



Formation of well-defined soluble aggregates upon fusion to MBP is a generic property of E6 proteins from various human papillomavirus species

Katia Zanier ^a, Yves Nominé ^a, Sebastian Charbonnier ^a, Christine Ruhlmann ^b,
Patrick Schultz ^b, Johannes Schweizer ^c, Gilles Travé ^{a,*}

^a Ecole Supérieure de Biotechnologie de Strasbourg, Institut Gilbert Laustriat (UMR 7175 LC1),
Boulevard Sébastien Brant, BP 10413, 67412 Illkirch, France

^b Institut de Génétique et de Biologie Moléculaire et Cellulaire, CNRS/INSERM/IULP, 1 rue Laurent Fries, BP 163, 67404 Illkirch, France

^c Arbor Vita Corporation, 772 Lucerne Drive, Sunnyvale, CA 94085, USA

Received 11 May 2006, and in revised form 27 June 2006

Abstract

Protein aggregation is a main barrier hindering structural and functional studies of a number of interesting biological targets. The E6 oncoprotein of Human Papillomavirus strain 16 (E6₁₆) is difficult to express under a native soluble form in bacteria. Produced as an unfused sequence, it forms inclusion bodies. Fused to the C-terminus of MBP, it is mainly produced in the form of soluble high molecular weight aggregates. Here, we produced as MBP-fusions seven E6 proteins from other HPV strains (5, 11, 18, 33, 45, 52, and 58) belonging to four different species, and we compared their aggregation state to that of MBP-E6₁₆. Using a fast mutagenesis method, we changed most non-conserved cysteines to the isosteric residue serine to minimize disulfide bridge-mediated aggregation during purification. Static and dynamic light scattering measurements, ultracentrifugation and electron microscopy demonstrated the presence in all MBP-E6 preparations of soluble high-molecular weight aggregates with a well-defined spherical shape. These aggregated particles are relatively monodisperse but their amount and their size vary depending on the conditions of expression and the strain considered. For all strains, minimal aggregate formation occurs when the expression is performed at 15 °C. Such observations suggest that the assembly of MBP-E6 aggregates takes place *in vivo* during protein biosynthesis, rather than occurring during purification. Finally, we show that all MBP-E6 preparations contain two zinc ions per protein monomer, suggesting that E6 domains within the high molecular weight aggregates possess a native-like fold, which enables correct coordination to the metal center.

© 2006 Published by Elsevier Inc.

Keywords: HPV; Oncoprotein; E6; Fusion protein; Protein aggregation; Protein overexpression; Rapid mutagenesis; Light scattering; Soluble inclusion bodies

Human papillomaviruses (HPV)¹ are small DNA viruses that infect epithelial tissues. 26 HPV species have been identified, comprising nearly 100 strains [1]. HPVs display distinct tropisms for different body sites (skin,

mouth and genitalia) as well as distinct levels of pathogenicity. Infections by “low-risk” HPVs generate benign proliferative lesions whereas infections by “high-risk” HPVs can progress into cancer. Cervical carcinomas are one of the main causes of cancer-related death for women and are caused by infection of high-risk genital HPVs [2]. HPV 16 is the most common high-risk genital type responsible for 50% of the cervical cancers, followed by HPV 18 covering 14% of the cancers [3]. Other oncogenic genital types include HPV 45, 31, 33, 52, and 58.

* Corresponding author. Fax: +33 3 90244718.

E-mail address: trave@esbs.u-strasbg.fr (G. Travé).

¹ Abbreviations used: HPV, human papillomavirus; MBP, maltose binding protein; SLS, static light scattering; DLS, dynamic light scattering; TEV, tobacco etch virus.

The high-risk cutaneous HPV types 5 and 8 have been instead associated to the occurrence of squamous cutaneous carcinoma (SCC) on sun-exposed areas of the skin of patients suffering from the rare genetic disease *epidermodysplasia verruciformis* (EV) [4].

HPV tumorigenesis is supported by the activities of two oncoproteins, E6 and E7. E6 proteins from high-risk genital HPVs form a trimeric complex with the tumor suppressor p53 [5] and the ubiquitin ligase E6-associated protein (E6AP) [6], which leads to ubiquitin-mediated degradation of p53. High-risk genital HPV E6 also target many other cellular proteins to degradation (reviewed in [7]), including the MAGI family of proteins [8,9], which recognise E6 *via* their multiple PDZ domains. In contrast, little is known about the targets of the high-risk cutaneous HPV E6 proteins, which might carry out their functions *via* alternative mechanisms. For instance, E6 from HPV 5 and 8 do not target p53 to degradation, but they promote the degradation of the Bak pro-apoptotic protein which is normally induced after UV irradiation of the skin [10].

HPV E6 are cysteine-rich small proteins of about 150 amino acids, which consist of two zinc-binding domains, E6-N and E6-C [11,12]. We recently solved the solution structure of the E6-C domain and proposed a model for the structure of the full-length protein [13]. Yet the high-resolution structure of the entire E6 protein remains to be solved. This is still hindered by difficulties in the recombinant production of full-length E6 under a native and soluble form. Attempts to produce the native full-length protein in *Escherichia coli* have mainly focussed on E6 from HPV 16 (E6₁₆) [11,14–16]. Most of these studies report on the aggregative tendencies of E6. We found that MBP-E6₁₆ fusions self-assemble during bacterial expression into high molecular weight particles (“soluble inclusion bodies”) [15,17]. The resulting particles are prevented from precipitation by the highly soluble MBP. When the E6 protein is separated from the MBP moiety *via* linker proteolysis, it immediately precipitates. We also found that the aggregates, most probably formed in the cells upon overexpression, further increased during purification due to disulfide bridging promoted by the numerous cysteine residues found in E6₁₆ sequence. This second cause of aggregation could be eliminated by mutating six non-conserved cysteines of E6₁₆ to the isosteric residue serine [15].

In the present work, we express as MBP fusions several E6 proteins of other HPV types, including major high-risk genital HPV 18, 33, 45, 52, 58, low-risk genital HPV 11 and high-risk cutaneous HPV 5. A fast mutagenesis approach is used to mutate non-conserved cysteines to serines. The influence of E6 sequence and expression temperature on the formation of soluble MBP-fused aggregates are investigated. Finally, p53 degradation and zinc binding by the MBP-E6 proteins are also characterised.

Materials and methods

A strategy for rapid multiple mutagenesis of E6 proteins

To rapidly mutate several E6 proteins at multiple positions, we applied an efficient PCR-based strategy, which was first proposed by Tomic et al. [18]. This strategy consists of performing PCR amplifications with primers containing the desired mutations, together with a site for a class II restriction enzyme placed at their 5'-ends (Fig. 1). Class II restriction enzymes recognize non-palindromic sequences and cleave the DNA on a single site situated at a short distance from their recognition sequence in the 3'-direction. We used here the enzyme *BsaI* which cleaves at N + 1 at the 3'-end and N + 5 at the 5'-end following the last base of its recognition site. We also routinely use two other enzymes, *BsmBI* and *BbsI*, which respectively cleave at N + 1 (3'-end) and N + 5 (5'-end), and at N + 2 (3'-end) and N + 6 (5'-end), following the last base pair of their recognition sites. A PCR fragment generated using primers containing at their 5' end a restriction site for any of these enzymes will present after cutting two sticky ends with four-base 5' overhangs. Since these enzymes do not cut on their own restriction sites, these sticky ends can adopt any sequence, which can be chosen upon primer design. The restriction sites, situated at the extreme ends of the PCR fragments, are eliminated during purification of the cut PCR fragments. It is therefore possible to engineer, on two distinct PCR fragments, sticky ends which are self-compatible and will either perfectly retain the original sequence or introduce any planned mutation. Alternatively, the primers can also be designed so that the generated sticky ends will be compatible with the sites of classical restriction enzymes such as *EcoRI*, *HindIII*, *BamHI*, *NeoI* or *Acc65I*.

