Structure of high-risk papillomavirus type 31 E6 oncogenic protein and characterization of E6/E6AP/p53 complex formation

Marcel Chris Conrady, Irina Suarez, Gergö Gogl, Desiree Isabella Frecot, Anna Bonhoure, Camille Kostmann, Alexandra Cousido-Siah, Andreá Mitschler, Jiawen Lim, Murielle Masson, et al.

JVI Accepted Manuscript Posted Online 28 October 2020
J. Virol. doi:10.1128/JVI.00730-20

Title: Structure of high-risk papillomavirus type 31 E6 oncogenic protein and characterization of E6/E6AP/p53 complex formation

Authors: Marcel Chris Conrady[1], Irina Suarez[2], Gergö Gogl[2], Desiree Isabella Frecot[1], Anna Bonhoure[2], Camille Kostmann[2], Alexandra Cousido-Siah[2], Andre´ Mitschler[2], JiaWen Lim[1], Murielle Masson[3], Thomas Iftner[1], Frank Stubenrauch[1], Gilles Trave[2]*, Claudia Simon[1]*

Running Title (54 characters)**:** HPV 31E6 structure and complex characterization

Affiliations:

[1]Institute of Medical Virology, Medical Faculty, Eberhard-Karls-University, Tuebingen, Germany

[2]Equipe Labellisée Ligue 2015, Department of Integrative Biology, Institut de Génétique et de Biologie Moléculaire et Cellulaire, CNRS, INSERM, UdS, 1 rue Laurent Fries, 67404 Illkirch CEDEX (France)

[3]UMR 7242 Biotechnologie et signalisation cellulaire, CNRS, UdS, ESBS, 300 Boulevard S. Brant, 67412 Illkirch, France

*__Address correspondence to__:

Gilles Trave, traveg@igbmc.fr, or Claudia Simon, claudia.simon@med.uni-tuebingen.de

Author contributions:

experimental design and data analysis/interpretation– MC, GG, AB, ACS, JL, FS, TI, GT, CS; Fluorescence Polarization – MC, GG; crystallography – IS, MC, ACS, AM, GG; MST – DIF, CS; GPCA – MC, MM

Abbreviations

E6AP	E3 ubiquitin ligase E6-associated protein
GPCA	*Gaussia princeps* protein complementation assay
His_6	Hexa-histidine
HPV	Human papillomavirus
MBP	Maltose binding protein
OD_{600}	Optical density at 600 nm
TCEP	Tris(2-carboxyethyl)phosphine

Abstract (173/250 words)

The degradation of p53 is a hallmark of “high-risk” HPV types of the alpha genus and HPV-related carcinogenicity. The oncoprotein E6 forms a ternary complex with the E3 ubiquitin ligase E6-associated protein (E6AP) and tumor suppressor protein p53 targeting p53 for ubiquitination. The extent of p53 degradation by different E6 proteins varies greatly, even for the closely-related HPV 16 and 31. 16E6 and 31E6 display high sequence identity (~67 %). We report here for the first time the structure of HPV31 E6 bound to the LxxLL motif of E6AP. 16E6 and 31E6 are structurally very similar in agreement with the high sequence conservation. Both E6 proteins bind E6AP and degrade p53. However, the binding affinities of 31E6 to E6AP and p53, respectively, are reduced 2-fold and 5.4-fold as compared to 16E6. The affinity of E6-E6AP-p53 ternary complex formation parallels the efficacy of the subsequent reaction, namely degradation of p53. Therefore, closely-related E6 proteins

addressing the same cellular targets may still diverge in their binding efficiencies, possibly explaining their different phenotypic or pathologic impact.

Importance (150/150 words)

Variations of carcinogenicity of Human papillomaviruses are related to variations of the E6 and E7 interactome. While different HPV species and genera are known to target distinct host proteins, the fine differences between E6 and E7 of closely-related HPV types, supposed to target the same cellular protein pools, remain to be addressed. We compare the oncogenic E6 proteins of the closely related “high-risk” HPV types 31 and 16 with regard to their structure and their efficiency of ternary complex formation with their cellular targets p53 and E6AP, which results in p53 degradation. We solved the crystal structure of 31E6 bound to the E6AP LxxLL-motif. 16E6 and 31E6 structures are highly similar but a few sequence variations lead to different protein contacts within the ternary complex and, as quantified here, an overall lower binding affinity of 31E6 compared with 16E6. These results align with the observed lower p53-degradation potential of 31E6.

Introduction

Human papillomaviruses (HPV) comprise over 200 types. In accordance to their sequence alignment of the major capsid protein L1 they are divided in 5 genera – alpha, beta, gamma, mu and nu. HPVs infect skin squamous epithelial cells (alpha and beta genus) and mucosal epithelial cells (alpha genus) in humans. An infection can either be asymptomatic or result in benign tumors or cancer. Cancer development occurs only in the rare cases of persistent infection and failure of viral clearance. However, so called “high-risk” HPVs are classified as most carcinogenic by the International Agency for Research on Cancer (IARC) (1, 2) because they are highly associated with cancer of cervix (100.0 %), oropharynx (30.8 %), vulvar (24.9 %), vagina (78.0 %), penis (50.0 %), and anus (88.0 %) (3). The carcinogenic potential varies between the different HPV types. All “high-risk” types belong to the alpha genus, but to four different species: alpha-5 (HPV 51), alpha-6 (HPV 56), alpha-7 (HPV 18, 39, 45, 59 and 68) and alpha-9 (HPV 16, 31, 33, 35, 52, 58). HPV 16 is associated with ~ 50 %, HPV 18 with 20 % whereas the HPV 31, 33, 35, 39, 45, 51, 52, 56, 58, 59 and 68 together are associated with ~ 30 % of cervical cancer (4-6).

Two viral oncoproteins E6 and E7, which are always expressed in HPV associated cancers (7), are responsible for the immortalization of human keratinocytes, the target cells of HPV (8). Current models suggest that protein-protein-interactions of both viral oncoproteins with their cellular targets contribute to carcinogenicity. Most E7 proteins share the capacity to bind to retinoblastoma (Rb) family members as well as phosphatases PTPN14 and PTPN21 (9-13), independently of the carcinogenic risk level of HPVs they belong to. In contrast, E6 proteins do not share any universal target cell protein conserved for all HPVs. However, binding to the E3 ubiquitin ligase

E6-associated protein (E6AP), the tumor suppressor protein p53 and PDZ domain proteins are key interactions that are specifically displayed only for E6 proteins of "high-risk" alpha mucosal HPVs. The simultaneous recruitment of E6AP and p53 by E6 results in the degradation of p53 (14, 15).

E6 is highly conserved among papillomaviruses and consists of two zinc-binding domains (figure 1). An E6-based phylogenic classification depicts similar relationships for E6 proteins within the alpha genus as the common L1-based classification. Within the alpha-9 group HPV 16, 35 and 31 are the closest related HPVs. Generally, E6 proteins bind to accessible LxxLL-motifs of various cellular target proteins. The LxxLL-binding profile of different E6 proteins varies within the alpha genus (16).

HPV 16 is the HPV type with highest carcinogenic risk. It has been shown previously, that 16E6 forms a ternary complex with the E6AP-LxxLL motif and the p53 core domain (17). Once formed, the E6/E6AP/p53 complex mediates ubiquitination and subsequent proteasomal degradation of p53 (18). Among all cellular interaction partners of E6, this ternary complex is the best characterized interaction in structure and function for 16E6. Studies on binding parameters of the ternary complex depicted binding affinities of the 16E6-LxxLL(E6AP) dimer to p53 in a µM-range (17) and defined crucial amino acids for complex formation and subsequent p53 degradation (table 2 and 3). Figure 1 shows a sequence alignment of E6 proteins of the alpha-9 genus and HPV 18 (alpha-7) as the second most prevalent "high-risk" type. Amino acids of 16E6, which significantly contribute to LxxLL(E6AP) or p53 binding are not completely conserved among the alpha-9 HPVs. In table 2 and 3 these amino acids are compared regarding their conservation between 16E6 and 31E6. Despite the high sequence homology of 16E6 and 31E6 of 67 % identity

(ClustalO (19)) both binding sites show sequence variations, which can potentially interfere with target binding, the formation of the ternary complex and subsequent p53 degradation. Additionally, quantitative assays using a luciferase-p53 fusion protein as substrate for E6 proteins revealed that HPV 16 E6 is more active in initiating p53 degradation than HPV 31 E6 (20). These data suggest that besides qualitative also quantitative differences among E6 and E7 protein interactions may explain the different carcinogenic behavior.

HPV 16 and 31 are closely related, belong to the same genus and species and show consistence in phylogeny and pathology. However, it is not completely understood, why HPV 16 is far more carcinogenic than other “high-risk” alpha-9 HPV types. In this work, the crystal structure of 31E6 was solved by X-ray crystallography. The binding properties of 31E6 to p53 and LxxLL(E6AP) were analyzed quantitatively in comparison to 16E6. Both HPV types represent “high-risk” types of alpha-9 genus infecting the mucosal keratinocytes but reveal a different p53 degradation potential in cell-based assays (2, 6, 20). Presumably, the degradation of p53 is greatly related to carcinogenicity and variations in p53 degradation could be one factor in this process. Hence, our structural and quantitative analysis bridges sequence, structure and function together and further, suggests explanations regarding the different p53 degradation potential of two very closely related HPV E6 proteins both targeting the same cellular targets E6AP and p53.

