NIH-PA Author Manuscript

Strategies for bacterial expression of protein-peptide complexes: application to solubilization of Papillomavirus E6

Abdellahi Ould M'hamed Ould Sidi[1], **Khaled Ould Babah**[1], **Nicole Brimer**[2], **Yves Nominé**[1], **Christophe Romier**[3], **Bruno Kieffer**[3], **Scott Vande Pol**[2], **Gilles Travé**[1],*, and **Katia Zanier**[1],*

[1]Ecole Supérieure de Biotechnologie de Strasbourg (UMR 7242), Boulevard Sébastien Brant, BP 10413, 67412 Illkirch, France.

[2]Department of Pathology, University of Virginia, Charlottesville, VA 22908, USA.

[3]Institut de Génétique et de Biologie Moléculaire et Cellulaire (UMR 7104, INSERM U 964, Université de Strasbourg), 1 rue Laurent Fries, BP 163, 67404 Illkirch, France.

Abstract

E6 is a small oncoprotein involved in tumorigenesis induced by papillomaviruses (PVs). E6 often recognises its cellular targets by binding to short motifs presenting the consensus LXXLL. E6 proteins have long resisted structural analysis. We found that Bovine Papillomavirus Type 1 (BPV1) E6 binds the N-terminal LXXLL motif of the cellular protein paxillin with significantly higher affinity as compared to other E6/peptide interactions. Although recombinant BPV1 E6 was poorly soluble in the free state, provision of the paxillin LXXLL peptide during BPV1 E6 biosynthesis greatly enhanced the protein's solubility. Expression of BPV1 E6/LXXLL peptide complexes was carried out in bacteria in the form of triple fusion constructs comprising, from N- to C-terminus, the soluble carrier protein MBP (Maltose-Binding-Protein), the LXXLL motif and the E6 protein. A TEV protease cleavage site was placed either between MBP and LXXLL motif or between LXXLL motif and E6. These constructs allowed us to produce highly concentrated samples of BPV1 E6, either covalently fused to the C-terminus of the LXXLL motif (intra-molecular complex) or non-covalently bound to it (inter-molecular complex). Heteronuclear NMR measurements were performed and showed that the E6 protein was folded with similar conformations in both covalent and non-covalent complexes. These data open the way to novel structural and functional studies of the BPV1 E6 in complex with its preferential target motif.

Keywords

BPV; E6; paxillin; LXXLL motif; protein aggregation; protein-peptide interactions

INTRODUCTION

Papillomaviruses (PV) are small DNA viruses that infect epithelial tissues and produce hyperplasias within mammals, birds, and reptiles. Almost 200 PV types have been identified

*Corresponding authors: Phone: 0033 3 68854406, fax: 0033 3 68854718 gilles.trave@unistra.fr, zanier@unistra.fr .

[1]. PVs display distinct tropisms for different body sites (skin, mouth and genitalia) as well as distinct levels of pathogenicity. Although papillomaviruses initially produce benign hyperplasias, in a subset of virus types termed "high-risk", these lesions can progress to cancer. In humans, cervical cancers are caused by infection with high-risk genital Human Papillomaviruses (HPVs) [2] of which HPV16 is the most common "high-risk" genital type [3]. Bovine Papillomavirus Type 1 (BPV1) is one of the most extensively studied animal papillomavirus, serving as a model system for papillomavirus transcription, transformation and replication [4].

PV tumorogenesis results primarily from the activities of two virally encoded oncoproteins, E6 and E7. E6 proteins from "high-risk" genital HPVs bind directly to the cellular E3 ubiquitin ligase E6-Associated Protein (E6AP) [5] and then recruits and forms a trimeric complex with the tumor suppressor p53, leading to ubiquitin–mediated degradation of p53 [6]. "High-risk" genital HPV E6 proteins also target many other cellular proteins to degradation (reviewed in [7]), including PDZ-containing proteins, which recognise a short motif at the extreme C-terminus of E6 [8].

HPV16 E6 binds directly to several of its targets proteins (including E6AP [9], E6BP [10], IRF-3 [11] and possibly p300/CBP [12, 13]) *via* "charged leucine" motifs, also called "LXXLL" motifs. These motifs contain the consensus LxxφLsh sequence, where L indicates conserved leucine residues, φ is an hydrophobic residue (usually a leucine), h is an amino acid residue with a side chain capable of accepting hydrogen bonds (Asp, Glu, Asn, or Gln), s represents a small amino acid residue (Gly or Ala) and xx is a dipeptide where one of the residues is Asp, Asn, Glu or Gln [10, 14]. BPV1 E6 also directly binds the LXXLL motif of E6AP, as well as LXXLL motifs found on the focal adhesion protein paxillin and possibly p300 and the AP1 adapter complex [15-17]. This interaction is required for cellular transformation by BPV1 E6 [17, 18].

Most mammalian PV E6 proteins are cysteine-rich small proteins of about 150 amino acids, consisting of two zinc-binding domains, E6-N and E6-C [19, 20]. We solved by NMR the solution structure of a mutant of HPV16 E6-C domain [21]. To date the three-dimensional structure of any entire E6 protein has not been determined, due to difficulties in its recombinant production. Attempts to produce the native full-length protein in *E. coli* have mainly focussed on E6 from HPV16 [19, 22-25]. Recently, we also explored the behavior of several E6 proteins of other HPV types, including major "high-risk" genital HPV18, 33, 45, 52, 58, "low-risk" genital HPV11 and high-risk cutaneous HPV5 [26, 27]. None of the HPV E6 proteins investigated raised samples amenable to structural analysis.

In the present work, we explored the suitability of various constructs of the BPV1 E6 protein for biophysical and structural analysis. We found that binding of BPV1 E6 to the to the LXXLL motif of paxillin, in either an inter-molecular complex (non-covalent) or intra-molecular complex (covalent, obtained by fusion of the LXXLL containing peptide to the N-terminus of the E6 protein), greatly enhanced the E6 protein's stability.

MATERIALS AND METHODS

MBP-E6 fusion constructs

Three constructs were generated where the DNA encoding for full-length BPV1 E6 (137, Genbank CAB 4650) was inserted in the pETM41 vector (kindly provided by G. Stier, EMBL, Heidelberg). In the first construct, BPV1 E6 was cloned at the C-terminus of 6His-tagged maltose binding protein (MBP) via a linker sensitive to TEV protease as previously described [26]. The second construct, referred to as MBP-tevs-LXXLL-E6, consists of the 10 amino acids of the N-terminal E6-binding LXXLL motif of paxillin (residues 1-10) fused

to the N-terminus of E6 via a short 7 residue linker (detailed sequence: [6HisMBP]-[MSENLYFQGAMDDLDALLADKESGGSA]-[E6]). In the third construct, named MBP-LXXLL-tevs-E6 we fused directly the LXXLL motif to the C-terminus of 6his-MBP and introduced the TEV sensitive linker at the junction between the LXXLL motif and the E6 sequence (detailed sequence: [6HisMBP]-[MDDLDALLADGGSGSNENLYFQGSG]-[E6]). All clones were verified by DNA sequencing.