We generally apply the following rules:

- (1) The engineered four-base cohesive site should ideally contain three (G,C) to facilitate ligation.
- (2) The engineered cohesive site should not be self-complementary.
- (3) The sequence of these four bases should be distinct for each end of all fragments which are going to be cloned together in the vector to avoid polymerisation or improper arrangement of the fragments. By taking this precaution we have successfully ligated together up to six mutated fragments of the cDNA of HPV16 E6 into the open expression vector.
- (4) After design of the oligomers one should always check that a new site for the classII enzyme has not been introduced.
- (5) The PCR amplification reactions should use a standard (therefore cheap) *Taq* polymerase with a large amount of template (200 ng of plasmid per 50 µl reaction) and a small number of amplification cycles (12

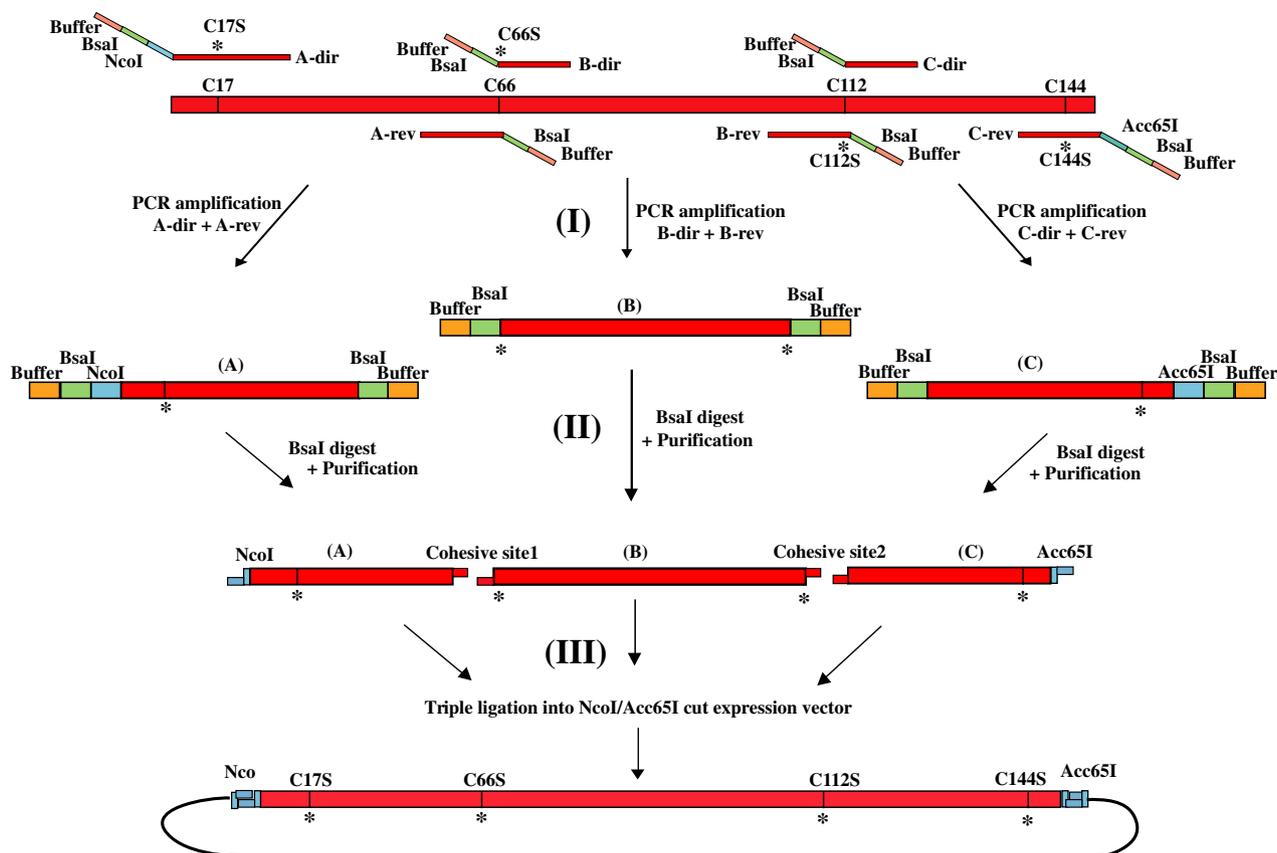


Fig. 1. PCR-based multiple-point mutagenesis strategy using class II restriction enzymes. Here, we show as an example the procedure we have followed for mutating into serines C17, C66, C112, and C144 of HPV 11 E6. Step (I): three PCR amplifications are performed using appropriate primers (see Fig. 2 for sequence details) bearing the planned mutations (indicated by asterisks) and *BsaI* restriction sites at their 5' extremities. The “dir” primers are designed to match the coding strand whereas the “rev” primers are designed to match the reverse (complementary) strand. Step (II): after digestion by *BsaI*, the fragments present the following characteristics: a *NcoI*-cohesive site and the (engineered) cohesive site 1 (fragment A); the cohesive site 1 and the cohesive site 2 (fragment B); the cohesive site 2 and a *Acc65I*-cohesive site (fragment C). Step (III): the three fragments are ligated together into the open *NcoI*-*Acc65I* vector, reconstituting HPV 11 E6 cDNA bearing the four planned mutations.

cycles). Such conditions generate reproducibly large amounts of product with very little chance of PCR-induced mutations.

- (6) After ligation, the selection of clones should be made based on their capability to overexpress a protein at the expected size. It is normally sufficient to check three–four clones. However, more clones may have to be screened if the multiple ligation step involves many fragments. Once this initial screening is performed, one or two clones expressing the protein at the right size can be sent for DNA sequencing. So far we did not have the case that a protein expressed at the right size would not present the exact expected sequence (we have already used this strategy for generating more than 100 point mutations in different proteins).

To illustrate the details of primer design, we describe in Fig. 2 the sequences of the primers which have been used for mutating four non-conserved cysteines (C17, C66, C112, and C144) of HPV 11 E6.

PCR, fragment processing and ligation into the expression vector

cDNA templates coding for HPV sequences were kindly provided by Prof. G. Orth (Institute Pasteur, Paris). The PCRs were performed in 50 μ l reaction mixes using RedTaq DNA Polymerase (Sigma–Aldrich) (standard enzyme and buffer conditions) with concentrations of 50 μ M of each nucleotide, 2 μ M of each primer and 200 ng of plasmid bearing the template cDNA. The PCR program comprised a starting 2 min. denaturation cycle at 95 $^{\circ}$ C followed by 12 cycles with the following steps: denaturation 94 $^{\circ}$ C for 20 s, annealing 60 $^{\circ}$ C for 30 s, elongation 72 $^{\circ}$ C for 50 s, followed by a single final 8-min step at 72 $^{\circ}$ C to complete elongation. The PCR products were separated on 2% agarose gel, purified and subsequently digested by *BsaI* (New England Biolabs, Ipswich, MA). Digests were separated on 2% agarose gel and purified. Series of digests allowing to reconstitute the entire cDNA of each mutant E6 were ligated together into the expression vector PET-M 41

summary, the Ragg parameters presented in this work are equal to the ratios of the intensities of scattered light at 350 nm and tryptophan fluorescence in denaturing conditions (i.e. I_{350}/I_{340}^*).

Preliminary experiments have shown that Ragg ratios are concentration-independent between 0.2 and 2.0 μM . (data not shown). For this reason, all SLS measurements in this work were conducted at protein concentrations ranging between 0.3 and 1.5 μM .