Results

Structural analysis of 31E6

Recombinant E6 proteins are notoriously prone to solubility issues (reviewed in (21)). To overcome this problem, we fused a crystallization-prone mutant of the bacterial maltose-binding protein (MBP) to the N-terminus of the HPV 31 E6 protein and the E6-binding LxxLL sequence of E6AP (figure 5) to the C-terminus of the HPV 31 E6 protein. The resulting MBP-31E6-LxxLL(E6AP) triple fusion protein was purified as a soluble monomer and yielded crystals that diffracted until 2.8 Å. The structure was solved by molecular replacement using the known structures of MBP and HPV 16 E6 as a template.

The overall domain organization and structure is very similar (figure 3A) to the published 16E6 structures. Two zinc ions are present in the structure, each coordinated by four cysteine residues of the highly conserved CxxC motifs. Notably, the amino acid R102 is conserved by sequence and structure. This residue is important, because it bridges the E6N and E6C domain via two hydrogen bonds to the backbone carbonyls of the E6N and contributes to LxxLL-motif binding in both E6 structures.

The structure of 16E6 was previously solved as a heterodimer 16E6/MBP-(LxxLL)E6AP complex (Protein Data Bank (PDB) 4GIZ, (22)) and as a heterotrimer (16E6/MBP-LxxLL(E6AP)/p53core (PDB 4XR8 (17)). An alignment of these structures with the heterodimeric structure of 31E6/LxxLL(E6AP) based on the LxxLL(E6AP)-peptide shows a mostly unperturbed E6N domain. Remarkably, the structure of the ternary complex 16E6/MBP-LxxLL(E6AP)/p53core contains two conformers in the asymmetric unit, trimer A and B, which differ in the position of the

E6C domain. 16E6C of trimer A superimposes well with E6C of the 16E6/MBP-(LxxLL)E6AP dimeric complex. In comparison, the E6C domain of trimer B is distorted. Notably, the crystal structure of 31E6-LxxLL(E6AP) only contains one conformer. Here, the 31E6C domain aligns well with the 16E6/MBP-LxxLL(E6AP)/p53core trimer B conformation. The differences of the E6C domain positions become obvious by comparing the root-mean-square deviation (RMSD) of the corresponding amino acid Cα positions in the different 16E6 protein structures related to the structural data of 31E6-LxxLL(E6AP) (figure 3B) as an indicator of protein backbone alignment. This shows that the E6C domain deviates up to 5.5 Å among the E6 proteins, whereas the E6N shows more similar RMSD values and, that 31E6-LxxLL(E6AP) aligns best with trimer B of 16E6/LxxLL(E6AP)/p53core. From now on the obtained structure of 31E6-LxxLL(E6AP) will be compared with 16E6 of the 16E6/LxxLL(E6AP)/p53core trimer B ternary complex as presented in PDB 4XR8 (16).

31E6 interacts specifically with E6AP and LxxLL(E6AP)-peptides

The interaction of E6 with the LxxLL-motif of the cellular E6AP is a requirement for the formation of the ternary complex E6/E6AP/p53 followed by proteasomal degradation of p53.

We examined the interaction of the oncoprotein E6 from HPV 31 with the host cell protein E6AP in comparison to HPV 16 E6 as previously published (23). The interaction was tested *in cellulo* using the qualitative *Gaussia princeps* protein complementation assay (GPCA) (24, 25), where proteins are expressed and assayed in a mammalian cellular environment. The GPCA signal of 31E6 (normalized luciferase ratio (NLR) = 55 ± 9) is decreased by 35 % compared to 16E6 (NLR = 84 ± 5) (figure 2G) clearly indicating, that 31E6 shows a reduced interaction with E6AP *in*

cellulo. We used 16E6 L50E as a negative control (NLR = 7 ± 1), which was previously shown not to interact with E6AP. The expression of all three tested E6 constructs was verified by western analysis (figure 2H).

Further, we used a fluorescence anisotropy assay to quantify the binding affinities of 16E6 and 31E6 to the LxxLL-motif of E6AP (figure 2). To eliminate the fluorophore-induced effects on the measured affinities, we used a fluorescein-labeled LxxLL(E6AP)-peptide as a tracer to monitor the binding of an unlabeled LxxLL(E6AP)-peptide. We observed a low-micromolar affinity between the unlabeled LxxLL-peptide of E6AP and 16E6 (K_d=7.8 +/- 0.4 µM) and a 2-fold weaker affinity with 31E6 (K_d=13.6 +/- 0.9 µM).

Even though the general architecture of 16E6 is maintained in 31E6, the structural conformation of R10, which is conserved by sequence in 16E6 and 31E6 (figure 3), is different. This is accompanied by a different orientation of the LxxLL(E6AP)-peptide at its C-terminal part. In 16E6 (PDB 4XR8, figure 3 D) R10 forms a hydrogen bond with carbonyl oxygen of G^6 of the LxxLL(E6AP)-peptide (trimer A) or a salt bridge with E^7 or E^8 of the LxxLL(E6AP)-peptide (trimer B) (The numbering of the LxxLL(E6AP)-peptide is: [$E^{-6}S^{-5}S^{-4}E^{-3}L^{-2}T^{-1}$**$L^1Q^2E^3L^4L^5$**$G^6E^7E^8R^9$]). In 31E6, R10 adopts a conformation which does not contribute to peptide binding (figure 3C). Further, some amino acids of the LxxLL(E6AP) binding pocket of 31E6 diverge from 16E6 (figure 1 and table 2). The difference in position 78 (H in 16E6 and W in 31E6) retains the polar nitrogen atom (N3 of the imidazole ring compared to the indol nitrogen) but the water-bridged hydrogen bond as in 16E6/LxxLL(E6AP) (PDB 4GIZ) or direct interaction as in 16E6/LxxLL(E6AP)/p53core (PDB 4XR8) is not retained. Another important difference between 31E6 and 16E6 is position 129. 16E6 R129 contacts two amino acids of the LxxLL(E6AP) peptide (trimer B, figure 3D). This interaction

was revealed by molecular dynamics simulation (22) based on the dimeric 16E6/LxxLL(E6AP) (PDB 4GIZ) structure and was later directly found in the crystal structure of the ternary 16E6/LxxLL(E6AP)/p53core complex (PDB 4XR8). 31E6 G129 is not able to establish these sidechain interactions (hydrogen bond and salt bridge). 16E6 C51 contributes to LxxLL(E6AP) binding via a hydrogen bond between the C51 amine and the LxxLL(E6AP) backbone carbonyl. C51 is not conserved among the alpha-9 HPV group (figure 1). 31E6 T51 contributes to LxxLL(E6AP) binding via a hydrogen bond between its sidechain hydroxyl group and the LxxLL(E6AP) backbone carbonyl. In order to analyze, if these minor amino acid variations impact LxxLL(E6AP) binding, a 16E6 mutant analogous to 31E6 (16E6-C51T/H78W/R129G) was tested by fluorescence anisotropy (figure 2) expecting a decreased binding to the LxxLL(E6AP)-peptide. Indeed, the binding decreased dramatically with a non-detectable interaction in competitive fluorescence anisotropy.

Taken together, minor amino acid variations together with structural differences between 16E6 and 31E6 accomplish an impaired binding of 31E6 to the LxxLL(E6AP)-peptide.

31E6-LxxLL(E6AP) interaction with p53

The ternary complex 31E6-LxxLL(E6AP) + p53core was modeled based on the published 16E6/LxxLL(E6AP)/p53core structure (PDB 4XR8, (17)) and the herein presented structure of 31E6-LxxLL(E6AP) by superimposing 31E6 onto 16E6 of trimer B. Compared to 16E6 the p53 interface of 31E6 shows amino acids essentially contributing to p53 binding of which some are not conserved (figure 1 and table 3). Above all, 16E6 E18 is not conserved in 31E6 and other alpha-9 HPV. The mutant 16E6 E18A has been described previously to reduce E6AP pulldown and p53 degradation efficiency by more than 50 % (17, 26). Our structural data clearly show,

that in 31E6 A18 the interaction with p53 would be disfavored because the loss of the salt bridge between 16E6 E18 and p53 K101 is not compensated in 31E6 (figure 3A). A second difference, which is unique to 31E6 in the alpha-9 genus, is a double amino acid variation of I23/Y43 in 16E6 and Y23/L43 in 31E6. This double amino acid variation seems to maintain an efficient hydrophobic packing at the E6/p53 interface in 31E6. In 16E6 I23/Y43 are important for binding of the α3 helix of p53 (amino acid 277-292) by hydrophobic interaction. Presumably, L289 of this helix is still capable to bind to the hydrophobic pocket in 31E6 (figure 3G). Remarkably, the hydroxyl groups of the tyrosines Y43 in 16E6 and Y23 in 31E6 are almost at the same position, indicating, that the association to the carbonyl oxygen of K292 of p53 is retained in 31E6. The amino acids Q6 and Q14 of 16E6, which contribute to p53 binding (17), are neither conserved in 31E6 nor in the alpha-9 HPV genus (figure 1). Q6 contributes backbone-backbone interactions which are likely impaired in 31E6 A6. Q14 of 16E6 is responsible for sidechain-sidechain interactions with T102 and N268 of p53. Potentially, E14 of 31E6 would still interact.