GST-peptide constructs

DNA oligomers coding for the negative control GSN-rich peptide (sequence : GSNSGNGNS), the paxillin (residues 2-10) peptide (sequence : DDLDALLADKE) and the E6AP (residues 403-417) peptide (sequence: ESSELTLQELLGEER) were cloned into the NcoI/KpnI sites of the pETM-30 expression vector containing a N-terminal 6HisGST tag and a TEV protease cleavage site.

Expression and purification procedures

BPV1 E6 fusion constructs were expressed overnight at 15 °C in *Escherichia coli* BL21(DE3) cells grown in either LB or M9 medium supplemented with ^{15}N labelled ammonium sulfate salts. All purification buffers were extensively degassed and equilibrated with argon. Expression cultures were harvested by centrifugation and resuspended in buffer A (Tris 50 mM pH 6.8, NaCl 400 mM, and DTT 2 mM) supplemented with 5% glycerol, 2.5 µg/ml DNase I, 2.5 µg/ml RNase I, 1 µg/ml lysozyme and anti-protease cocktail (Roche Applied Science) at the concentration recommended by the manufacturer. Cells were broken by sonication on ice and then centrifugated at 38,000 rpm for 1h at 4 °C. The supernatant was loaded on an affinity amylose resin column (New England Biolabs) equilibrated in buffer A. After extensive washing, the protein fusion was eluted using buffer A containing 10 mM maltose. Hence, the affinity purified sample was ultracentrifuged overnight at 110,000 *g* and 4 °C in a SW41 rotor (Beckman). Subsequently, the MBP moiety was removed by TEV protease digestion carried out at 8 °C for 12 up to 36 hours. In the case of the 6His-MBP-LXXLL-tevs-E6 TEV protease digestion was performed in the presence of a four-fold stoichiometric excess of synthetic peptide derived from paxillin (see Results section for further details). The digested product was concentrated and loaded on a HiLoad 16/60 Superdex 75 gel filtration column (Amersham Biosciences) previously equilibrated in buffer B (20 mM sodium phosphate buffer (pH 6.8), NaCl 200 mM, and DTT 2 mM). Protein concentration was estimated by absorption at 280 nm.

GST-peptide fusions for SPR binding experiments (see below) were over-expressed from 20 ml LB cultures of BL21 DE3 *Escherichia coli* cells for 3 hours at 37 °C. Cells were harvested and resuspended in 1.5 ml of buffer A supplemented with 5% glycerol, 1% NP-40, 1 µg/ml DNase I, 1 µg/ml Rnase I, 1 µg/ml lysozyme, anti-protease inhibitor cocktail, broken by sonication on ice in 2.5 ml eppendorf tubes and centrifuged in a microfuge at maximal speed and at 4 °C for 15 min. Supernatants were thus coupled to the Biacore sensor surface.

Surface Plasmon Resonance experiments

Surface Plasmon Resonance (SPR) experiments were performed at 25 °C on a BIAcore 2000 instrument (Biacore) equilibrated in buffer B. The experimental set-up followed was described in a previous publication [28]. Briefly, peptides containing the LXXLL motives of paxillin and E6AP and a negative control peptide were expressed fused to GST (see above) and captured onto the Biacore's surface *via* a polyclonal anti-GST antibody. The analyte protein was monomeric BPV1 E6 fused to MBP (MBP-E6) isolated by gel-filtration 24 hours prior to injection and concentrated to a stock concentration below the aggregation threshold (i.e. 20 µM). We previously demonstrated that the MBP tag does not significantly

alter binding properties of E6 proteins to their target peptide while permitting the preservation of the E6 as a soluble monomer and an increase of the SPR signal [29]. The analyte MBP-E6 protein was therefore injected and the interaction with the captured peptides monitored. Immediately after synthetic peptide containing the LXXLL motif from either E6AP or paxillin (depending on whether GST-E6AP or GST-paxillin was immobilized on the surface) was injected at a concentration of 20 μM during the analyte's dissociation phase, thus avoiding a regeneration step. The negative control peptide sensorgram was used to correct for non-specific binding whereas buffer effects were corrected by subtracting buffer injection sensorgrams. All data were interpreted using the BiaEvaluation 3.2 software (Biacore). Kinetic data were interpreted by global fit using a simple Langmuir interaction model. The kinetic parameters reported were obtained from fits with χ^2 values smaller than 4% of the globally fitted Rmax values.

Synthetic peptides

The paxillin and E6AP derived synthetic peptide presented the following sequences: EFTMDDLDALLADKE and ESSELTLQELLGEER.

NMR spectroscopy

^{1}H,^{15}N HSQC spectra were acquired on a Bruker DRX600 spectrometer equipped with a z-gradient triple-resonance cryoprobe at 296 K. Samples were buffer exchanged against 20 mM sodium phosphate buffer (pH 6.8), 200 mM NaCl, 2 mM DTT, 10% D_2O.

Mass spectrometry

Samples for MALDI-TOF mass spectrometry were recovered from 15% SDS-PAGE gels and subjected to trypsin digestion according to standard protocols. Measurements were performed of a Reflex IV sperctrometer (Bruker, Germany). Samples for electrospray ionization mass spectrometry were buffer exchanged against 100 mM ammonium acetate at pH 7.0. Native spectra were recorded at a sample concentration of 12 μM. For the measurements in denaturing conditions the sample was diluted in a 1:1 water/acetonitrile (v/v) mixture acidified with 1% formic acid to achieve a concentration of 5 μM. Measurements were performed on an MicOTOF mass spectrometer (Bruker, Germany).

RESULTS

Domain architecture and stability of BPV1 E6

Sequence alignement of BPV1 E6 versus E6 proteins from a subset of HPV species (Fig. 1A) shows that zinc-binding cystein residues as well as most of key buried hydrophobic positions are conserved across human and bovine PV species. These observations suggest that BPV1 E6 contains two zinc-binding domains with folds similar to the HPV counterparts.

We therefore applied to BPV1 E6 the same expression and purification protocols developed for the other HPV E6 proteins [26]. The wild-type sequence of full-length BPV1 E6 was fused to the C-terminus of Maltose Binding Protein (MBP) *via* a linker sensitive to TEV protease, over-expressed and affinity purified on amylose resin. Overnight ultracentrifugation to clear soluble MBP-E6 aggregates led to a significant loss of material in the supernatant (Fig. 1B, lanes 2 and 3). The cleared sample was then subject to TEV protease cleavage. The identity of the BPV1 E6 band generated after TEV digestion and migrating at 16 kDa (Fig. 1B, lane 4) was confirmed by MALDI-TOF mass spectrometry analysis (Table S1, Supplementary data). The digested sample was subsequently loaded on a gel-filtration column (Fig. 1C). We could only recover low amounts of protein at the elution volume expected for monomeric BPV1 E6 separated from the MBP tag. In addition gel

filtrated E6 could not even stand further concentration due to precipitation onto the concentrator's membrane.