Dynamic light scattering

Dynamic light scattering (DLS) was performed using a DynaPro™ instrument (Protein Solutions™, Wyatt Technology Corporation). Protein preparations were adjusted to 5 μM in buffer A. Sample temperature was maintained at 20 °C. During the measurements, the total scattering signal was maintained between 200 and 900 dpi by adjusting the instrument sensitivity to 10–20%. Scattering data were analyzed using the Dynamics software (Protein Solutions™). The hydrodynamic radii reported in this work were obtained using the regularisation algorithm in Dynamics, which performs a multiple exponential fit of the raw data and provides a size distribution analysis for samples containing two or more populations. Interpretation using single and double exponential mono- and bimodal models in Dynamics were performed to further corroborate the findings of the regularisation algorithm.

Ultracentrifugation

The sedimentation velocity v of a spherical particle of radius r is proportional to the sedimentation constant s , the distance x from the center of rotation and the angular frequency ω :

$$v = s\omega^2x$$

$$s = \frac{M(1 - \bar{V}_2\rho)}{N_0f}$$

where f is described by the Stokes law [20] and is linearly dependent with r ; M is the molecular weight; $1 - \bar{V}_2\rho$ is the resulting buoyancy term; N_0 is the Avogadro's number.

According to this relation, monomeric MBP-E6 (MW = 60 kDa) subject to a force of 110,000g has a sedimentation velocity v of approximately 0.16 cm/h when situated in the middle of the centrifuge tube (i.e. at a distance $x = 8$ cm from the centre of rotation). On the other hand, a particle consisting of 10 MBP-E6 units ($M_W = 600$ kDa) in the same conditions will travel at approximately 0.5 cm/h. However, the sedimentation velocity is counterbalanced by the diffusion, which in the case of a monomeric MBP-E6 particle at 4 °C in water is around 0.1 cm/h.

Therefore, to separate the fraction of non-aggregated MBP-E6 from soluble aggregates, we choose to subject MBP-E6 preparations to a 14-h ultracentrifugation at

110,000g (i.e. 36,000 rpm) at 4 °C in a swing rotor SW-41 (Beckman Coulter, Fullerton, CA).

Electron microscopy

MBP-E6 preparations were expressed at 15 °C and affinity-purified. Samples were diluted to a final concentration of 30 ng/ μl in buffer A. Five microliters were deposited on an air glow-discharged carbon-coated grid. After 2 min adsorption samples were negatively stained with a 2% (w/v) uranyl acetate solution. Images were acquired on a Philips CM120 transmission electron microscope equipped with a LaB6 filament and operating at 100 kV. Areas covered with individual molecules were recorded under low dose condition, at a magnification of 45,000 \times on a Pelletier-cooled CCD camera (Model 794, Gatan, Pleasanton, CA).

Determination of Zn^{2+} content (TSQ assay)

Zn^{2+} concentrations in E6 samples were determined monitoring changes in fluorescence of the zinc selective fluorophore TSQ ((*N*-6-methoxy-8-quinolyl)-*p*-toluenesulfonamide) (Molecular Probes, Eugene, OR). The protocols employed followed published procedures [16], with minor modifications. In brief, the protein bound zinc ions were ejected by incubating 2.5–5.0 μM of MBP-E6 with 0.6% of H_2O_2 in buffer T (10 mM sodium phosphate buffer, pH 7.0, and 10% glycerol) for 2 h at room temperature. Hence, oxidized MBP-E6 samples were diluted 4-fold in TSQ buffer and 100 μl aliquots mixed with equal volumes of buffer TS (10 mM sodium phosphate buffer, pH 7.0, 10% glycerol, and 200 μM TSQ). Immediately after, the fluorescence at 477 nm (excitation wavelength 346 nm) was measured. The background fluorescence of the unliganded TSQ dye was measured and subtracted from the values obtained for the MBP-E6 samples. The zinc concentrations were calculated relative to ZnSO_4 standards covering the concentration range between 0.2–4.0 μM . Protein concentrations were determined by denaturing protein samples in 7 M urea and measuring absorbance at 280 nm. The eight tryptophan residues of MBP contribute to most of the absorption signal. However, depending on the HPV strain considered, E6 proteins contain zero to two tryptophans. Extinction coefficients were calculated by adding the contributions of tyrosine and tryptophan residues present in the sequence of the different MBP-E6 constructs.

p53 degradation assays

p53 proteins were synthesized *in vitro* using the TNT T7/SP6 reticulocyte lysate transcription/translation system (Promega) as previously described [21]. p53 degradation was carried out by incubating 4 μl of p53 translation product with 0.4 μM MBP-E6 proteins in 25 mM Tris–HCl, pH 7.5, 100 mM NaCl, and 2 mM DTT at 29 °C for 3 h (total

reaction volume 5 µl). Control reactions were made by replacing the MBP-E6 proteins with equivalent volumes of buffer A. Reactions were stopped by adding 5 µl of 2× SDS-loading buffer and analyzed by SDS-PAGE (12% acrylamide). Gels were dried on Whatman 3M paper and analyzed by autoradiography.

Results

Subcloning and fast cysteine/serine mutagenesis of a panel of HPV E6 proteins for expression as MBP-fusions

We cloned a panel of E6 proteins from various HPV strains (5, 11, 16, 18, 33, 45, 52, and 58) for bacterial expression as fusions to the C-terminus of MBP. All HPV E6 proteins contain eight strictly conserved cysteines that are required for zinc-binding. In addition, an unusually high proportion of cysteines, which are not conserved, are interspaced at various positions in the sequences of E6 proteins. To minimize the risk of protein aggregation deriving from disulfide cross-bridging we mutated to serines most of non-conserved cysteines in the E6 proteins investigated in this work. Given the high number of cysteine/serine mutations required for each protein we used a strategy which merges the steps of sub-cloning and of mutagenesis [18]. This strategy consists of generating PCR fragments containing the required mutations, which carry at their ends restriction sites for class II enzymes. These sites are positioned so that the digested PCR products present cohesive ends that enable their multiple ligation into the expression vector, therefore leading to a perfectly reconstituted cDNA containing all the planned mutations. The principle of this approach is illustrated in Figs. 1 and 2 and the details are fully described in Materials and methods. The aligned

sequences of the E6 proteins sub-cloned and the sites of cysteine/serine mutagenesis are reported in Fig. 3.

Effects of induction temperature on MBP-E6₁₆ soluble aggregation as monitored by light scattering measurements

Lower induction temperatures are thought to improve folding by slowing down over-expression in the bacterial cell and consequently to minimize aggregation events deriving from protein misfolding. In a previous study we have observed a decrease of the aggregation levels of MBP-E6₁₆ when expression temperature was decreased from 37 to 27 °C [15]. Here, we further investigate the effects of lower temperatures. Small-scale (25–50 ml) bacterial cultures of MBP-E6₁₆ were induced at 37 °C for 3 h, at 27 °C for 5 h, at 15 °C for 17 h and at 8 °C for 2 days. Similar expression levels were obtained at all conditions (data not shown). Protein preparations were obtained by amylose affinity-chromatography. We previously introduced an “aggregation ratio” parameter (Ragg) which consists of the ratio between the intensities of scattered light and tryptophan fluorescence both measured in a spectrofluorometer [15]. In the present work, we slightly modified this assay. We have performed fluorescence measurements in denaturing conditions, so that these measurements, which quantify protein concentrations in solution, are no more influenced by folding and/or aggregation (see Materials and methods). Dramatic differences in static light scattering by MBP-E6₁₆ produced at the various temperatures are observed. Maximal and minimal Ragg values are seen for the proteins expressed at, respectively, 37 and 15 °C (Fig. 4A). Yet even the lowest light scattering value measured (Ragg = 0.5) indicates the presence of soluble aggregates. By comparison, a monodisperse solution of