To complement the analysis of the ternary complex E6/E6AP/p53, we analyzed the interaction of the MBP-31E6-LxxLL(E6AP) construct, which was used for crystallization, to the p53 core domain quantitatively using microscale thermophoresis (MST). The direct fusion of the LxxLL-motif to the E6 protein mimics E6/E6AP complex formation, which is required for p53 binding. As a comparison, an analogous MBP-16E6-LxxLL(E6AP) construct was used. The affinity of the MBP-31E6-LxxLL(E6AP) construct to the p53 core domain (K_d = 91.7±1.26 µM) was 5-fold lower compared with the analogous MBP-16E6-LxxLL construct (K_d = 18.1±2.47 µM) (figure 4). The binding of 31E6 to the p53 core domain appears to be reduced, which agrees with the structural differences observed between 16E6 and 31E6 and the

255 sequence variations (table 3). Since it was previously shown, that the 16E6 E18A
256 mutant binds to and degrades p53 less efficiently (17, 26), we designed a 31E6 A18E
257 mutant analogous to 16E6 expecting an increased binding to the p53 core domain.
258 Here the binding increased slightly to 66.6±20.7 µM, indicating that the E18
259 contributes to the binding of the p53 core domain but that additional protein features
260 account for p53 core domain binding.

261

Discussion

All “high-risk” HPV types inactivate the tumor suppressor protein p53 via E6AP-dependent proteasomal degradation, which promotes cell immortalization. However, they differ in their ability to degrade p53, what may affect carcinogenic potential. E6 recruits E6AP and p53 to form the ternary complex E6/E6AP/p53, which is required for p53 degradation. The inactivation of p53 via proteasomal degradation is based on the formation of the ternary complex E6/E6AP/p53 (17). Differences in the assembly of this complex can alter the p53 degradation efficiency. The scope of this work was to characterize the ternary complex of two very closely related alpha-9 “high-risk” HPV types 16 and 31 in order to investigate whether (I) phylogenetic similarity results in structural conservation and (II) the binding of E6 proteins to the same cellular targets differs structurally and quantitatively between 16E6 and 31E6.

As expected, the overall structure of 31E6 resembles that of 16E6: two zinc-binding domains E6N and E6C forming a binding cleft for LxxLL-motifs (figure 3). Due to the high sequence conservation, especially of the zinc-binding motifs of the HPV E6 proteins this can very likely be claimed also for other HPV E6 proteins. The sequence alignment of HPV alpha-9 E6 (figure 1 and table 2) shows, that the LxxLL-motif binding site of E6 is highly conserved. Amino acid 16E6L50 in the hydrophobic LxxLL-motif binding pocket, which abolishes LxxLL(E6AP) binding if mutated (27, 28), is conserved between 16E6 and 31E6, and within all alpha species. Additionally, mutations of 16E6 R102 and R131 to alanine largely impair E6AP interaction. These amino acids are also conserved in alpha-9 HPV types and contribute to LxxLL(E6AP) binding by polar interactions. However, we found that the affinity of 31E6 to the LxxLL(E6AP)-peptide is 2-fold lower than of 16E6 to the same peptide. Structural comparison of 16E6/LxxLL(E6AP) and 31E6-LxxLL(E6AP) showed a slightly shifted

287 E6C domain. Indeed, the two different conformations of the E6C domain can be
288 related to the heterogeneous dynamic behavior of the E6C domain, which was
289 previously reported in NMR solution studies performed on various E6C domains (29-
290 31). The E6C domain is one building block of the LxxLL(E6AP) binding cleft.
291 Subsequently, flexibility of the E6C domain can be one reason for different binding
292 affinities to the LxxLL(E6AP)-peptide. Moreover, sequence differences between 16E6
293 and 31E6 as described in table 2, participating in LxxLL(E6AP) binding, result in less
294 protein contacts in 31E6. 16E6, if mutated to the 31E6 analogous amino acids
295 (C51T, H78W, R129G), resulted in a tremendously reduced binding affinity to the
296 LxxLL(E6AP)-peptide. These amino acids are not conserved in the alpha-9 genus at
297 all (Figure 1). As a conclusion, minor amino acid variations are another possibility of
298 the lower affinity of 31E6 to LxxLL(E6AP).

299 Notably, 16E6 mutants showing impaired binding to LxxLL(E6AP) also showed less
300 efficient p53 degradation (22, 28). As neither E6 nor E6AP alone are able to interact
301 directly with p53 (32-35), the binding to p53 requires the formation of E6/E6AP
302 complex. The binding of E6 to the LxxLL(E6AP)-peptide is sufficient to recruit the
303 core domain of p53 (17). However, it was recently reported, that additional binding
304 sites at the N-terminal region of E6AP are necessary to stimulate the ubiquitin-ligase
305 activity of E6AP by 16E6 (28). Here, we focused on one interaction site of E6 and
306 E6AP, the LxxLL(E6AP) motif, necessary for p53 binding, but not sufficient for p53
307 degradation. The structure and binding affinities of the functional complex in terms of
308 p53 ubiquitination are still elusive.

309 The binding of E6AP to E6 is required prior to binding of p53 to E6. In order to
310 investigate the binding of p53 to E6, we mimicked an “p53-ready” E6 by fusing the
311 LxxLL(E6AP)-peptide to the C-term of E6 (Figure 5). In this proximity, the

LxxLL(E6AP)-peptide is bound to E6 and therefore, the measured binding of the p53 core domain is presumably independent from the required binding of LxxLL(E6AP). Of course, in the cellular environment the sequential binding of E6AP and p53 to E6 finally determines p53 degradation.

Some amino acids (D44, F47, D49), crucial for p53 core domain interaction in the 16E6/LxxLL/p53 complex are conserved within 16E6 and 31E6, suggesting that 31E6 can bind to p53, table 3).

However, the 31E6 binding site shows striking amino acid differences compared to 16E6. Of these amino acid variations, it was shown, that 16E6 mutants Q6A and Q14A (not conserved in alpha HPV) bind to E6AP and degrade p53 similarly to wildtype 16E6 *in cellulo* (17) indicating that variations at these positions have minor influence on p53 binding and degradation *in cellulo*. In contrast the 16E6 E18A mutant showed 75 % lower binding to p53 and decrease in p53 degradation efficiency (17). Strikingly, in 31E6 this position is an alanine residue (A18). Indeed, the mutation of A18 to E18 in 31E6 resulted in an increased affinity to p53 core domain. This position may have an influence on p53 binding affinity and subsequent degradation and accordingly it is subject to variation across alpha species, where it is not conserved, neither in “high-risk”, nor in “low-risk” HPV types. The gain of affinity of the 31E6A18E mutant is rather low, indicating that additional variations between 16E6 and 31E6 influence the binding to the p53 core domain, like the observed shift in the E6C domain or other sequence variation. For example, our structural analysis indicates that the sequence variation 16E6Y43 and 31E6L43 is compensated by 16E6I23 and 31E6Y23, retaining the hydrophobic pocket for p53 binding. However, slight variations in the p53 binding pocket can also lead to different affinities. Overall,

the sequence differences and structural analysis parallels the obtained 5.4-fold lower K_d for 31E6-LxxLL(E6AP) binding to the p53 core domain.

The formation of the ternary complex is presumably stronger for 16E6, because it shows higher affinities to both, LxxLL(E6AP) and the p53 core domain. Our affinity analysis strongly agrees with the previously reported >2-fold less efficient degradation of p53 in HPV 31 E6 transfected cells compared to 16E6 transfected cells, even though 31E6 shows an almost 3-fold higher cellular level than 16E6 in these experiments (20).

HPV 18 is the second most prevalent HPV type associated with cervical cancer and belongs to the alpha-7 HPV species. The intracellular level of 18E6 resembles 16E6, but it shows almost 2-fold less efficient p53 degradation (20). Overall, 18E6 shares less sequence identity (~57 %) with 16E6 compared to 31E6 with 16E6 (~66 %). Slight structural differences, e.g. the position of E6C domain are not predictable but can change the binding to the p53 core domain and E6AP. 18E6 does possess all crucial amino acids, necessary for p53 core domain binding, only 16D44 is found as the homologous amino acid E in 18E6. On the other hand, LxxLL(E6AP) binding 16R131, which shows ~ 50 % reduced binding to LxxLL(E6AP) if mutated to A, is not conserved in 18E6 (18E6 H131, see figure 1). These variations can potentially influence the efficiency of ternary complex formation and subsequent p53 degradation. “Low-risk” HPV types already show a much lower sequence identity to 16E6 (e.g. 11E6 36 %) and already possess amino acid differences, which neglect binding to p53 (e.g.11E6 has no conserved 16E6 E18, D44, F47, D49 crucial for p53 core domain binding). Consequently, “low-risk” HPV types are inactive in E6AP-dependent p53 degradation at all (20).