We have previously observed that the MBP tag slows down the formation of higher-order multimeric species of E6 proteins, including BPV1 E6 [27]. Indeed monomeric MBP-BPV1 E6 fusion previously purified by gel filtration chromatography could be concentrated up to 20 μM without undergoing precipitation (*data not shown*).

BPV1 E6 binds with high-affinity to the LXXLL motif of paxillin

The BPV1 E6 protein has been shown to bind to several peptides bearing the LXXLL consensus motif [17, 18]. We investigated the binding of BPV1 E6 to two distinct peptides containing respectively the LXXLL motif of E6AP and the N-terminal LXXLL motif of paxillin by SPR (Biacore). Freshly purified monomeric MBP-E6 was injected at varying concentrations on sensor surfaces capturing the GST-peptide fusions and the interactions monitored.

The sensorgrams shown in Fig. 2 denote a remarkably slow dissociation phase for the E6/paxillin interaction as compared to E6/E6AP. Kinetic analysis using a simple bimolecular Langmuir interaction model yielded k_{off} rate values of 0.2 and $1.7 \times 10^{-3}\,M^{-1}\,s^{-1}$ for the E6AP and paxillin peptides respectively. The equilibrium (K_D), values estimated from the k_{off}/k_{on} ratios are 1.3 μM and 35.1 nM (Table 1). Therefore, whereas the affinity of the BPV1 E6 / E6AP interaction approximates that of HPV16 E6 / E6AP (4 μM) [29], the affinity of BPV1 E6 for the paxillin peptide is significantly stronger.

Strategies for the purification of complexes of BPV1 E6 bound to N-terminal LXXLL motif of paxillin

The strong affinity of BPV1 E6 for the N-terminal LXXLL motif of paxillin suggested us the possibility that this interaction might stabilize E6 in a conformation that is not any more prone to self-association. Based on this assumption we have designed a number of expression and purification strategies for the production of samples of BPV1 E6/paxillin complexes.

Initially we have tested a number of co-expression and co-purification protocols. Co-purification of GST-paxillin peptide and MBP-BPV1 E6 fusion expressed in separate bacterial cultures was attempted on either amylose (MBP-binding) or gluthatione (GST-binding) resins. This approach yielded very little amounts of complex, most probably because MBP-E6 aggregates pre-formed upon bacterial expression were not proficient for binding to the peptide (*data not shown*). A related approach, consisting of adding an excess of synthetic paxillin peptide to MBP-E6 expressions during cell lysis was also unsuccessful (*data not shown*). Finally, co-expression within the same bacterial cell of GST-paxillin peptide and untagged E6, both constructs being expressed by plasmids bearing different replication origins [30], led to co-purification of BPV1 E6 and GST-paxillin peptide, albeit at levels too modest for biophysical and structural studies (*data not shown*).

We therefore tried to stabilize BPV1 E6 by expressing the paxillin LXXLL motif and the E6 protein "*in cis*" on the same fusion construct. Two alternative constructs were generated. The first fusion protein (MBP-tevs-LXXLL-E6) was constructed as follows: [6HisMBP]-[TEV site]-[paxillin peptide]-[E6] (Fig. 3, left panel). The MBP-tevs-LXXLL-E6 fusion protein was expressed and purified exactly as described above for MBP-E6. The affinity purification on amylose resin raised amounts of soluble MBP-tevs-LXXLL-E6 fusion equivalent to the amounts of soluble MBP-E6 previously obtained. However, the ultracentrifugation step indicated that the MBP-tevs-LXXLL-E6 preparations contained very little amounts of soluble aggregates as compared to MBP-E6 preparations (Fig. 4A, lanes 2

and 3). TEV digestion of the MBP-tevs-LXXLL-E6 fusion was unusually slow (up to 36 hours), suggesting that the conformation of the LXXLL-E6 construct hindered accessibility to the proteolytic site. The identity of the LXXLL-E6 band (Fig. 4A, lanes 3) was confirmed by MALDI-TOF analysis (Table S2, Supplementary data). By contrast to what was previously observed for the MBP-E6 construct, the LXXLL-E6 moiety released by TEV proteolysis did not undergo precipitation neither during the proteolysis reaction nor during the subsequent concentration step. Gel filtration raised large amounts of monomeric LXXLL-E6 (Fig. 4B). In addition LXXLL-E6 samples could be concentrated up to 300 μM without undergoing any detectable aggregation.

To extend this approach, a second fusion protein (MBP-LXXLL-tevs-E6) was constructed as follows: [6HisMBP]-[paxillin peptide]- [TEV site]-[E6] (Fig. 3 right panel). This fusion protein was highly expressed as a soluble form. Like for the MBP-tevs-LXXLL-E6 construct, affinity purified MBP-LXXLL-tevs-E6 contained only minor amounts of aggregated material when submitted to overnight ultracentrifugation (Fig. 4C, lanes 2 and 3). The MBP-LXXLL-tevs-E6 preparation was further processed as the MBP-tevs-LXXLL-E6 construct, with the following modification: TEV digestion was performed in the presence of a four-fold stoichiometric excess of synthetic LXXLL paxillin peptide, to allow the released E6 moieties to be displaced from the MBP-fused LXXLL peptides to the small synthetic peptides. The E6/peptide complex was then separated from MBP-paxillin peptide and TEV by gel-filtration, eluting at the volume expected for a monomeric complex (Fig. 4D). The gel-filtrated sample could be concentrated up to 300 μM just like the covalent LXXLL-E6 fusion described before. To further verify the integrity of the E6 and LXXLL peptide in the non-covalent complex, we performed electrospray mass spectrometry analysis in both denaturing and native conditions. In denaturing conditions we identified two species of 1724.8 and 16051.4 Da, corresponding respectively to the LXXLL peptide (predicted molecular weight 1725.8 Da) and BPV1 E6 (predicted molecular weight 16051.6 Da). When the analysis was performed in native conditions we obtained a single species of 17904.3 Da, which corresponds to the sum of the molecular weights of the LXXLL peptide, BPV1 E6 and the two zinc atoms (predicted molecular weight of the complex: 17911.2 Da) (Table 2).

We also designed a GST-LXXLL-tevs-E6 construct that we attempted to purify in a similar manner. However, the solubility of this construct was poor as compared to that of the MBP-LXXLL-tevs-E6 (*data not shown*). This suggests that MBP acts in synergy with the LXXLL peptide to increase the soluble amounts of BPV1 E6 protein during biosynthesis in bacteria.