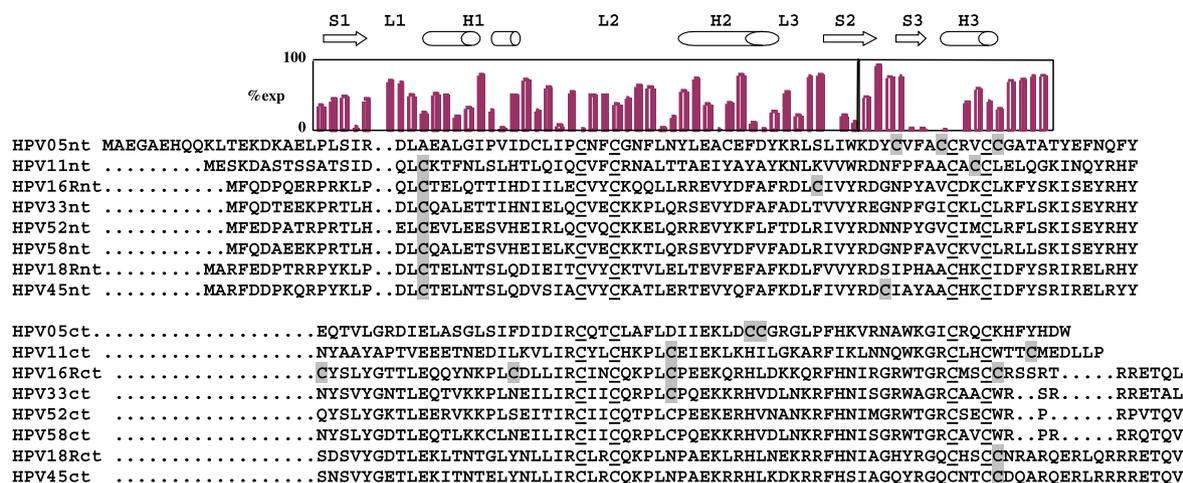


Fig. 3. Sequence alignment of E6 proteins investigated in this work. HPV strain type is reported on the left hand side; nt and ct indicate respectively the N-terminal and C-terminal Zn²⁺ binding domains (E6-N and E6-C). The histogram indicates the percentage of side chain exposure to the solvent of the corresponding amino acid positions and it is derived from the structure of the E6-C domain [13]. Secondary structure elements of the E6-C domain are indicated. C stands for conserved cysteines involved in Zn²⁺ coordination; S indicates non-conserved cysteines which have been replaced by serines in the constructs used in the present work.

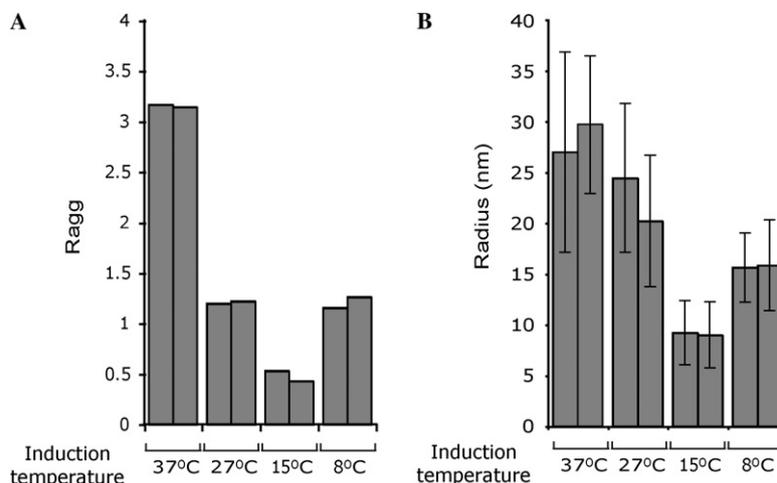


Fig. 4. Influence of induction temperature on the formation of soluble aggregates of MBP-E6 as measured by static (A) and dynamic (B) light scattering (SLS and DLS). Expression temperatures are indicated. To show the reproducibility of the data, experiments are performed in duplicate, using two distinct protein preparations. (A) Ragg parameters as derived from SLS measurements. Ragg parameters are equal to I_{350}/I_{340}^* , that is the intensity of the light scattered by the sample at 350 nm divided by the tryptophan fluorescence at 340 nm measured in denaturing conditions. Fluorescence values are normalized by the number of tryptophan residues of each MBP-E6 fusions. All measurements are performed at low protein concentrations, ranging between 0.3 and 1.5 μM . (B) Hydrodynamic radii obtained from DLS. Error bars represent size polydispersity and are derived from the half-maximal width of the Gaussian distribution of the radius measured. The expected radius for a monomeric 60 kDa MBP-E6 protein is 3.0 nm. Sample concentrations were adjusted to 5 μM .

monomeric MBP displays a Ragg value of ~ 0.002 (data not shown).

MBP-E6₁₆ preparations were also analyzed by dynamic light scattering (DLS). DLS data were acquired on the same samples and on the same day as the SLS measurements. Sample concentration was adjusted to 5 μM . Hydrodynamic radii were derived from the scattering data assuming a spherical model as described in Materials and methods. Hydrodynamic radii of the aggregated MBP-E6 samples vary from 28 to 10 nm corresponding to molecular weights of respectively, 8000–750 kDa, according to the calibration curve provided by the program used for data analysis (see Materials and methods). This, in turn, corresponds to particles consisting of approximately 130 and 12 MBP-E6 monomer units. In agreement with previous DLS data on MBP-E6₁₆ [17] the polydispersity values reported in Fig. 4B are rather low (approximately 30% of the hydrodynamic radius) indicating that multimeric assemblies of MBP-E6₁₆ are fairly homogeneous in size. The DLS data show the same general trend as the SLS results, with the maximal and minimal average aggregate sizes being observed for the samples expressed at 37 and 15 °C, respectively (Fig. 4B).

Light scattering analysis of the aggregation properties of HPV E6 proteins

We therefore proceeded to investigate the aggregation properties of other E6 proteins, including those from HPV strains 5, 11, 16, 18, 33, 45, 52, and 58. To study the specific effects of induction temperature on the different MBP-E6 proteins, we compared MBP-E6 preparations

obtained at 27 and 15 °C. Samples were purified as described in the previous paragraph and analyzed by SLS immediately followed by DLS measurements.

Interestingly, lowering of induction temperature significantly reduces Ragg parameters (by approximately 2-fold) only for some of the E6 proteins, including those from HPV 5, 16, 52, and 58. In contrast, Ragg parameters from other E6 proteins, such as 11, 18, 33, and 45, do not vary considerably within this temperature range (Fig. 5A). MBP-E6₁₆ preparations display the lowest scattering values, with an average Ragg of 0.5. Intermediate Ragg values, between 1 and 1.25, are found for E6₁₁ and E6₁₈ and E6₃₃ preparations. The highest Ragg values, between 1.5 and 2.0, are instead found for samples of E6₅, E6₃₃, E6₄₅ and E6₅₂, and E6₅₈ (Fig. 5A). Overall, over-expression at 15 °C leads to Ragg values ranging between 0.5 and 2.0 for all E6 proteins analyzed. This indicates that all these preparations contain soluble MBP-E6 aggregates.