It is important to note, that apart from defined crucial amino acids in 16E6 (table 2 and 3) the individual subset of minor sequence differences and the flexibility of the E6C domain position, influence E6 structure and the binding to E6AP and p53. These sequence variations increase with decreasing phylogenetic relation of E6 proteins. Their impact on the E6 structure and binding affinities is not predictable. As a conclusion, binding affinities certainly vary between the E6 proteins but must be analyzed individually.

Further, it must be pointed out, that the alpha-9 “high-risk” HPV types 52 and 58 E6 proteins show higher p53 degradation efficiencies *in cellulo* than 16E6 despite similar intracellular protein levels of E6 proteins (20), but possess a lower carcinogenic potential. Here, p53 degradation potential does not correlate with the cancerogenic risk (20, 36). The physiological context likely represents a more complex situation. Apart from p53 degradation, many other factors contribute to viral persistence and HPV-associated cancer which further differ among different HPV genus, species and types. Carcinogenicity of HPV may be influenced by many parameters, including the entire viral interactome of E1 to E7, transcription regulation, half-life of proteins, deregulation of posttranslational modifications (37-39); all playing a role in DNA-damage response (reviewed in (40-42)), persistence and immune response ((43) and reviewed in (44, 45)), E6-mediated degradation of other cell proliferation regulatory proteins, e.g. NHERF1 (27) and still elusive factors. Notably, human keratinocytes can be immortalized by 16E7 alone (46). Co-expression with 16E6 increases the immortalization rate (47), highlighting the important concomitant role of E7 in HPV-associated carcinogenesis. Nonetheless, the inactivation of p53 remains a very important process with respect to cell immortalization. A 16E6 mutant deficient in p53

384 interaction showed tremendously decreased potency of cell immortalization though
385 being co-expressed with 16E7 (22).

386 Further, the multiple interactions of E6 play important roles in cell immortalization
387 (48). E6 proteins can bind to various LxxLL-motifs of other cellular targets, e.g.
388 hTERT (49) and IRF3 (50, 51). Additionally, E6 PDZ-binding motifs (PBM) differ even
389 within HPV species (reviewed in (52)). The last 4 amino acids of the PBM of 31E6
390 (ETQV) differ only slightly from 16E6 (ETQL). However, it was shown previously, that
391 HPV 18 E6 (ETQV) has a different PDZ-binding profile compared with 16E6 (55).
392 Assumingly the PDZ-binding profile of 31E6 is also different from 16E6. It was
393 shown, that an interaction of E6 with the PDZ-containing protein MAGI-1 results in
394 degradation of MAGI-1, for carcinogenic as well as non-carcinogenic types (36, 53) of
395 alpha papillomaviruses in a very similar efficiency. The authors draw the conclusion
396 that MAGI1 degradation alone cannot result in carcinogenesis but in concert with p53
397 degradation and hTERT (54). In conclusion, different PDZ-binding partners can also
398 influence the carcinogenic potential of E6.

399 The interactome of both oncoproteins, E6 and E7, facilitates cell transformation. The
400 p53 degradation potential is one important factor in carcinogenesis especially for
401 “high-risk” types. Semi-quantitative analysis revealed a link between carcinogenicity
402 and p53 degradation (20) for some HPV types, as HPV 16 and 31. Further, HPV 16
403 is associated with ~ 50 % cervical cancers, in contrast HPV 31 is only associated
404 with 3-8 % (56, 57). This difference is even more significant in HPV positive tumors of
405 the oropharynx, where HPV 16 accounts for 93 % and HPV 31 together with 12 other
406 HPV types for 4 % of these cancers (58). In this context it is interesting, that 16E6
407 and 31E6 share the same structural fold but 31E6 forms the ternary complex
408 E6/E6AP/p53 less efficiently. Consequently, the E6-mediated proteasomal

409 degradation of p53 can be impaired. In principle, these findings are likely conferrable
410 to the alpha-9 species and beyond, and are not limited to the proteins of ternary
411 complex, analyzed here. In summary, additionally to the diverse interactions of E6
412 with different interaction partners (qualitative differences like 16E6 binds to E6AP and
413 8E6 binds to MAML1 (16, 59, 60)), divergence of E6 proteins could also be explained
414 by different affinities of very closely related proteins to the same cellular targets
415 (quantitative differences like 16E6 and 31E6 bind to LxxLL(E6AP) and p53 with
416 different affinities).

417

Materials and Methods

Recombinant Protein Production and Purification

The E6 proteins possess cysteine mutations to decrease oxidation and oligomerization. An overview of used E6 constructs is given in figure 5. Proteins are fused N-terminally to maltose-binding protein (MBP) to increase solubility and C-terminally to LxxLL(E6AP)-peptides in pETxM1 plasmids. The proteins were produced in BL21(DE3) at 20 °C after addition of 100 µM $ZnCl_2$ and induction with 1 mM IPTG overnight at OD_{600} ~0.8. MBP-31E6 was expressed in TB-medium (12 g/L tryptone, 24 g/L yeast extract, 4 mL/L glycerol, 5 g/L NaCl, 0.017 M KH_2PO_4, 0.072 M K_2HPO_4), MBP-16E6 and MBP-31E6-LxxLL(E6AP) in LB-medium (10 g/L tryptone, 10 g/L NaCl, 5 g/L yeast extract).

Proteins were lysed in 50 mM Tris-HCl pH 8 at 8 °C, 400 mM NaCl, 5 % (w/v glycerol, 2 mM Tris(2-carboxyethyl)phosphine (TCEP) using a micro-fluidizer or French press. Lysate was cleared by centrifugation (1 h, 100 000 x g, 4°C) and applied to an equilibrated affinity column (self-packed amylose column, 30 mL amylose resin (New England Biolabs) for MBP-16E6, MBP-16E6-LxxLL(E6AP), MBP-31E6A18E-LxxLL(E6AP), MBP-16E6mut and MBP-31E6-LxxLL(E6AP)) or MBPTrap (GE Healthcare) 3x 5 mL in a row for MBP-31E6. Proteins were eluted with lysis buffer containing 10 mM maltose. Elution fractions were pooled and centrifuged overnight at 100 000 x g to sediment agglomerates. The supernatant was applied to a S200 XK16/60 or XK26/60 column (GE healthcare) pre-equilibrated with 50 mM Tris-HCl pH 8 at 8 °C, 400 mM NaCl, 2 mM TCEP. Fractions containing monomeric protein were pooled, if needed concentrated (>45 µM for FP, >60 mg/mL for Xtal, >20 µM for Microscale Thermophoresis (MST)) and stored at -80 °C.

442 The p53 core domain construct is produced as a His_6-MBP-p53core following the
443 protocol previously described (61). After MBP-affinity chromatography (MBPTrap (GE
444 Healthcare) 3x 5 mL in a row) the N-terminal His_6 -MBP-fusion was cleaved by His_6 -
445 TEV-Protease followed by Heparin column (GE healthcare, 5 mL) to separate
446 cleaved MBP and TEV-protease from p53core and size exclusion chromatography
447 (Superdex S200 XK 16/60, GE healthcare) using the same buffers as for E6.

448 **Fluorescence Anisotropy**

449 Fluorescence anisotropy (FA) was measured with a PHERAstar (BMG Labtech,
450 Offenburg, Germany) microplate reader by using 485 ± 20 nm and 528 ± 20 nm
451 band-pass filters. In direct measurements, a dilution series of the MBP-E6 protein
452 was prepared in 96 well plates (96 well skirted PCR plate, 4ti-0740, 4titude, Wotton,
453 UK) in a 20 mM HEPES pH 7.5 buffer that contained 150 mM NaCl, 0.5 mM TCEP,
454 0.005 % Tween 20 and 50 nM fluorescently labeled E6AP-peptide (fE6AP, Fl-ttds-
455 PESSELTLQELLGEER). The volume of the dilution series was 40 µL, which was
456 later divided into three technical replicates of 10 µL during transferring to 384 well
457 micro-plates (low binding microplate, 384 well, E18063G5, Greiner Bio-One,
458 Kremsmünster, Austria). In total, the anisotropy of the probe was measured at 8
459 different protein concentrations (whereas one contained no protein and corresponded
460 to the free peptide). In competitive FA measurements, the same buffer was
461 supplemented with the E6 protein and fluorescently labeled LxxLL(E6AP)-peptide to
462 achieve a complex formation of 60-80 %, at concentrations based on the titration of
463 direct binding. Then, this mixture was used for creating a dilution series of the non-
464 fluorescent competitor (i.e. the biotinylated LxxLL-peptide of E6AP, Biotin-ttds-
465 PESSELTLQELLGEER) and the measurement was carried out identically as in the
466 direct experiment. Analyses of FA experiments were carried out in ProFit (62).

GPCA

GPCA was performed as previously described (23, 25).