NMR analysis of BPV1 E6/paxillin complexes

The expression of both the MBP-tevs-LXXLL-E6 and MBP-LXXLL-tevs-E6 fusion constructs were reproduced in appropriate minimal M9 media, giving rise to ^{15}N-labeled samples that were subsequently purified according to the protocols described above and concentrated up to 300 μM. The ^{1}H, ^{15}N-HSQC spectrum of the covalent LXXLL-E6 fusion complex (Fig. 5A) displays and the chemical shift dispersion typical for a folded protein. Approximately 147 amide peaks can be counted in the HSQC spectrum. This number is close to that predicted from the sequence of the LXXLL-E6 construct (consisting of 155 amino acids of which 6 are proline residues). The ^{1}H, ^{15}N-HSQC spectrum of the non-covalent complex of BPV1 E6 with non-labelled synthetic LXXLL peptide (E6/LXXLL) is also typical for a folded protein (Fig. 5B). The number of amide peaks counted in the spectrum is 130, which is close to the number of amide resonances in the ^{15}N labelled BPV1 E6 counterpart (140 residues of which 6 are prolines) of the complex. Interestingly, signal line-width in the latter spectrum appears more homogeneous when compared to that of the first spectrum. This suggests that the covalent peptide fusion to the N-teminus of BPV1 E6 might introduce some constraints to complex dynamics and stability. Superimposition of the

two spectra (Fig. 5C) indicates that the majority of amide signals in the better dispersed regions of the spectrum have very similar chemical shift values, suggesting that the conformations of E6 in the covalent and non covalent complexes are very similar.

DISCUSSION

Papillomavirus E6 proteins are notoriously difficult to express and purify. Unfused E6 proteins form insoluble aggregates upon bacterial overexpression. Usual purification tags such as the 6His-tag or the Gluthatione-S-Transferase (GST) generate no or little solubilization of E6. By contrast, fusion to the Maltose-Binding Protein (MBP) is particularly efficient in solubilising E6 proteins. However, a large fraction of the soluble MBP-E6 proteins may exist as soluble aggregates. We have previously developed various assays to detect these soluble aggregates and used ultracentrifugation approaches to eliminate them [23, 26, 31]. However, monomeric E6 proteins obtained by these procedures generally remain prone to self-association at high concentration, hindering their biophysical and structural analysis [27].

The BPV1 E6 protein studied in this work generally behaves in a manner comparable to that of the HPV E6 proteins studied to date. GST-BPV1 E6 fusions are poorly soluble (*data not shown*) whereas MBP-BPV1 E6 fusions are soluble but largely under the form of soluble aggregates and/or small size oligomers [27]. These aggregates can be cleared by ultracentrifugation, allowing one to isolate, after removal of MBP by proteolysis, only diluted amounts of monomeric E6 that cannot stand further concentration required for biophysical analysis.

To turn around this problem, we have searched for means of improving the solubility of BPV1 E6. Previous studies had shown that a construct of BPV1 E6 fused to the C-terminus of the same paxillin-derived LXXLL peptide used in the present study was inactivated *in vivo* for biological functions requiring interactions with partners possessing a LXXLL E6-binding motif. This suggested that the fused peptide was efficiently blocking the LXXLL-binding site of BPV1 E6 [18]. SPR measurements showed that BPV1 E6 binds to the paxillin LXXLL peptide with high affinity and slow dissociation rate kinetics. Given its stability, we reasoned that the protein/peptide complex may persists through the purification steps and ensure the stabilization of the E6 protein at high concentrations. Indeed the conserved leucine residues of the LXXLL motif are likely to interact with a hydrophobic patch exposed on the surface of E6, which may promote aggregation of the unbound protein. We tested several alternative strategies. Approaches based on co-expression and co-purification of BPV1 E6 and the cognate peptide were unsuccessful or yielded insufficient amounts of complexes. By contrast, approaches based on *"in cis"* expression of MBP, peptide and E6, in the form of triple fusion, turned out to be remarkably successful. Two alternative strategies were used to express and purify samples of BPV1 E6/paxillin complexes. One strategy, based on the MBP-tevs-LXXLL-E6 construct, led to a covalent (intra-molecular) LXXLL-E6 fusion contruct, which proved soluble at high concentrations and amenable to biophysical analysis by NMR. The second strategy, based on the MBP-LXXLL-tevs-E6 construct, allowed us to obtain samples of a non-covalent (inter-molecular) E6/LXXLL complex, which was equally soluble at high concentrations and also amenable to NMR studies. This latter approach has the advantage of allowing for subunit specific labelling, thus reducing the complexity of NMR spectra. The success of the triple fusion approach indicates that both peptidic ligand and MBP tag are required during biosynthesis for the solubilization of BPV1 E6.

In our previous work, we have shown that mutagenesis of non-conserved cysteines to the isosteric non-oxidable residue serine notably increased the solubility of recombinant HPV16

E6 protein [23]. We have tried to reproduce this approach by producing various mutants of BPV1 E6 in which all or some of the non-conserved cysteins were mutated to serines. None of these mutants displayed improved solubility as compared to the wild-type BPV1 E6 protein (*data not shown*). In parallel, we also generated constructs where the LXXLL motif of E6AP was fused to the N-terminus of HPV16 E6 (MBP-tevs-LXXLL-HPV16E6) *via* linkers of different sizes. None of these constructs led to a significant improvement of HPV16 E6 solubility (*data not shown*), a result that might be related to the weaker affinity of this interaction.

Therefore, a similar problem (aggregation of protein upon expression and concentration) occurring for two proteins belonging to the same family (papillomavirus E6 proteins) had to be treated by two different approaches (cystein mutagenesis and stabilisation by peptidic ligand for HPV E6 and BPV E6 respectively). This shows that, even for closely related proteins, there are no definite standard recipes for enhancing sample solubility.

Both BPV1 E6 and HPV16 E6 both bind to peptides possessing the consensus LXXLL [18]. We have found here that a LXXLL peptide from paxillin forms with BPV1 E6 a tight complex (K_D=35 nM). On the other hand, the interactions of BPV1 E6 and other "high-risk" genital HPV E6 proteins with the LXXLL peptide of E6AP have 30-to-1000-fold lower equilibrium binding affinity constants (K_D=1-30 μM) [29]. These observations on the affinities of E6 proteins towards different LXXLL containing peptides may have functional implications. The E6/paxillin complex may exist permanently in BPV-infected cells, whereas the E6/E6AP complex may be more transient, allowing for the formation of alternative complexes.

