother WJ, Palmer AG, Reizer J, Saier MH & Wright PE (1992) Backbone dynamics of the *Bacillus subtilis* glucose permease IIA domain determined from ^{15}N NMR relaxation measurements. *Biochemistry* **31**, 4394–4406.
45 Kraulis PJ (1991) MOLSCRIPT: a program to produce both detailed and schematic plots of protein structures. *J Appl Crystallogr* **24**, 946–950.
46 Sayle RA & Milner-White EJ (1995) RasMol: biomolecular graphics for all. *Trends Biochem Sci* **20**, 374–376.
47 Humphrey W, Dalke A & Schulten K (1996) VMD – Visual Molecular Dynamics. *J Mol Graph* **14**, 33–38.

Supplementary material

The following supplementary material is available online:

Fig. S1. Main chain rmsd as a function of time of HPV-16 E2 (upper panel) and BPV-1 E2 (bottom panel).

Fig. S2. S^2 order parameters evaluated in three windows of the simulation time for the NH groups of HPV-16 E2.

Fig. S3. S^2 order parameters evaluated in three windows of the simulation time for the NH groups of BPV-1 E2.

Fig. S4. rmsf evaluated in three sections of the simulation and averaged over each residue for each subunit of HPV-16 E2.

Fig. S5. rmsf evaluated in three sections of the simulation and averaged over each residue for each subunit of BPV-1 E2.

This material is available as part of the online article from http://www.blackwell-synergy.com

Please note: Blackwell Publishing is not responsible for the content or functionality of any supplementary materials supplied by the authors. Any queries (other than missing material) should be directed to the corresponding author for the article.