HEK-293T Pasteur cells were reverse-transfected in white 96-well plates at a concentration of 4.2 x 10^5 cells per well using PEI MAX (Polysciences) with 100 ng of pSPICA-N2-E6 and 100 ng of pSPICA-N1-target protein E6APΔ1-290. These plasmids encode E6 and E6AP fused to the split fragments G1 or G2 (figure 5) of the Gaussia luciferase (G), which complement the enzyme upon interaction of E6 and E6AP. At 48 h post-transfection, cells were washed with 50 µL of PBS and lysed with 40 µL of Renilla Lysis Buffer (Promega, Madison, WI, USA, E2820) for 30 min. Split *Gaussia princeps* luciferase enzymatic activity was measured using a Berthold Centro LB960 luminometer by injecting 50 µL of luciferase substrate reagent (Promega, E2820) per well and counting luminescence for 10 s. Results were expressed as a x-fold change of signal normalized over the sum of controls, specified herein as normalized luciferase ration (NLR). For a given protein pair A/B: NLR = (G1-A + G2-B)/ [(G1-A + G2) + (G1 + G2-B)] as described in (24).

Crystallization and structure refinement

The MBP-31E6-GSSGSGSGSGSGSAAA-LxxLL(E6AP) fusion protein was cloned into pETxM1, purified as described, concentrated to 30 mg/mL in 50 mM Tris pH 8.0, 150 mM NaCl, 2 mM TCEP, and crystallized. Crystals were obtained at 20 °C in 200 mM tri-lithium citrate, 200 mM ammonium citrate tribasic pH 7.0, 20% (w/v) polyethylene glycol 3350. All crystals were flash-cooled in a cryo-protectant solution containing 25% (w/v) glycerol in crystallization buffer and stored in liquid nitrogen before data collection. X-ray diffraction data were collected at the Synchrotron Swiss Light Source (SLS) (Switzerland) on the X06DA (PXIII) beamline and processed with

the program XDS (63). The crystal structure was solved by molecular replacement with a high resolution MBP structure (PDB ID: 5H7Q, (64)) and 16E6 (PDB ID: 4XR8, (17)) using Phaser (65) and structure refinement was carried out with PHENIX (66). TLS refinement was applied during the refinement. Coot was used for model building (67). The crystallographic parameters and the statistics of data collection and refinement are shown in Table 1. The refined model and the structure factor amplitudes have been deposited in the PDB with the accession code 6SLM. Figures were prepared using PyMOL 2.3.3. RMSD (root-mean-square deviation) per residue was calculated on superpositioned structures by calling the rms_cur function in PyMOL on every pair of corresponding Cα atoms using a custom Python script. Rms_cur returns the rms difference of the selection (here Cα distance) without performing any superpositioning. The structure of 31E6 was used as the reference. RMSD (here Cα distance) was plotted against the residue position. Superposition was based on the ligand (LxxLL-peptide). Trimer B of the ternary complex in PDB 4XR8 (17) was used for comparison.

Microscale Thermophoresis

The affinity of MBP-E6-LxxLL(E6AP) to p53 was determined by the K_d using Microscale Thermophoresis (MST), a small-scale method based on the principle of thermophoresis (68-70). The purified proteins MBP-31E6-LxxLL(E6AP), MBP-31E6A18E-LxxLL(E6AP) and MBP-16E6-LxxLL(E6AP) were labeled at lysine residues with the Monolith NT^{TM} Protein labeling Kit RED-NHS first generation (NanoTemper Technologies GmbH, Munich, Germany) following the user manual instructions. Note, that (I) here the 16E6 does not carry the F47R mutation, because this mutation diminishes p53 interaction (22) and (II) p53 core domain is used, because it was previously shown, that only this region binds to E6 (17, 61).

In order to ensure a sufficient fluorescence signal, to prevent adsorption to the capillary walls and to exclude auto fluorescence of the measurement buffer (20 mM HEPES pH 6.8 at room temperature, 200 mM NaCl, 0.5 mM TCEP, 0.05 % Tween 20), a pre-test was performed according to the manufacturer manual and evaluated positive. Binding was measured by incubating 200 nM of the labeled MBP-16E6-LxxLL(E6AP) or MBP-31E6-LxxLL(E6AP) with a 2-fold serial dilution of p53 core domain, starting with 87.5 µM or 600 µM, respectively. Samples at higher ligand concentrations showed aggregation. The sample preparation was performed as recommended by MO.Control software (NanoTemper Technologies GmbH, Munich, Germany). The measurements were performed at 25 °C using 5 % excitation power for MBP-16E6-LxxLL(E6AP) or 20 % excitation power for MBP-31E6-LxxLL(E6AP) and MBP-31E6A18E-LxxLL(E6AP). Standard treated capillaries (Monolith NT.115 Capillary) were used. Thermophoresis was measured using the Monolith NT.115 (NanoTemper Technologies GmbH, Munich, Germany) and analyzed by MO.Affinity analysis software (NanoTemper Technologies GmbH, Munich, Germany). Samples leading to heterogenous fluorescence intensity were neglected. The K_d was calculated using the MO.Affinity analysis software (K_d-model) only fixing the template concentrations and implying fluorescence signals at 19 to 20 s of thermophoresis for all samples.

Data availability

The structural data of the MBp-31E6-LxxLL(E6AP) are deposited in PDB with the accession code 6SLM.

539 **Acknowledgments**

540 We gratefully thank Joerg Martin and Lorena Voehringer (MPI Tuebingen) for their
541 kind support with MST measurements. X-ray data collection was performed on the
542 PXIII beamline at the Swiss Light Source synchrotron, P. Scherrer Institute, Villigen,
543 Switzerland. We thank V. Olieric and C.-Y. Huang for their help on the beamline.

544 Further this work received institutional support from le Centre National de la
545 Recherche Scientifique (CNRS), Université de Strasbourg, Institut National de la
546 Santé et de la Recherche Médicale (INSERM) and Région Alsace. The work was
547 supported in part by grants from Ligue contre le Cancer (équipe labellisée 2015 and
548 fellowship to A.B.), Ligue contre le Cancer CCIR-GE, ANR (Infect-ERA program,
549 project “HPV motiva”), Fondation recherche Médicale (fellowship to A.B.), National
550 Institutes of Health (Grant R01CA134737), Instruct (ESFRI), and the French
551 Infrastructure for Integrated Structural Biology (FRISBI, ANR-10-INBS-05) and
552 Instruct-ERIC. The authors declare that the content is solely their responsibility and
553 does not represent the official views of the National Institutes of Health.

554 G.G. was supported by the Post-doctorants en France program of the Fondation
555 ARC Pour La Recherche Sur Le Cancer.

556

References

1. **de Villiers EM, Gunst K, Stein H, Scherubl H.** 2004. Esophageal squamous cell cancer in patients with head and neck cancer: Prevalence of human papillomavirus DNA sequences. Int J Cancer **109:**253-258.
2. **IARC.** 2007. Human papillomaviruses. IARC Monographs on the evaluation of carcinogenic risks to humans **90:**1-636.
3. **de Martel C, Plummer M, Vignat J, Franceschi S.** 2017. Worldwide burden of cancer attributable to HPV by site, country and HPV type. Int J Cancer **141:**664-670.
4. **Munoz N, Bosch FX, Castellsague X, Diaz M, de Sanjose S, Hammouda D, Shah KV, Meijer CJ.** 2004. Against which human papillomavirus types shall we vaccinate and screen? The international perspective. Int J Cancer **111:**278-285.
5. **Munoz N, Bosch FX, de Sanjose S, Herrero R, Castellsague X, Shah KV, Snijders PJ, Meijer CJ, International Agency for Research on Cancer Multicenter Cervical Cancer Study G.** 2003. Epidemiologic classification of human papillomavirus types associated with cervical cancer. N Engl J Med **348:**518-527.
6. **Cogliano V, Baan R, Straif K, Grosse Y, Secretan B, El Ghissassi F, Cancer WHOIAfRo.** 2005. Carcinogenicity of human papillomaviruses. Lancet Oncol **6:**204.
7. **Cancer Genome Atlas Research N, Albert Einstein College of M, Analytical Biological S, Barretos Cancer H, Baylor College of M, Beckman Research Institute of City of H, Buck Institute for Research on A, Canada's Michael Smith Genome Sciences C, Harvard Medical S, Helen FGCC, Research Institute at Christiana Care Health S, HudsonAlpha Institute for B, Ilsbio LLC, Indiana University School of M, Institute of Human V, Institute for Systems B, International Genomics C, Leidos B, Massachusetts General H, McDonnell Genome Institute at Washington U, Medical College of W, Medical University of South C, Memorial Sloan Kettering Cancer C, Montefiore Medical C, NantOmics, National Cancer I, National Hospital AN, National Human Genome Research I, National Institute of Environmental Health S, National Institute on D, Other Communication D, Ontario Tumour Bank LHSC, Ontario Tumour Bank OIfCR, Ontario Tumour Bank TOH, Oregon H, Science U, Samuel Oschin Comprehensive Cancer Institute C-SMC, International SRA, St Joseph's Candler Health S, Eli, Edythe LBIoMIoT, Harvard U, Research Institute at Nationwide Children's H, Sidney Kimmel Comprehensive Cancer Center at Johns Hopkins U, University of B, University of Texas MDACC, University of Abuja Teaching H, University of Alabama at B, University of California I, University of California Santa C, et al.** 2017. Integrated genomic and molecular characterization of cervical cancer. Nature **543:**378-384.
8. **Hudson JB, Bedell MA, McCance DJ, Laiminis LA.** 1990. Immortalization and altered differentiation of human keratinocytes in vitro by the E6 and E7 open reading frames of human papillomavirus type 18. J Virol **64:**519-526.
9. **White EA, Munger K, Howley PM.** 2016. High-Risk Human Papillomavirus E7 Proteins Target PTPN14 for Degradation. mBio **7**.
10. **Szalmas A, Tomaic V, Basukala O, Massimi P, Mittal S, Konya J, Banks L.** 2017. The PTPN14 Tumor Suppressor Is a Degradation Target of Human Papillomavirus E7. J Virol **91**.
11. **Hatterschide J, Bohidar AE, Grace M, Nulton TJ, Kim HW, Windle B, Morgan IM, Munger K, White EA.** 2019. PTPN14 degradation by high-risk human papillomavirus E7 limits keratinocyte differentiation and contributes to HPV-mediated oncogenesis. Proc Natl Acad Sci U S A **116:**7033-7042.
12. **Yun HY, Kim MW, Lee HS, Kim W, Shin JH, Kim H, Shin HC, Park H, Oh BH, Kim WK, Bae KH, Lee SC, Lee EW, Ku B, Kim SJ.** 2019. Structural basis for recognition of the tumor suppressor protein PTPN14 by the oncoprotein E7 of human papillomavirus. PLoS Biol **17:**e3000367.