To summarize, we have shown that, like Human Papillomavirus E6, the E6 protein of Bovine Papillomavirus is prone to aggregation, hindering its biophysical analysis. However, the protein becomes highly soluble and amenable to biophysical analysis provided that it is complexed, covalently or non-covalently, with its main peptidic ligand. This finding opens the way to novel structural and functional studies of this potent mammalian oncoprotein. In addition, the original approaches that we have developed in this work might be applicable to other cases of proteins that are only stable and/or soluble in the presence of a peptidic partner.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

Abbreviations

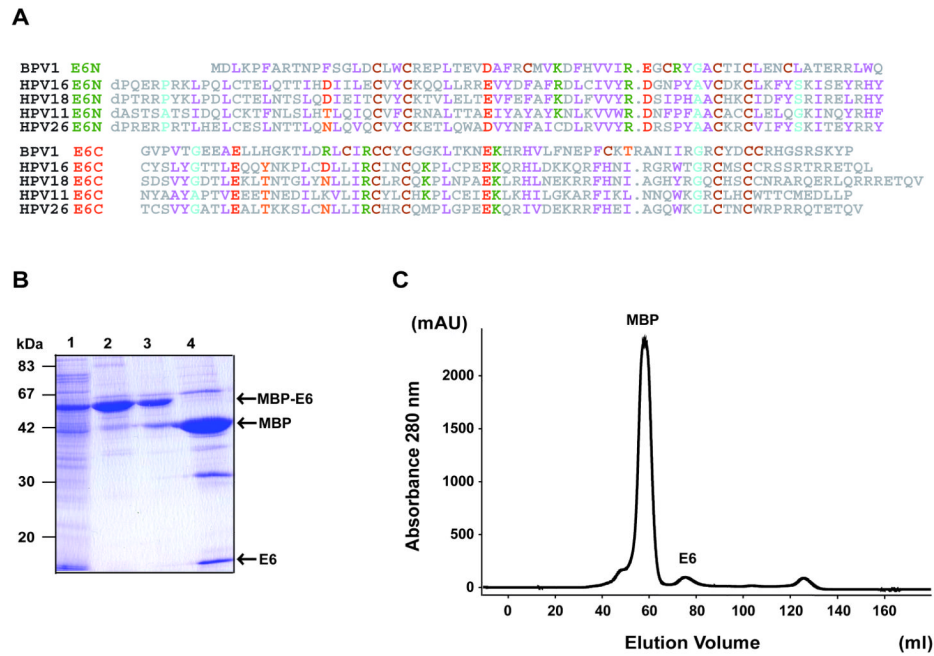|	
MBP	maltose binding protein
HPV	human papillomavirus
BPV	bovine papillomavirus
TEV	tobacco etch virus
SPR	surface plasmon resonance

RU   response units

NMR   nuclear magnetic resonance spectroscopy.

REFERENCES

[1]. de Villiers EM, Fauquet C, Broker TR, Bernard HU, zur Hausen H. Classification of papillomaviruses. Virology. 2004; 324:17–27. [PubMed: 15183049]

[2]. zur Hausen H. Papillomaviruses in human cancers. Proc. Assoc. Am. Physicians. 1999; 111:581–587. [PubMed: 10591087]

[3]. Bosch FX, Manos MM, Munoz N, Sherman M, Jansen AM, Peto J, Schiffman MH, Moreno V, Kurman R, Shah KV. Prevalence of human papillomavirus in cervical cancer: a worldwide perspective. International biological study on cervical cancer (IBSCC) Study Group. J. Natl. Cancer Inst. 1995; 87:796–802. [PubMed: 7791229]

[4]. Nasir L, Campo MS. Bovine papillomaviruses: their role in the aetiology of cutaneous tumours of bovids and equids. Vet. Dermatol. 2008; 19:243–254. [PubMed: 18927950]

[5]. Scheffner M, Huibregtse JM, Vierstra RD, Howley PM. The HPV-16 E6 and E6-AP complex functions as a ubiquitin-protein ligase in the ubiquitination of p53. Cell. 1993; 75:495–505. [PubMed: 8221889]

[6]. Werness BA, Levine AJ, Howley PM. Association of human papillomavirus types 16 and 18 E6 proteins with p53. Science. 1990; 248:76–79. [PubMed: 2157286]

[7]. Chakrabarti O, Krishna S. Molecular interactions of 'high risk' human papillomaviruses E6 and E7 oncoproteins: implications for tumour progression. J. Biosci. 2003; 28:337–348. [PubMed: 12734411]

[8]. Thomas M, Dasgupta J, Zhang Y, Chen X, Banks L. Analysis of specificity determinants in the interactions of different HPV E6 proteins with their PDZ domain-containing substrates. Virology. 2008; 376:371–378. [PubMed: 18452965]

[9]. Huibregtse JM, Scheffner M, Howley PM. Localization of the E6-AP regions that direct human papillomavirus E6 binding, association with p53, and ubiquitination of associated proteins. Mol. Cell. Biol. 1993; 13:4918–4927. [PubMed: 8393140]

[10]. Chen JJ, Hong Y, Rustamzadeh E, Baleja JD, Androphy EJ. Identification of an alpha helical motif sufficient for association with papillomavirus E6. J. Biol. Chem. 1998; 273:13537–13544. [PubMed: 9593689]

[11]. Ronco LV, Karpova AY, Vidal M, Howley PM. Human papillomavirus 16 E6 oncoprotein binds to interferon regulatory factor-3 and inhibits its transcriptional activity. Genes Dev. 1998; 12:2061–2072. [PubMed: 9649509]

[12]. Patel D, Huang SM, Baglia LA, McCance DJ. The E6 protein of human papillomavirus type 16 binds to and inhibits co-activation by CBP and p300. EMBO J. 1999; 18:5061–5072. [PubMed: 10487758]

[13]. Zimmermann H, Degenkolbe R, Bernard HU, O'Connor MJ. The human papillomavirus type 16 E6 oncoprotein can down-regulate p53 activity by targeting the transcriptional coactivator CBP/p300. J. Virol. 1999; 73:6209–6219. [PubMed: 10400710]

[14]. Elston RC, Napthine S, Doorbar J. The identification of a conserved binding motif within human papillomavirus type 16 E6 binding peptides, E6AP and E6BP. J. Gen. Virol. 1998; 79:371–374. [PubMed: 9472622]

[15]. Das K, Bohl J, Vande Pol SB. Identification of a second transforming function in bovine papillomavirus type1 E6 and the role of E6 interactions with paxillin, E6BP, and E6AP. J. Virol. 2000; 74:812–816. [PubMed: 10623743]

[16]. Tong X, Boll W, Kirchhausen T, Howley PM. Interaction of the bovine papillomavirus E6 protein with the clathrin adaptor complex AP-1. J. Virol. 1998; 72:476–482. [PubMed: 9420248]