Interestingly, changes in Ragg values of MBP-E6 preparations analyzed by SLS (Fig. 4A) do not always correlate with changes in particle size analyzed by DLS (Fig. 4B). The marked decrease in particle size occurring for HPV 16 E6 upon reduction of induction temperature is only observed for variants 11 and 45. These E6 variants do not, on the other hand, display strong temperature-dependent variations in their Ragg values such as those observed for HPV16 E6. Overall, the hydrodynamic radii of soluble aggregates in MBP-E6 preparations obtained at 15 °C vary between 9 and 17 nm (corresponding to molecular weights of approximately 600 and 2500 kDa, respectively). Similarly to HPV 16 E6, low polydispersity values are

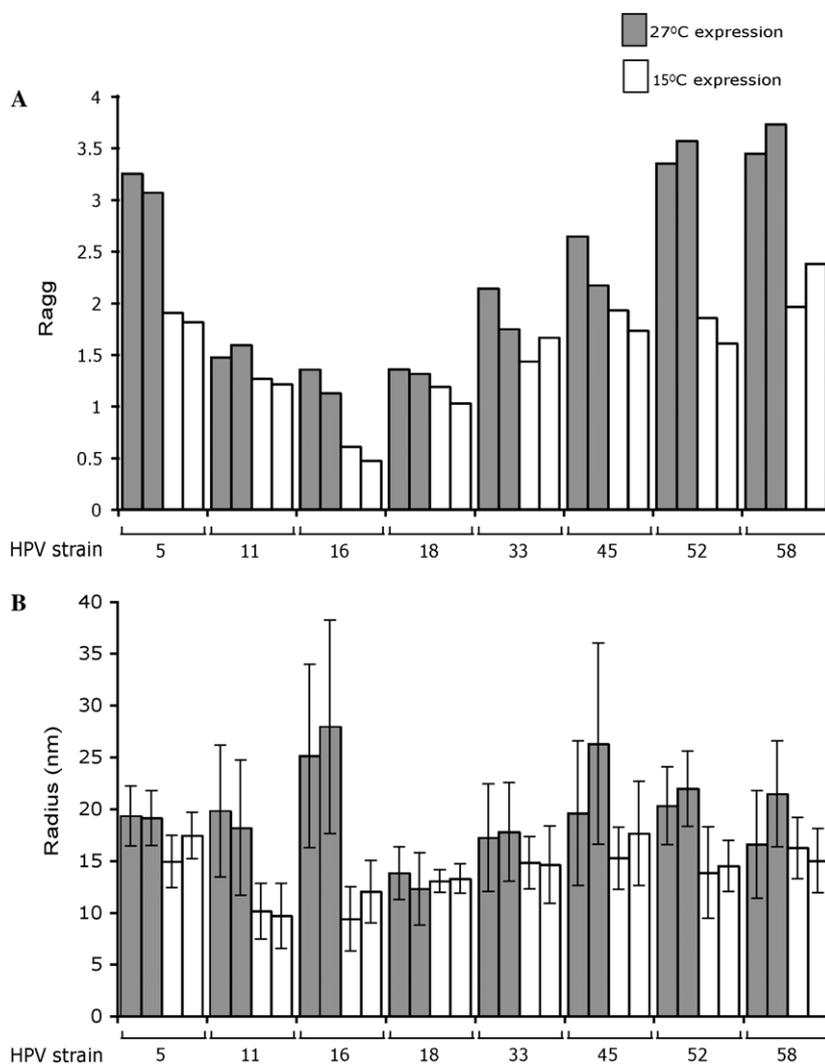


Fig. 5. Comparison of the aggregation properties of different E6 variant proteins. Ragg parameters (A) and hydrodynamic radii (B) are presented for the different MBP-E6 samples. HPV strain types are indicated. Grey and white histograms indicate MBP-E6 preparations over-expressed, respectively, at 27 and 15 °C. Experiments are performed in duplicate, using two different protein preparations, to report on the reproducibility of the data. See also legend of Fig. 4.

observed for the other HPV E6 proteins, particularly for preparations expressed at 15 °C (between 15 and 30% of the average radius, see Fig. 5B).

Ultracentrifugation quantifies the amounts of protein involved in aggregated particles within E6 preparations

We used ultracentrifugation to pellet MBP-E6 soluble aggregates and analyzed the protein recovered in the supernatant. We limited our analysis to samples over-expressed at 15 °C. MBP-E6 preparations from the different HPV variants were affinity purified by amylose chromatography and their concentration adjusted to 12 μ M. Samples were subject to a 14-h ultracentrifugation at 4 °C and 110,000g in a swing rotor SW-41 (Beckman). According to the calculations (see Materials and methods) such centrifugation conditions, applied to a protein preparation diluted in a buffer presenting approximately the

viscosity of water, are suitable for pelleting large molecular weight particles (above 600 kDa) whereas low molecular weight particles (150 kDa or below) should remain in the supernatant. After ultracentrifugation, pellets containing MBP-E6 protein were visible and supernatants displayed negligible light scattering, confirming clearance from MBP-E6 aggregates (data not shown). The percentages of recovery of MBP-E6 in the supernatants are reported in Fig. 6. The recovery varies for the different E6 variants analyzed, with E₁₆ preparations displaying the largest fraction of non-aggregated protein (36%). For E₁₁, E₁₈, and E₃₃ preparations the percentages of recovery vary between 15 and 23%. For E₅, E₄₅, E₅₂, and E₅₈ preparations the recovery is instead below 10%. Altogether, these results are in agreement with the Ragg parameters obtained from the SLS assay and confirm that variations in light scattering are associated with different levels of aggregation.

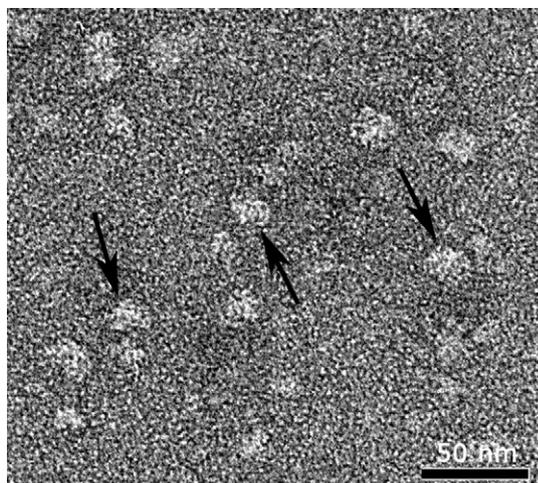


Fig. 6. Partial field of a micrograph showing negatively stained MBP-E6₁₆ soluble aggregates. Spherical isolated particles are indicated (arrows). Bar = 50 nm.

MBP-E6 preparations do not contain major amounts of GroEL chaperone

MBP-fusion proteins have often been reported to co-purify with GroEL [22,23]. This abundant bacterial chaperone consists of two heptameric rings of a single polypeptide unit of approximately 55 kDa. We reasoned that the large soluble particles observed in MBP-E6 preparations might be the result of GroEL sequestering misfolded E6 moieties. The molecular weight of the GroEL polypeptide is close to that of MBP-E6 proteins (~60 kDa), thus preventing unambiguous detection of GroEL by SDS-PAGE. To investigate this possibility, we cleaved affinity-

purified MBP-E6 preparations by TEV protease. In all MBP-E6 samples analyzed protease digestion proceeded mostly to completion, yielding two bands of respectively ~40 kDa (MBP) and ~20 kDa (E6) on a SDS gel. Residual bands at 60 kDa detected in some of the preparations, which may be associated with incomplete protease digestion of MBP-E6 and/or possible GroEL contamination, account for less than 5% of the total protein in the sample (Fig. 7). Such findings rule out the possibility of a major GroEL contamination in the MBP-E6 samples.

Soluble aggregates of MBP-E6 form defined spherical assemblies as observed in electron microscopy

We proceeded with the investigation of the structure of soluble MBP-E6 aggregates by direct electron microscopy observation. Affinity purified MBP-E6₁₆ preparations expressed at 15 °C were adjusted to a concentration of 30 ng/μl, deposited on a carbon-coated grid and negatively stained using a uranyl acetate solution. Fig. 8 shows representative images from these preparations, which reveal the presence of remarkably homogenous particles with a well-defined spherical structure and diameters of 18nm ± 2 nm. Likewise, spherical structures of similar size are observed in preparations of MBP-E6₁₁ expressed at 15 °C (data not shown).