13. **White EA, Sowa ME, Tan MJ, Jeudy S, Hayes SD, Santha S, Munger K, Harper JW, Howley PM.** 2012. Systematic identification of interactions between host cell proteins and E7 oncoproteins from diverse human papillomaviruses. Proc Natl Acad Sci U S A **109:**E260-267.
14. **Hiller T, Poppelreuther S, Stubenrauch F, Iftner T.** 2006. Comparative analysis of 19 genital human papillomavirus types with regard to p53 degradation, immortalization, phylogeny, and epidemiologic risk classification. Cancer Epidemiol Biomarkers Prev **15:**1262-1267.
15. **Fu L, Van Doorslaer K, Chen Z, Ristriani T, Masson M, Trave G, Burk RD.** 2010. Degradation of p53 by human Alphapapillomavirus E6 proteins shows a stronger correlation with phylogeny than oncogenicity. PLoS One **5**.
16. **Brimer N, Drews CM, Vande Pol SB.** 2017. Association of papillomavirus E6 proteins with either MAML1 or E6AP clusters E6 proteins by structure, function, and evolutionary relatedness. PLoS Pathog **13:**e1006781.
17. **Martinez-Zapien D, Ruiz FX, Poirson J, Mitschler A, Ramirez J, Forster A, Cousido-Siah A, Masson M, Vande Pol S, Podjarny A, Trave G, Zanier K.** 2016. Structure of the E6/E6AP/p53 complex required for HPV-mediated degradation of p53. Nature **529:**541-545.
18. **Talis AL, Huibregtse JM, Howley PM.** 1998. The role of E6AP in the regulation of p53 protein levels in human papillomavirus (HPV)-positive and HPV-negative cells. J Biol Chem **273:**6439-6445.
19. **Madeira F, Park YM, Lee J, Buso N, Gur T, Madhusoodanan N, Basutkar P, Tivey ARN, Potter SC, Finn RD, Lopez R.** 2019. The EMBL-EBI search and sequence analysis tools APIs in 2019. Nucleic Acids Res **47:**W636-W641.
20. **Mesplede T, Gagnon D, Bergeron-Labrecque F, Azar I, Senechal H, Coutlee F, Archambault J.** 2012. p53 degradation activity, expression, and subcellular localization of E6 proteins from 29 human papillomavirus genotypes. J Virol **86:**94-107.
21. **Suarez I, Trave G.** 2018. Structural Insights in Multifunctional Papillomavirus Oncoproteins. Viruses **10**.
22. **Zanier K, Charbonnier S, Sidi AO, McEwen AG, Ferrario MG, Poussin-Courmontagne P, Cura V, Brimer N, Babah KO, Ansari T, Muller I, Stote RH, Cavarelli J, Vande Pol S, Trave G.** 2013. Structural basis for hijacking of cellular LxxLL motifs by papillomavirus E6 oncoproteins. Science **339:**694-698.
23. **Poirson J, Biquand E, Straub ML, Cassonnet P, Nomine Y, Jones L, van der Werf S, Trave G, Zanier K, Jacob Y, Demeret C, Masson M.** 2017. Mapping the interactome of HPV E6 and E7 oncoproteins with the ubiquitin-proteasome system. FEBS J **284:**3171-3201.
24. **Cassonnet P, Rolloy C, Neveu G, Vidalain PO, Chantier T, Pellet J, Jones L, Muller M, Demeret C, Gaud G, Vuillier F, Lotteau V, Tangy F, Favre M, Jacob Y.** 2011. Benchmarking a luciferase complementation assay for detecting protein complexes. Nat Methods **8:**990-992.
25. **Neveu G, Cassonnet P, Vidalain PO, Rolloy C, Mendoza J, Jones L, Tangy F, Muller M, Demeret C, Tafforeau L, Lotteau V, Rabourdin-Combe C, Trave G, Dricot A, Hill DE, Vidal M, Favre M, Jacob Y.** 2012. Comparative analysis of virus-host interactomes with a mammalian high-throughput protein complementation assay based on Gaussia princeps luciferase. Methods **58:**349-359.
26. **Li S, Hong X, Wei Z, Xie M, Li W, Liu G, Guo H, Yang J, Wei W, Zhang S.** 2019. Ubiquitination of the HPV Oncoprotein E6 Is Critical for E6/E6AP-Mediated p53 Degradation. Front Microbiol **10:**2483.
27. **Drews CM, Case S, Vande Pol SB.** 2019. E6 proteins from high-risk HPV, low-risk HPV, and animal papillomaviruses activate the Wnt/beta-catenin pathway through E6AP-dependent degradation of NHERF1. PLoS Pathog **15:**e1007575.
28. **Drews CM, Brimer N, Vande Pol SB.** 2020. Multiple regions of E6AP (UBE3A) contribute to interaction with papillomavirus E6 proteins and the activation of ubiquitin ligase activity. PLoS Pathog **16:**e1008295.
29. **Nomine Y, Charbonnier S, Miguet L, Potier N, Van Dorsselaer A, Atkinson RA, Trave G, Kieffer B.** 2005. 1H and 15N resonance assignment, secondary structure and dynamic

behaviour of the C-terminal domain of human papillomavirus oncoprotein E6. J Biomol NMR **31:**129-141.