[17]. Vande Pol SB, Brown MC, Turner CE. Association of Bovine Papillomavirus Type 1 E6 oncoprotein with the focal adhesion protein paxillin through a conserved protein interaction motif. Oncogene. 1998; 16:43–52. [PubMed: 9467941]

[18]. Bohl J, Das K, Dasgupta B, Vande Pol SB. Competitive binding to a charged leucine motif represses transformation by a papillomavirus E6 oncoprotein. Virology. 2000; 271:163–170. [PubMed: 10814581]

[19]. Lipari F, McGibbon GA, Wardrop E, Cordingley MG. Purification and biophysical characterization of a minimal functional domain and of an N-terminal Zn^{2+}-binding fragment from the human papillomavirus type 16 E6 protein. Biochemistry. 2001; 40:1196–1204. [PubMed: 11170444]

[20]. Nominé Y, Charbonnier S, Ristriani T, Stier G, Masson M, Cavusoglu N, Van Dorsselaer A, Weiss E, Kieffer B, Travé G. Domain substructure of HPV E6 protein: biophysical characterization of E6 C-terminal DNA-binding domain. Biochemistry. 2003; 42:4909–4917. [PubMed: 12718532]

[21]. Nomine Y, Masson M, Charbonnier S, Zanier K, Ristriani T, Deryckere F, Sibler AP, Desplancq D, Atkinson RA, Weiss E, Orfanoudakis G, Kieffer B, Trave G. Structural and functional analysis of E6 oncoprotein: insights in the molecular pathways of human papillomavirus-mediated pathogenesis. Mol. Cell. 2006; 21:665–678. [PubMed: 16507364]

[22]. Daniels PR, Sanders CM, Coulson P, Maitland NJ. Molecular analysis of the interaction between HPV type 16 E6 and human E6-associated protein. FEBS Letters. 1997; 416:6–10. [PubMed: 9369221]

[23]. Nominé Y, Ristriani T, Laurent C, Lefèvre JF, Weiss E, Travé G. A strategy for optimizing the monodispersity of fusion proteins: application to purification of recombinant HPV E6 oncoprotein. Protein Eng. 2001; 14:297–305. [PubMed: 11391022]

[24]. Degenkolbe R, Gilligan P, Gupta S, Bernard HU. Chelating agents stabilize the monomeric state of the zinc binding huma, papillomavirus 16 E6 oncoprotein. Biochemistry. 2003; 42:3868–3873. [PubMed: 12667077]

[25]. Liu Y, Cherry JJ, Dineen JV, Androphy EJ, Baleja JD. Determinants of stability for the E6 protein of papillomavirus type 16. J. Mol. Biol. 2009; 386:1123–1137. [PubMed: 19244625]

[26]. Zanier K, Nomine Y, Charbonnier S, Ruhlmann C, Schultz P, Schweizer J, Trave G. Formation of well-defined soluble aggregates upon fusion to MBP is a generic property of E6 proteins from various human papillomavirus species. Protein Expr. Purif. 2007; 51:59–70. [PubMed: 17055740]

[27]. Zanier K, Ruhlmann C, Melin F, Masson M, Sidi A. Ould M'hamed Ould, Bernard X, Fischer B, Brino L, Ristriani T, Rybin V, Baltzinger M, Vande Pol S, Hellwig P, Schultz P, Trave G. E6 proteins from diverse papillomaviruses self-associate both in vitro and in vivo. J. Mol. Biol. 2010; 396:90–104. [PubMed: 19917295]

[28]. Fournane S, Charbonnier S, Chapelle A, Kieffer B, Orfanoudakis G, Trave G, Masson M, Nomine Y. Surface plasmon resonance analysis of the binding of high-risk mucosal HPV E6 oncoproteins to the PDZ1 domain of the tight junction protein MAGI-1. J. Mol. Recognit. 2010

[29]. Zanier K, Charbonnier S, Baltzinger M, Nomine Y, Altschuh D, Trave G. Kinetic analysis of the interactions of Human Papillomavirus E6 oncoproteins with the ubiquitin ligase E6AP using Surface Plasmon Resonance. J. Mol. Biol. 2005; 349:401–412. [PubMed: 15890204]

[30]. Romier C, Jelloul M. Ben, Albeck S, Buchwald G, Busso D, Celie PH, Christodoulou E, De Marco V, van Gerwen S, Knipscheer P, Lebbink JH, Notenboom V, Poterszman A, Rochel N, Cohen SX, Unger T, Sussman JL, Moras D, Sixma TK, Perrakis A. Co-expression of protein complexes in prokaryotic and eukaryotic hosts: experimental procedures, database tracking and case studies. Acta Crystallogr. D Biol. Crystallogr. 2006; 62:1232–1242. [PubMed: 17001100]

[31]. Nominé Y, Ristriani T, Laurent C, Lefèvre JF, Weiss E, Travé G. Formation of soluble inclusion bodies by HPV E6 oncoprotein fused to Maltose-binding protein. Protein Expr. Purif. 2001; 23:22–32. [PubMed: 11570842]

A

```
BPV1  E6N           MDLKPFARTNPFSGLDCLWCREPLTEVDAFRCMVKDFHVVIR.EGCRYGACTICLENCLATERRLWQ
HPV16 E6N  dPQERPRKLPQLCTELQTTIHDIILECVYCKQQLLRREVYDFAFRDLCIVYR.DGNPYAVCDKCLKFYSKISEYRHY
HPV18 E6N  dPTRRPYKLPDLCTELNTSLQDIEITCVYCKTVLELTEVFEFAFKDLFVVYR.DSIPHAACHKCIDFYSRIRELRHY
HPV11 E6N  dASTSATSIDQLCKTFNLSLHTLQIQCVFCRNALTTAEIYAYAYKNLKVVWR.DNFPFAACACCLELQGKINQYRHF
HPV26 E6N  dPRERPRTLHELCESLNTTLQNLQVQCVYCKETLQWADVYNFAICDLRVVYR.DRSPYAACKRCVIFYSKITEYRRY

BPV1  E6C  GVPVTGEEAELLHGKTLDRLCIRCCYCGGKLTKNEKHRHVLFNEPFCKTRANIIRGRCYDCCRHGSRSKYP
HPV16 E6C  CYSLYGTTLEQQYNKPLCDLLIRCINCQKPLCPEEKQRHLDKKQRFHNI.RGRWTGRCMSCCRSSRTRRETQL
HPV18 E6C  SDSVYGDTLEKLTNTGLYNLLIRCLRCQKPLNPAEKLRHLNEKRRFHNI.AGHYRGQCHSCCNRARQERLQRRRETQV
HPV11 E6C  NYAAYAPTVEEETNEDILKVLIRCYLCHKPLCEIEKLKHILGKARFIKL.NNQWKGRCLHCWTTCMEDLLP
HPV26 E6C  TCSVYGATLEALTKKSLCNLLIRCHRCQMPLGPEEKQRIVDEKRRFHEI.AGQWKGLCTNCWRPRRQTETQV
```