MBP-E6 soluble aggregates do not contain disulfide cross-bridges and possess native-like zinc content

We analyzed MBP-E6 preparations containing soluble aggregates by SDS-PAGE in the presence and in the absence of the reducing agent β-mercaptoethanol. No

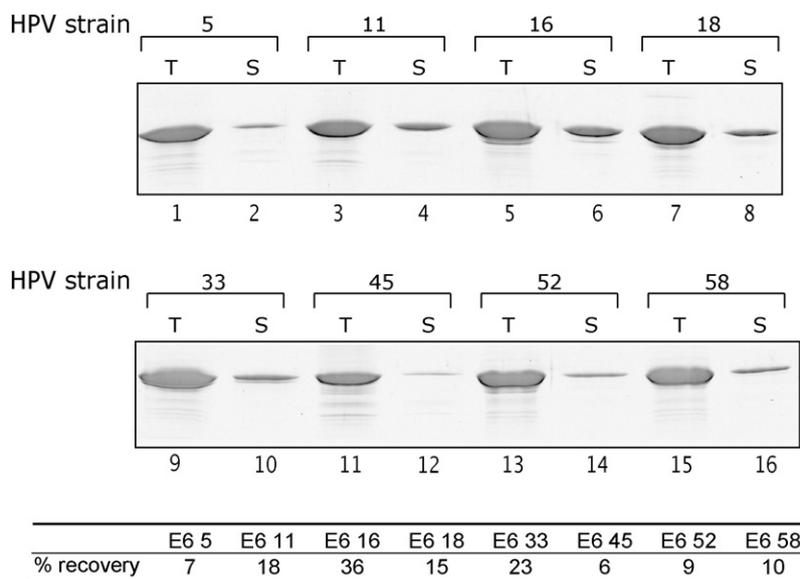


Fig. 7. Recovery of non-aggregated MBP-E6 from different variant preparations after ultracentrifugation. HPV strain types are indicated. MBP-E6 proteins are expressed at 15 °C, affinity-purified and their concentration adjusted to 12 μM. Subsequently, protein samples are subjected to 14-h ultracentrifugation at 110,000g. For all MBP-E6 proteins, equal sample volumes are taken before (T, total) and after ultracentrifugation (S, supernatant) and loaded on a 12% SDS gel (lanes 1–16). The percentages of MBP-E6 recovery after ultracentrifugation are calculated from protein concentrations measured by absorbance at 280 nm.

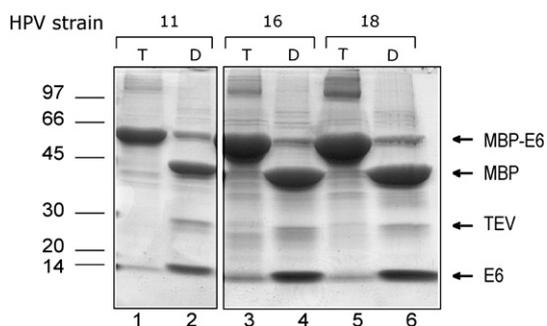


Fig. 8. TEV cleavage of affinity-purified MBP-E6 preparations containing soluble aggregates. Samples were taken before (T, total) and after protease digestion (D, digested) and loaded on lanes 1–6 of a 12% SDS-PAGE gel. HPV strain types are indicated. The migration of the molecular weight markers (in kDa) is reported on the left hand side.

bands could be detected at molecular weights higher than that of a single MBP-E6 polypeptide chain (60 kDa) (data not shown). This indicates that soluble aggregates arise from processes other than chemical cross-linking. Therefore the Cys/Ser mutagenesis approach resulted in the complete suppression of cysteine cross-bridging during purification.

Evidence from work on other zinc binding proteins, such as HPV 16 E7, has shown that loss of metal coordination can induce sample aggregation. We monitored the Zn^{2+} /E6 stoichiometry on a selected set of MBP-E6 preparations (which includes E6 variants 11, 16, and 18), using the selective fluorophore TSQ. MBP-E6 proteins were expressed at 15 °C, purified by amylose affinity chromatography and subjected to overnight ultracentrifugation to pellet soluble aggregates (see above). Samples were taken both before (indicated as T) and after (indicated as S) ultracentrifugation and the Zn^{2+} concentrations measured by the TSQ assay. We found that both samples containing soluble aggregates (which account for 65–85% of total protein, see Fig. 6) and cleared preparations display a Zn^{2+} /E6 stoichiometry close to 2:1 (Table 1). Such finding is expected from the common architecture of HPV E6 proteins, all consisting of two zinc-binding domains. The data show that E6 self-assembly processes observed in this work are not related to loss of zinc ions.

Table 1
Experimentally derived Zn^{2+} :protein stoichiometries for MBP-E6 samples from variants 11, 16 and 18. T (total) and S (supernatant) indicate respectively samples taken before and after ultracentrifugation

MBP-E6 samples	Zn^{2+} : protein stoichiometry
11 T	1.95 ± 0.05
11 S	1.93 ± 0.10
16 T	1.79 ± 0.07
16 S	1.97 ± 0.23
18 T	1.81 ± 0.06
18 S	1.85 ± 0.08

Zinc concentrations are measured by the TSQ assay and calculated relative to $ZnSO_4$ standards (see Materials and methods). Protein concentrations are determined by denaturing protein samples in 7 M urea and measuring absorbance at 280 nm.

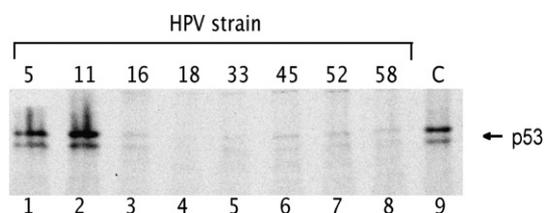


Fig. 9. p53 degradation in the presence of 0.4 μ M of MBP-E6 proteins, which are previously affinity-purified and separated from soluble aggregates by ultracentrifugation. HPV strain types of the MBP-E6 proteins are indicated (lanes 1–8). Lane 9: control.

p53 degradation activities of MBP-E6 preparations cleared from soluble aggregates

E6 proteins containing the Cys/Ser mutations were tested for p53 degradation activity. MBP-E6 proteins from HPV strains 5, 11, 16, 18, 33, 45, 52, and 58 were expressed at 15 °C, affinity-purified and cleared from soluble aggregates by ultracentrifugation. MBP-E6 proteins were incubated with ^{35}S labelled p53 at a constant concentration of 0.4 μ M for 3 h at 29 °C. Fig. 9 shows that all E6 variants from “high-risk” genital HPV strains degrade p53 (lanes 3–8), while the cutaneous E6₅ (lane 1) and the “low-risk” genital E6₁₁ (lane 2) do not seem to affect p53 levels. These data are fully consistent with the known activities of these various E6 proteins relative to p53.