30. **Zanier K, ould M'hamed ould Sidi A, Boulade-Ladame C, Rybin V, Chappelle A, Atkinson A, Kieffer B, Trave G.** 2012. Solution structure analysis of the HPV16 E6 oncoprotein reveals a self-association mechanism required for E6-mediated degradation of p53. Structure **20:**604-617.
31. **Mischo A, Ohlenschlager O, Hortschansky P, Ramachandran R, Gorlach M.** 2013. Structural insights into a wildtype domain of the oncoprotein E6 and its interaction with a PDZ domain. PLoS One **8:**e62584.
32. **Huibregtse JM, Scheffner M, Howley PM.** 1994. E6-AP directs the HPV E6-dependent inactivation of p53 and is representative of a family of structurally and functionally related proteins. Cold Spring Harb Symp Quant Biol **59:**237-245.
33. **Huibregtse JM, Scheffner M, Howley PM.** 1993. Localization of the E6-AP regions that direct human papillomavirus E6 binding, association with p53, and ubiquitination of associated proteins. Mol Cell Biol **13:**4918-4927.
34. **Scheffner M, Huibregtse JM, Howley PM.** 1994. Identification of a human ubiquitin-conjugating enzyme that mediates the E6-AP-dependent ubiquitination of p53. Proc Natl Acad Sci U S A **91:**8797-8801.
35. **Scheffner M, Huibregtse JM, Vierstra RD, Howley PM.** 1993. The HPV-16 E6 and E6-AP complex functions as a ubiquitin-protein ligase in the ubiquitination of p53. Cell **75:**495-505.
36. **Muench P, Hiller T, Probst S, Florea AM, Stubenrauch F, Iftner T.** 2009. Binding of PDZ proteins to HPV E6 proteins does neither correlate with epidemiological risk classification nor with the immortalization of foreskin keratinocytes. Virology **387:**380-387.
37. **Thatte J, Banks L.** 2017. Human Papillomavirus 16 (HPV-16), HPV-18, and HPV-31 E6 Override the Normal Phosphoregulation of E6AP Enzymatic Activity. J Virol **91**.
38. **Hsu CH, Peng KL, Jhang HC, Lin CH, Wu SY, Chiang CM, Lee SC, Yu WC, Juan LJ.** 2012. The HPV E6 oncoprotein targets histone methyltransferases for modulating specific gene transcription. Oncogene **31:**2335-2349.
39. **Langsfeld ES, Bodily JM, Laimins LA.** 2015. The Deacetylase Sirtuin 1 Regulates Human Papillomavirus Replication by Modulating Histone Acetylation and Recruitment of DNA Damage Factors NBS1 and Rad51 to Viral Genomes. PLoS Pathog **11:**e1005181.
40. **Bristol ML, Das D, Morgan IM.** 2017. Why Human Papillomaviruses Activate the DNA Damage Response (DDR) and How Cellular and Viral Replication Persists in the Presence of DDR Signaling. Viruses **9**.
41. **Anacker DC, Moody CA.** 2017. Modulation of the DNA damage response during the life cycle of human papillomaviruses. Virus Res **231:**41-49.
42. **Spriggs CC, Laimins LA.** 2017. Human Papillomavirus and the DNA Damage Response: Exploiting Host Repair Pathways for Viral Replication. Viruses **9**.
43. **Chiang C, Pauli EK, Biryukov J, Feister KF, Meng M, White EA, Munger K, Howley PM, Meyers C, Gack MU.** 2018. The Human Papillomavirus E6 Oncoprotein Targets USP15 and TRIM25 To Suppress RIG-I-Mediated Innate Immune Signaling. J Virol **92**.
44. **Westrich JA, Warren CJ, Pyeon D.** 2017. Evasion of host immune defenses by human papillomavirus. Virus Res **231:**21-33.
45. **Hong S, Laimins LA.** 2017. Manipulation of the innate immune response by human papillomaviruses. Virus Res **231:**34-40.
46. **Halbert CL, Demers GW, Galloway DA.** 1991. The E7 Gene of Human Papillomavirus Type 16 Is Sufficient for Immortalization of Human Epithelial Cells. J Virol **65:**473-478.
47. **Hawley-Nelson P, Vousden KH, Hubbert NL, Lowy DR, Schiller JT.** 1989. HPV16 E6 and E7 proteins cooperate to immortalize human foreskin keratinocytes. EMBO J **8:**3905-3910.
48. **Liu Y, Chen JJ, Gao Q, Dalal S, Hong Y, Mansur CP, Band V, Androphy EJ.** 1999. Multiple functions of human papillomavirus type 16 E6 contribute to the immortalization of mammary epithelial cells. J Virol **73:**7297-7307.

49. **Liu X, Dakic A, Zhang Y, Dai Y, Chen R, Schlegel R.** 2009. HPV E6 protein interacts physically and functionally with the cellular telomerase complex. Proc Natl Acad Sci U S A **106:**18780-18785.
50. **Ronco LV, Karpova AY, Vidal M, Howley PM.** 1998. Human papillomavirus 16 E6 oncoprotein binds to interferon regulatory factor-3 and inhibits its transcriptional activity. Genes Dev **12:**2061-2072.
51. **Shah M, Anwar MA, Park S, Jafri SS, Choi S.** 2015. In silico mechanistic analysis of IRF3 inactivation and high-risk HPV E6 species-dependent drug response. Sci Rep **5:**13446.
52. **Ganti K, Broniarczyk J, Manoubi W, Massimi P, Mittal S, Pim D, Szalmas A, Thatte J, Thomas M, Tomaic V, Banks L.** 2015. The Human Papillomavirus E6 PDZ Binding Motif: From Life Cycle to Malignancy. Viruses **7:**3530-3551.
53. **Van Doorslaer K, DeSalle R, Einstein MH, Burk RD.** 2015. Degradation of Human PDZ-Proteins by Human Alphapapillomaviruses Represents an Evolutionary Adaptation to a Novel Cellular Niche. PLoS Pathog **11:**e1004980.
54. **Van Doorslaer K, Burk RD.** 2012. Association between hTERT activation by HPV E6 proteins and oncogenic risk. Virology **433:**216-219.
55. **Vincentelli R, Luck K, Poirson J, Polanowska J, Abdat J, Blemont M, Turchetto J, Iv F, Ricquier K, Straub ML, Forster A, Cassonnet P, Borg JP, Jacob Y, Masson M, Nomine Y, Reboul J, Wolff N, Charbonnier S, Trave G.** 2015. Quantifying domain-ligand affinities and specificities by high-throughput holdup assay. Nat Methods **12:**787-793.
56. **Arbyn M, Verdoodt F, Snijders PJ, Verhoef VM, Suonio E, Dillner L, Minozzi S, Bellisario C, Banzi R, Zhao FH, Hillemanns P, Anttila A.** 2014. Accuracy of human papillomavirus testing on self-collected versus clinician-collected samples: a meta-analysis. Lancet Oncol **15:**172-183.
57. **de Oliveira CM, Fregnani JH, Carvalho JP, Longatto-Filho A, Levi JE.** 2013. Human papillomavirus genotypes distribution in 175 invasive cervical cancer cases from Brazil. BMC Cancer **13:**357.
58. **Stein AP, Saha S, Kraninger JL, Swick AD, Yu M, Lambert PF, Kimple RJ.** 2015. Prevalence of Human Papillomavirus in Oropharyngeal Cancer: A Systematic Review. Cancer J **21:**138-146.
59. **Meyers JM, Spangle JM, Munger K.** 2013. The human papillomavirus type 8 E6 protein interferes with NOTCH activation during keratinocyte differentiation. J Virol **87:**4762-4767.
60. **Tan MJ, White EA, Sowa ME, Harper JW, Aster JC, Howley PM.** 2012. Cutaneous beta-human papillomavirus E6 proteins bind Mastermind-like coactivators and repress Notch signaling. Proc Natl Acad Sci U S A **109:**E1473-1480.
61. **Bernard X, Robinson P, Nomine Y, Masson M, Charbonnier S, Ramirez-Ramos JR, Deryckere F, Trave G, Orfanoudakis G.** 2011. Proteasomal degradation of p53 by human papillomavirus E6 oncoprotein relies on the structural integrity of p53 core domain. PLoS One **6:**e25981.
62. **Simon MA, Ecsedi P, Kovacs GM, Poti AL, Remenyi A, Kardos J, Gogl G, Nyitray L.** 2019. High-throughput competitive fluorescence polarization assay reveals functional redundancy in the S100 protein family. FEBS J doi:10.1111/febs.15175.
63. **Kabsch W.** 2010. Xds. Acta Crystallogr D Biol Crystallogr **66:**125-132.
64. **Jin T, Chuenchor W, Jiang J, Cheng J, Li Y, Fang K, Huang M, Smith P, Xiao TS.** 2017. Design of an expression system to enhance MBP-mediated crystallization. Sci Rep **7:**40991.
65. **McCoy AJ, Grosse-Kunstleve RW, Adams PD, Winn MD, Storoni LC, Read RJ.** 2007. Phaser crystallographic software. J Appl Crystallogr **40:**658-674.
66. **Adams PD, Afonine PV, Bunkoczi G, Chen VB, Davis IW, Echols N, Headd JJ, Hung LW, Kapral GJ, Grosse-Kunstleve RW, McCoy AJ, Moriarty NW, Oeffner R, Read RJ, Richardson DC, Richardson JS, Terwilliger TC, Zwart PH.** 2010. PHENIX: a comprehensive Python-based system for macromolecular structure solution. Acta Crystallogr D Biol Crystallogr **66:**213-221.
67. **Emsley P, Cowtan K.** 2004. Coot: model-building tools for molecular graphics. Acta Crystallogr D Biol Crystallogr **60:**2126-2132.

68. **Wienken CJ, Baaske P, Rothbauer U, Braun D, Duhr S.** 2010. Protein-binding assays in biological liquids using microscale thermophoresis. Nat Commun **1:**100.
69. **Alexander CG, Wanner R, Johnson CM, Breitsprecher D, Winter G, Duhr S, Baaske P, Ferguson N.** 2014. Novel microscale approaches for easy, rapid determination of protein stability in academic and commercial settings. Biochim Biophys Acta **1844:**2241-2250.
70. **Mueller AM, Breitsprecher D, Duhr S, Baaske P, Schubert T, Langst G.** 2017. MicroScale Thermophoresis: A Rapid and Precise Method to Quantify Protein-Nucleic Acid Interactions in Solution. Methods Mol Biol **1654:**151-164.

Figure legends

Figure 1 - Multiple sequence alignment of E6 proteins sequences of alpha 9 HPV types and alpha 7 HPV type 18: Except HPV 67 all α-9 HPV types are classified as "high-risk". After HPV16, HPV 18 (α-7) is the second most "high-risk" HPV type. E6 proteins are ordered based on the phylogenetic relation of E6 (depicted left). Residue positions identified in the 16E6 ternary complex structure to be responsible for 16E6/p53 interaction, for 16E6/ LxxLL(E6AP) motif interaction or for both, p53 and LxxLL(E6AP), are indicated below as $, # or &, respectively. E6 domain organization and structural features based on the published structure of 16E6 (PDB 4rxn) are shown above, illustrating the two zinc-finger domains E6N and E6C, and the C-terminal PDZ Binding motif (PBM). The CxxC motifs (zinc binding) are depicted below.