B

C

Fig. 1.
(**A**) Alignment of E6 sequences (upper half E6-N and lower half E6-C) from the BPV1 and several HPV species. Residues are numbered as in the shorter transcript of HPV 16 E6 (151 residues). Positions exhibiting conserved physicochemical characteristics are colored as follows: magenta, hydrophobic (W, F, Y, L, I, V, M); green, basic (K, R, H); red, acidic (E, D); orange, polar (Q, N, T, S); brown, cystein (C); cyan, small (G, A, P). (**B**) 15% SDS-PAGE gel showing the different purifications steps of BPV1 E6. Affinity-purified MBP-E6 was cleared from soluble aggregates by ultracentrifugation and digested using TEV protease. Subsequently, the sample was concentrated and injected on a HighLoad Superdex75 16/60 column. Lane 1: total cell lysate; lane 2: eluate from amylose resin before overnight ultracentrifugation; lane 3: eluate from amylose resin after overnight ultracentrifugation; lane 4: TEV protease digest. (**C**) Gel filtration after separation from the MBP by TEV digestion. Peaks corresponding to the MBP and E6 proteins are indicated. Note that very small amounts of monomeric E6 are recovered after gel filtration.
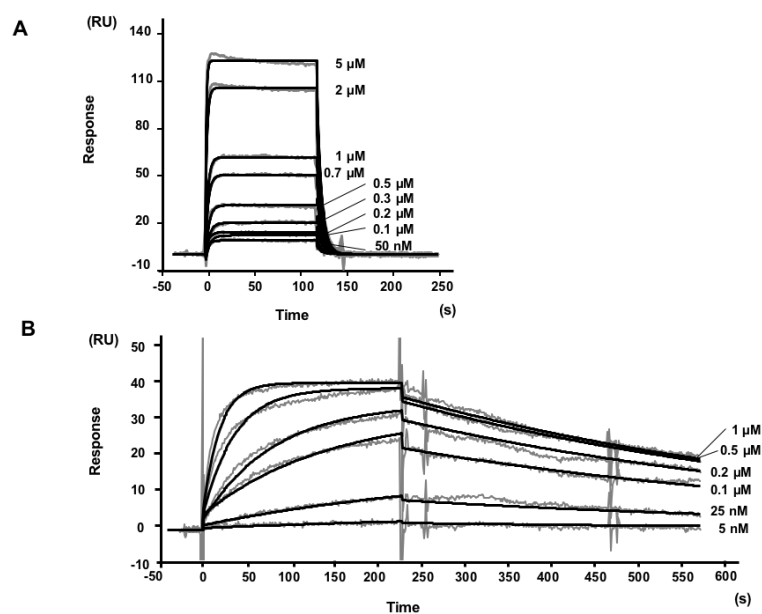
Fig. 2.
Biacore datasets used for kinetic analysis of the interactions of MBP-BPV1 E6 with peptides containing (**A**) the LXXLL motif of E6AP and (**B**) the N-terminal LXXLL motif of paxillin. Gray lines represent SPR sensorgrams obtained by injecting analyte MBP-E6 on surfaces capturing 60 (**A**) and 30 (**B**) RUs of GST-peptide ligand. The concentration of analyte is indicated on the right hand side of each sensorgram. Black lines are the global fits of the sensorgrams to a simple bimolecular Langmuir model. The ratio of the globally fit Rmax ($Rmax_{global}$) over the theoretical Rmax ($Rmax_{theoretical}$ -estimated from the amount of fusion peptide captured-) $Rmax_{global}/Rmax_{theoretical}$ was 0.9 and 0.6 for the GST-paxillin and GST-E6AP ligands. Kinetic parameters are reported in Table 1.
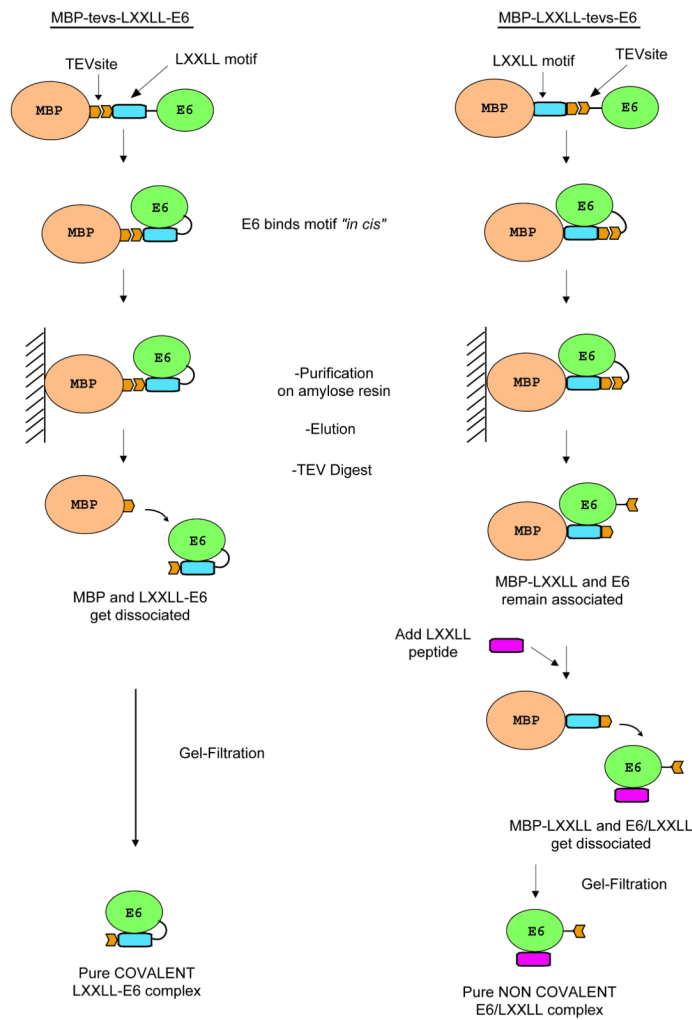
Fig. 3.