Discussion

When dealing with problematic proteins, MBP-induced solubilisation is often accompanied by the presence of large molecular weight aggregates [24–28]. Such multimeric structures can be purified from bacterial lysates using amylose-affinity chromatography. In our previous work we showed that aggregation of MBP-E6₁₆ originated from at least two distinct phenomena: formation of soluble large molecular weight particles, most probably taking place *in vivo* during protein expression, and intermolecular disulfide bridge cross-linking, occurring *in vitro* during extraction, purification and storage of the proteins [15]. The unusually high number of cysteines found in E6 sequences leads to intermolecular disulfide bridge formation even at mild oxidative conditions. In the present work, we have applied the same strategy that formerly allowed us to isolate stable monomeric E6₁₆ proteins [15], i.e., the mutagenesis of non-conserved cysteines into the isosteric residue serine. The introduction of multiple Cys/Ser mutations in each E6 construct seems to efficiently prevent disulfide bridge formation. On the other hand it does not appear to alter the protein’s sequence and/or structural features required for wild-type activities, such as coordination to Zn^{2+} ions and p53 degradation. Here, we demonstrate that E6 samples have a Zn^{2+} : protein stoichiometry of 2:1, which is predicted from the common architecture of all HPV E6 proteins. We also show that all high-risk genital E6 Cys/Ser mutants analyzed have retained their capability

to degrade p53 *in vitro*. In addition, in a previous work [19] we have shown that some of these E6 mutants (including those from HPV strains 16, 18, 33, and 58) interact with an E6AP derived 18-peptide with affinities in the low micromolar range.

Elimination of oxidative cross-linking during purification has allowed us to proceed to the study of the properties of soluble MBP-E6 aggregates originating from *in vivo* self-assembly pathways. To carry out this task we have combined techniques, such as dynamic and static light scattering and ultracentrifugation, which provide complementary information on the amounts and sizes of large soluble aggregates in the preparations. DLS provides information on the size and dispersity of the soluble aggregated particles, while the SLS signal varies linearly with concentration and with the cube of the particle's radius. Ultracentrifugation, performed in the conditions indicated above, provides instead a quantitative estimate of the amounts of protein associated with a population of large molecular weight aggregates. It is to be noticed that other methods, less demanding in terms of instrumentation, have also been proposed to monitor the presence of aggregated particles in preparations of fusion proteins, such as sandwich ELISA [17], native gel analysis [29] and sucrose gradients ultracentrifugation [30]. A recent review published in [31] includes a comprehensive list of methods used to study protein aggregates.

A striking feature of MBP-E6 soluble multimeric assemblies is their remarkable size homogeneity, revealed by the low polydispersity of all preparations analyzed by DLS. Based on results from immunochemical assays, we previously proposed a micellar model, in which the E6 moieties are buried and hidden from the solvent, whereas MBP units are exposed to the solvent [17]. In the current study, the regular spherical shape observed by electron microscopy provides further support to the micellar model. When protein over-expression is carried out at 15 °C, MBP-E6 multimeric particles from the different HPV strains have similar average hydrodynamic radii. When expression temperature is raised to 27 °C, larger particle radii are observed for some, but not all, MBP-E6 proteins. This direct effect of the expression temperature on particle size further supports the view that the soluble MBP-E6 aggregates result from *in vivo* self-assembly pathways.

We previously proposed that the E6 moieties are not properly folded within the soluble MBP-E6 aggregated particles [17]. This was supported by the fact that proteolysis of the linker separating MBP and E6 generated rapid formation of E6 precipitates which could only be solubilised using strong denaturing agents. On the other hand, we also observed a sharp cooperative transition upon chemical denaturation of MBP-E6 soluble aggregates, which may reflect the fact that E6 moieties within these structures interact *via* a network of identical molecular interfaces [17]. In the present work, we demonstrate that preparations including a large excess of soluble aggregates contain two zinc molecules for each MBP-E6 polypeptide

chain. These findings indicate that the two zinc-binding domains of E6 moieties within soluble aggregates possess a native-like fold, which enables correct coordination to the metal center. These observations also suggest that the E6 self-assembly processes described in this work do not result from loss of zinc binding, a mechanism which has been describe to promote aggregation of a number of zinc binding proteins such as HPV E7 [32] and members of the RING finger family [33].

In the present work, we show that eight E6 proteins of four phylogenetically distinct HPV species form large molecular weight aggregates with similar properties. Such finding suggests that soluble aggregation is a general behavior of MBP fused E6 proteins of most HPV species. The solution structure of HPV16 E6 C-terminal domain has shown that whereas the general fold of HPV E6 proteins seems to be extremely conserved, their surface properties appear extremely variable [13]. Here, we observe that the levels of soluble aggregates vary among the E6 proteins analyzed, ranging between 64% of the total protein, in the case of MBP-E6₁₆, and 93–94%, for MBP-E6₅ and MBP-E6₄₅. We therefore wondered whether surface properties such as hydrophobicity explain the different aggregation tendencies observed for these variants. It is generally accepted that surface hydrophobicity increases the propensity of a protein to aggregate by promoting local unfolding events which destabilize the native three-dimensional structure [34,35]. We therefore examined the number of exposed hydrophobic residues (including Val, Leu, Ile, Trp, Phe, Met, and Tyr) on the surface of the pseudo-dimeric model proposed in [13] for the different E6 proteins. As a result we found no straightforward correlation between surface properties and specific aggregation tendencies of each variant. For instance, E6₁₆, which displays the lowest tendency to polymerize, has fewer exposed hydrophobic residues (10% of the total number of residues) than E6₅ (17%). On the other hand, E6₁₆ surface hydrophobicity is very similar to that of E6₄₅ (with 11% of hydrophobic exposed residues), which, similarly to E6₅, has a high tendency to form soluble aggregates. Hence, it appears that the phenomenon leading to the formation of the large MBP-E6 aggregates may not be driven by the surface properties of these proteins.

The work presented here provides evidence for structural organization within MBP-E6 soluble aggregates. It will be extremely interesting to further analyse the structural features of the E6 moieties present in these aggregated particles. These studies may provide further understanding into the processes of protein over-expression, folding and aggregation during biosynthesis in the host cell.

Acknowledgments

We are grateful to members of the Structural Biology laboratory at IGBMC (Strasbourg) for access and help with the light scattering equipment. We thank Prof. G. Orth (Institut Pasteur, Paris) for providing the plasmids

containing the HPV genomes. K.Z. and S.C. are supported by a fellowship from the Fondation pour la Recherche Medicale (FRM) and Ligue Nationale Contre le Cancer. This work was funded by the following organizations: Arbor Vita Corporation (Sunnyvale, CA), Association pour la Recherche sur le Cancer (ARC), Institut National de la Santé et de la Recherche Médicale, Centre National pour la Recherche Scientifique (CNRS) and Université Louis Pasteur (Strasbourg).

References

- [1] E.M. de Villiers, C. Fauquet, T.R. Broker, H.U. Bernard, H. zur Hausen, Classification of papillomaviruses, *Virology* 324 (2004) 17–27.
- [2] H. zur Hausen, Papillomaviruses in human cancers, *Proc. Assoc. Am. Physicians* 111 (6) (1999) 581–587.
- [3] F. Bosch, M. Manos, N. Munoz, M. Sherman, A. Jansen, J. Peto, M. Schiffman, V. Moreno, R. Kurman, K. Shah, Prevalence of human papillomavirus in cervical cancer: a worldwide perspective. International biological study on cervical cancer (IBSCC) Study Group, *J. Natl. Cancer Inst.* 87 (1995) 796–802.
- [4] H. Pfister, Chapter 8: human papillomavirus and skin cancer, *J. Natl. Cancer Inst. Monogr.* 31 (2003) 52–56.
- [5] B.A. Werness, A.J. Levine, P.M. Howley, Association of human papillomavirus types 16 and 18 E6 proteins with p53, *Science* 248 (1990) 76–79.
- [6] M. Scheffner, J.M. Huibregtse, R.D. Vierstra, P.M. Howley, The HPV-16 E6 and E6-AP complex functions as a ubiquitin–protein ligase in the ubiquitination of p53, *Cell* 75 (1993) 495–505.
- [7] O. Chakrabarti, S. Krishna, Molecular interactions of ‘high risk’ human papillomaviruses E6 and E7 oncoproteins: implications for tumour progression, *J. Biosci.* 28 (3) (2003) 337–348.
- [8] B. Glaunsinger, S. Lee, M. Thomas, L. Banks, R. Javier, Interactions of the PDZ-protein MAGI-1 with adenovirus E4-ORF1 and high-risk papillomavirus E6 oncoproteins, *Oncogene* 19 (46) (2000) 5270–5280.
- [9] M. Thomas, R. Laura, K. Hepner, E. Guccione, C. Sawyers, L. Lasky, L. Banks, Oncogenic human papillomavirus E6 proteins target the MAGI-2 and MAGI-3 proteins for degradation, *Oncogene* 21 (2002) 5088–5096.
- [10] S. Jackson, C. Harwood, M. Thomas, L. Banks, A. Storey, Role of Bak in UV-induced apoptosis in skin cancer and abrogation by HPV E6 proteins, *Genes Dev.* 14 (23) (2000) 3065–3073.
- [11] F. Lipari, G.A. McGibbon, E. Wardrop, M.G. Cordingley, Purification and biophysical characterization of a minimal functional domain and of an N-terminal Zn²⁺-binding fragment from the human papillomavirus type 16 E6 protein, *Biochemistry* 40 (5) (2001) 1196–1204.
- [12] Y. Nominé, S. Charbonnier, T. Ristriani, G. Stier, M. Masson, N. Cavusoglu, A. Van Dorselaer, E. Weiss, B. Kieffer, G. Travé, Domain substructure of HPV E6 protein: biophysical characterization of E6 C-terminal DNA-binding domain, *Biochemistry* 42 (2003) 4909–4917.
- [13] Y. Nominé, M. Masson, S. Charbonnier, K. Zanier, T. Ristriani, F. Deryckere, A.P. Sibling, D. Desplancq, R.A. Atkinson, E. Weiss, G. Orfanoudakis, B. Kieffer, G. Trave, Structural and functional analysis of E6 oncoprotein: insights in the molecular pathways of human papillomavirus-mediated pathogenesis, *Mol. Cell.* 21 (5) (2006) 665–678.
- [14] P.R. Daniels, C.M. Sanders, P. Coulson, N.J. Maitland, Molecular analysis of the interaction between HPV type 16 E6 and human E6-associated protein, *FEBS Lett.* 416 (1997) 6–10.
- [15] Y. Nominé, T. Ristriani, C. Laurent, J.F. Lefèvre, E. Weiss, G. Travé, A strategy for optimizing the monodispersity of fusion proteins: application to purification of recombinant HPV E6 oncoprotein, *Protein Eng.* 14 (4) (2001) 297–305.
- [16] R. Degenkolbe, P. Gilligan, S. Gupta, H.U. Bernard, Chelating agents stabilize the monomeric state of the zinc binding human papillomavirus 16 E6 oncoprotein, *Biochemistry* 42 (2003) 3868–3873.
- [17] Y. Nominé, T. Ristriani, C. Laurent, J.F. Lefèvre, E. Weiss, G. Travé, Formation of soluble inclusion bodies by HPV E6 oncoprotein fused to maltose-binding protein, *Protein Expr. Purif.* 23 (1) (2001) 22–32.
- [18] M. Tomic, I. Sunjevaric, E.S. Savtchenko, M. Blumenberg, A rapid and simple method for introducing specific mutations into any position of DNA leaving all other positions unaltered, *Nucleic Acids Res.* 18 (6) (1990) 1656.
- [19] K. Zanier, S. Charbonnier, M. Baltzinger, Y. Nominé, D. Altschuh, G. Trave, Kinetic analysis of the interactions of human papillomavirus E6 oncoproteins with the ubiquitin ligase E6AP using surface plasmon resonance, *J. Mol. Biol.* 349 (2005) 401–412.
- [20] C.R. Cantor, P.R. Schimmel, *Biophysical Chemistry, Part II: Techniques for the Study of Biological Structure and Function*, Freeman and Company, New York, 1998.
- [21] T. Ristriani, Y. Nominé, L. C. E. Weiss, G. Trave, Protein mutagenesis with monodispersity-based quality probing: selective inactivation of p53 degradation and DNA-binding properties of HPV E6 oncoprotein, *Protein Expr. Purif.* 26 (3) (2002) 357–367.
- [22] B. Bjorndal, G. Trave, I. Hageberg, J. Lillehaug, A. Raae, Expression and purification of receptor for activated C-kinase 1 (RACK1), *Protein Expr. Purif.* 31 (2003) 47–55.
- [23] P. Douette, R. Navet, F. Bouillenne, A. Brans, C. Sluse-Goffart, A. Matagne, F.E. Sluse, Secondary-structure characterization by far-UV CD of highly purified uncoupling protein 1 expressed in yeast, *Biochem. J.* 380 (Pt 1) (2004) 139–145.
- [24] J.M. Louis, R.A. McDonald, N.T. Nasheed, E.M. Wondrak, D.M. Jerina, S. Oroszlan, P.T. Mora, Autoprocessing of the HIV-1 protease using purified wild-type and mutated fusion proteins expressed at high levels in *Escherichia coli*, *Eur. J. Biochem.* 199 (1991) 361–369.
- [25] V.M. Saavedra-Alanis, P. Rysavy, L.E. Rosenberg, F. Kalousek, Rat liver mitochondrial processing peptidase: both α - and β -subunits are required for activity, *J. Biol. Chem.* 269 (1994) 9284–9288.
- [26] H.K. Lorenzo, D. Farber, V. Germain, O. Acuto, P.M. Alzari, The MBP fusion protein restores the activity of the first phosphatase domain of CD45, *FEBS Lett.* 411 (1997) 231–235.
- [27] D. Sachdev, J.M. Chirgwin, Properties of soluble fusions between mammalian aspartic proteases and bacterial maltose-binding protein, *J. Protein Chem.* 18 (1999) 127–136.
- [28] S. Swalley, B. Baker, L. Calder, S. Harrison, J. Skehel, D. Wiley, Full-length influenza hemagglutinin HA2 refolds into the trimeric low-pH-induced conformation, *Biochemistry* 43 (2004) 5902–5911.
- [29] J. Stegemann, R. Ventzki, A. Schrodell, A. de Marco, Comparative analysis of protein aggregates by blue native electrophoresis and subsequent sodium dodecyl sulfate–polyacrylamide gel electrophoresis in a three-dimensional geometry gel, *Proteomics* 5 (8) (2005) 2002–2009.
- [30] A. Schrodell, A. de Marco, Characterization of the aggregates formed during recombinant protein expression in bacteria, *BMC Biochem.* 6 (10) (2005).
- [31] W. Wang, Protein aggregation and its inhibition in biopharmaceutics, *Int. J. Pharm.* 289 (1–2) (2005) 1–30.
- [32] L. Alonso, M. Garcia-Alai, C. Smal, J. Centeno, R. Iacono, E. Castano, P. Gualfetti, G. de Prat-Gay, The HPV16 E7 viral oncoprotein self-assembles into defined spherical oligomers, *Biochemistry* 43 (12) (2004) 3310–3317.
- [33] A. Melnick, J.D. Licht, Deconstructing a disease: RAR α , its fusion partners, and their roles in the pathogenesis of acute promyelocytic leukemia, *Blood* 93 (10) (1999) 3167–3215.
- [34] G. Calloni, S. Zoffoli, M. Stefani, C.M. Dobson, F. Chiti, Investigating the effects of mutations on protein aggregation in the cell, *J. Biol. Chem.* 280 (11) (2005) 10607–10613.
- [35] S.Y. Patro, T.M. Przybycien, Simulations of kinetically irreversible protein aggregate structure, *Biophys. J.* 66 (5) (1994) 1274–1289.