Purification scheme for the MBP-tevs-LXXLL-E6 construct (left) and the MBP-LXXLL-tevs-E6 construct (right). Both constructs stabilise the BPV1 E6 protein during expression by allowing it to form a complex *in cis* with its cognate LXXLL motif peptide. TEV digestion is carried out for both constructs after the first affinity purification step. In the case of the MBP-tevs-LXXLL-E6 construct, this leads to the isolation of pure covalent LXXLL-E6 complex after gel filtration. On the other hand, TEV proteolysis of the MBP-LXXLL-tevs-E6 construct is performed in the presence of a 4-fold excess of LXXLL synthetic peptide, which binds to the E6 moiety, allowing the displacement of the MBP-LXXLL fusion and leading to isolation of the non-covalent E6/LXXLL complex after gel filtration.
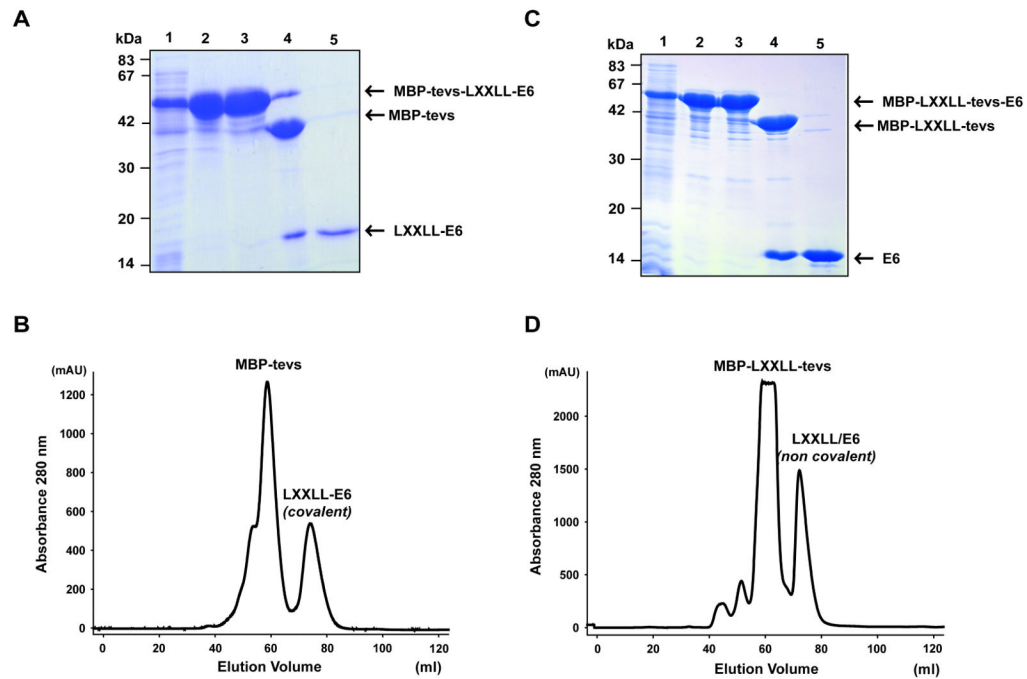
Fig. 4.
Purification of the construct of MBP-tevs-LXXLL-E6 (**A-B**) and MBP -LXXLL-tevs-E6
(**C-D**) constructs. (**A** and **C**) 15% SDS-PAGE gels for the two purifications. Lane 1: total
cell lysate; lane 2: eluate from amylose resin before overnight ultracentrifugation; lane 3:
eluate from amylose resin before overnight ultracentrifugation; lane 4: TEV protease digest;
lane 5: pooled gel-filtration fractions at the elution volume expected for the E6/paxillin
complex. (**B** and **D**) Gel filtration profiles of LXXLL-E6 and E6/LXXLL samples after
separation from the MBP. Peaks corresponding to MBP-tevs or MBP-LXXLL-tevs and to
E6/paxillin complexes are indicated. See also legend of Fig. 1.
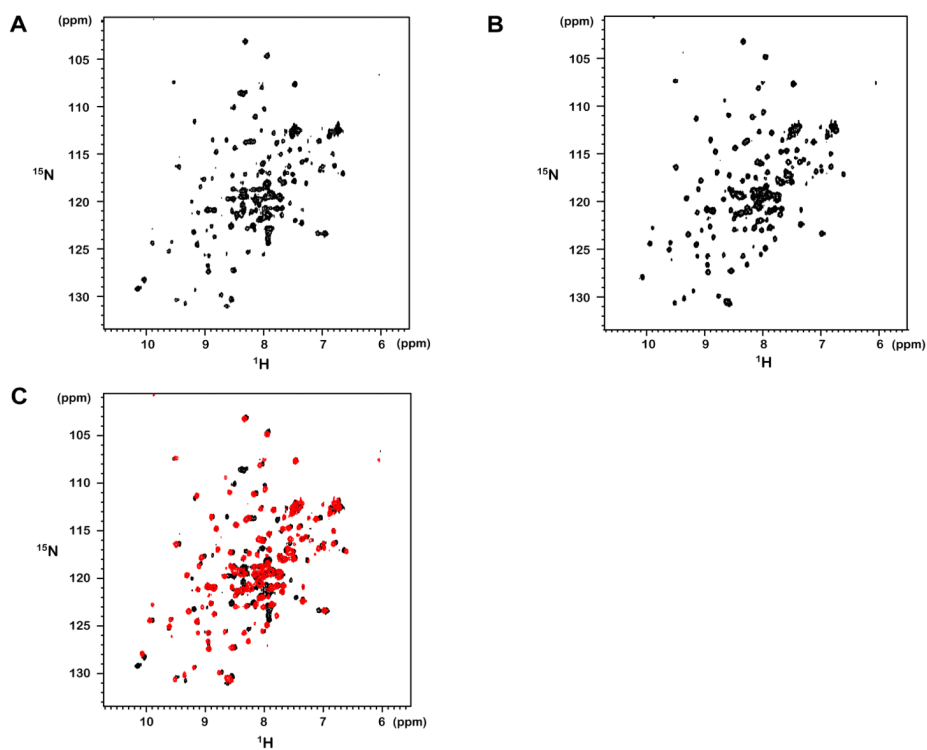
Fig. 5.
^{1}H-^{15}N HSQC spectra of (**A**) ^{15}N labelled LXXLL-E6 fusion (covalent) and of (**B**) ^{15}N labelled E6 bound to unlabelled LXXLL peptide (non-covalent). Spectra were acquired with 32 scans at 296 K. (**C**) Overlay of spectra shown in **A** (black) and in **B** (red).

Table 1

Kinetic parameters derived from Biacore sensorgrams shown in Figure 2.

	k_{on} (M^{-1} s^{-1})	k_{off} (s^{-1})	k_{off}/k_{on} (M)
BPV1 E6/E6AP	1.5×10^5	0.2	1.3×10^{-6}
BPV1 E6/paxillin	5.0×10^4	1.7×10^{-3}	3.5×10^{-8}

Table 2

Molecular weights of the non-covalent BPV1 E6/paxillin complex and its components as derived from both denaturing and native mass spectrometry measurements.

	BPV1 E6 ($-Zn^{2+}$) (Da)	paxillin (Da)	BPV1 E6 ($+Zn^{2+}$) / LXXLL (Da)
predicted	16051.6	1725.8	17903.2
denaturing mass-spec	16051.4	1724.8	-
native mass-spec	-	-	17904.3