ClustalX coloring scheme in Jalview, sequences from PAVE database, aligned using T-Coffee.

Figure 2 - Interaction of E6 proteins with E6AP: Purified E6 proteins show binding to a fluorescein-labeled LxxLL peptide of E6AP (fE6AP) in fluorescence anisotropy measurements in direct (A, C, E) and competitive (B, D, F; competition with unlabeled peptide) measurements. Concluding from the competitive measurements, 31E6 has a 2-fold lower affinity to the E6AP peptide than 16E6. The 31E6 analogous triple mutant of 16E6 (16E6_mut, C51T/H78W/R129G) shows a largely decreased binding to LxxLL(E6AP). (G) GPCA analyses the interaction of proteins by complementation of Gaussia-princeps split fragments *in cellulo*. 31E6 shows only 55% of the normalized luciferase ratio (NLR) compared with 16E6 indicating a lower interaction of 31E6 with E6AP protein. 16E6 L50E is a negative control. The expression of the E6 proteins was verified by Western blot analysis (H) using an antibody, which detects the luciferase split fusion (G2) of the E6 protein. Actin serves as a loading control.

Figure 3 - A comparison of protein structures of 16E6 and 31E6 with the LxxLL (E6AP) peptide: (A) A superposition of the E6 structures of 31E6/LxxLL(E6AP) in cyan, 16E6/LxxLL(E6AP)/p53 Trimer B in grey and 16E6/LxxLL(E6AP) in green shows that all E6 proteins adopt a similar overall structural fold. (B) The root-mean-square deviation (RMSD) of the Cα-atoms of the 16E6 protein structures in relation to 31E6, as an indicator of protein backbone alignment, shows that the E6C domain for the dimeric 16E6/LxxLL(E6AP) and ternary complex 16E6/LxxLL(E6AP)/p53 trimer A deviates most from 31E6-LxxLL(E6AP), whereas the 16E6 ternary complex trimer B shows very low deviation to the 31E6-LxxLL(E6AP) structure. This indicates, that 31E6-LxxLL(E6AP) adopts the conformation of 16E6/LxxLL(E6AP)/p53 trimer B. (C+D) Polar interactions between E6 proteins and the LxxLL(E6AP) peptide. MBP

and p53 molecules are omitted in the representation for clarity. Most E6 interacting amino acids are conserved (Figure 1, table 2) but the C-terminus of the LxxLL(E6AP) peptide (top) is differently organized accompanied by a loss of interaction in 31E6 with the R10 conformers and R102 compared to 16E6. (F) The 31E6/LxxLL(E6AP) structure was superimposed onto the ternary complex structure of 16E6/LxxLL(E6AP)/p53core trimer B based on the LxxLL peptide as in (A). Key amino acid differences are shown in (E) where the interaction of 16E6 residues Q14 and E18 to p53 residues N268, T102 and K101 will properly not be supported by 31E6 E14 and A18. The hydrophobic packing of the p53 core domain C-terminus is very similar in the model even tough different residues can be found in 16E6 (I23, Y43) and 31E6 (Y23, L43) which complement each other on the structural level.

Figure 4 - Interaction of 31E6 and 16E6 with p53: p53 core domain was titrated from 87.5 µM or 600 µM against fluorescently labeled 200 nM MBP-16E6-LxxLL(E6AP), MBP-31E6-LxxLL(E6AP) or MBP-31E6 A18E-LxxLL(E6AP) in two-fold serial dilutions (n=3), at 25 °C measured by microscale thermophoresis (MST). Data analysis was performed using the manufacturers MO.affinity analysis software. The fraction of formed complex was calculated at 19-20 s of thermophoresis and plotted against ligand concentration. Data points showing heterogenous fluorescence intensity were neglected. Concomitant fitting applying the softwares K_d-model yielded a K_d of 18.1±2.47 µM for 16E6-p53core and a 5-times higher K_d of 91.7±1.26 µM for 31E6-p53core. Mutating A18 in 31E6 to the 16E6 analogous E18 increased the affinity to the p53 core domain slightly to a K_d = 66.6±20.7 µM. The standard error of regression (S) is 5.5 for MBP-31E6-LxxLL(E6AP), 3.8 for MBP-31E6A18E-LxxLL(E6AP) and 5.0 for MBP-16E6-LxxLL(E6AP).

Figure 5: Representation of the used constructs and mutants with respect to their application in Fluorescence anisotropy, Microscale thermophoresis (MST), GPCA and crystallization: 31E6 has two, 16E6 has four surface exposed cysteines, which were mutated to alanines (lines). The dashed line in 16E6 indicates the F47R mutation, which decreases E6 oligomerization. Since this mutation abolishes p53 interaction, it was not applied for p53 interaction studies using MST. Fusion tags (MBP or luciferase fragments), linkers, proteins of interest and fused peptide ligands are colored in orange or cyan, blue, lily and magenta, respectively.

844 **Table 1 data collection and refinement statistics (highest-resolution shell is**
845 **shown in parentheses)**

	MBP-31E6-LxxLL(E6AP)
Wavelength	1.0
Resolution range (Å)	49.21 - 2.8 (2.9 - 2.8)
Space group	P 61 2 2
Unit cell (Å/°)	113.64 113.64 185.97 90 90 120
Total reflections	350689 (33992)
Unique reflections	18044 (1756)
Multiplicity	19.4 (19.3)
Completeness (%)	99.34 (98.82)
Mean I/sigma (I)	14.71 (1.60)
R-meas	0.2318 (2.025)
CC1/2	0.998 (0.599)
R-work	0.2308 (0.3296)
R-free	0.2698 (0.3355)
Number of non-hydrogen atoms	4281
macromolecules	4213
ligands	37
solvent	31
Protein residues	539
RMS (bonds)	0.003
RMS (angles)	0.62
Ramachandran favored (%)	96.07
Ramachandran allowed (%)	3.93
Ramachandran outliers (%)	0
Rotamer outliers (%)	2.53
Clashscore	6.76
Average B-factor	85.21
macromolecules	85.56
ligands	67.19
solvent	59.1

846

847 **Table 2 conservation of amino acids of 16E6 which participate in LxxLL(E6AP)**
848 **binding (22, 28) – variations between 16E6 and 31E6 are highlighted in**
849 **grey/bold, * mutations show a significant decrease in binding**

16E6	R10	K11	L50*	C51	R55	S74	R77	H78	R129	R102*	R131*
31E6	R10	K11	L50*	**T51**	R55	S74	R77	**W78**	**G129**	R102	R131

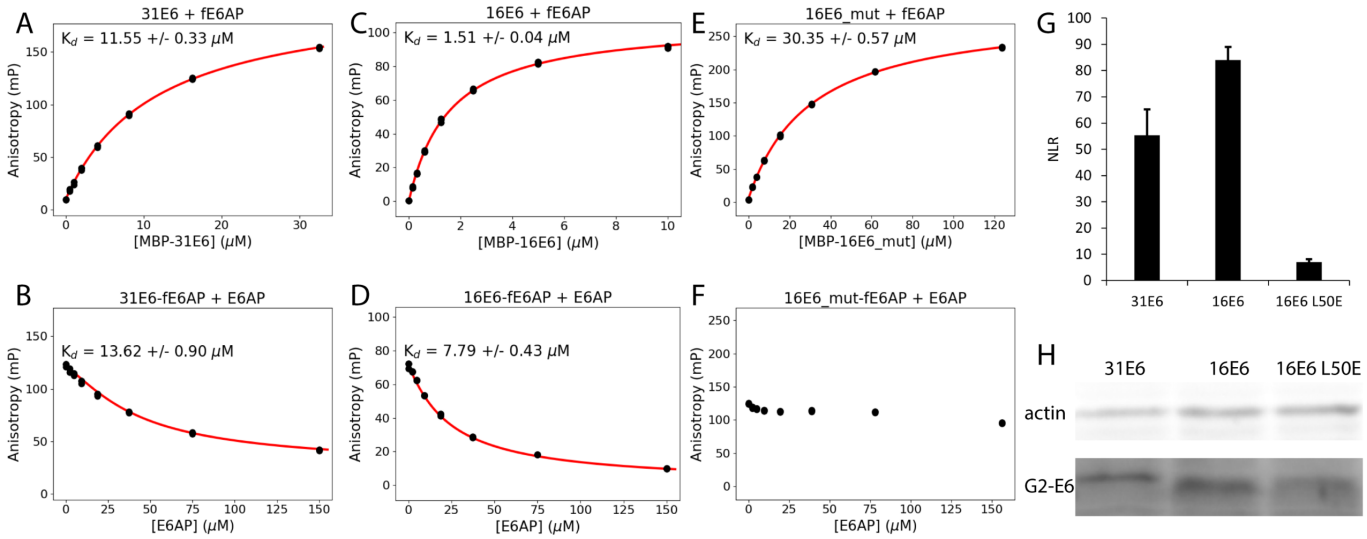
850

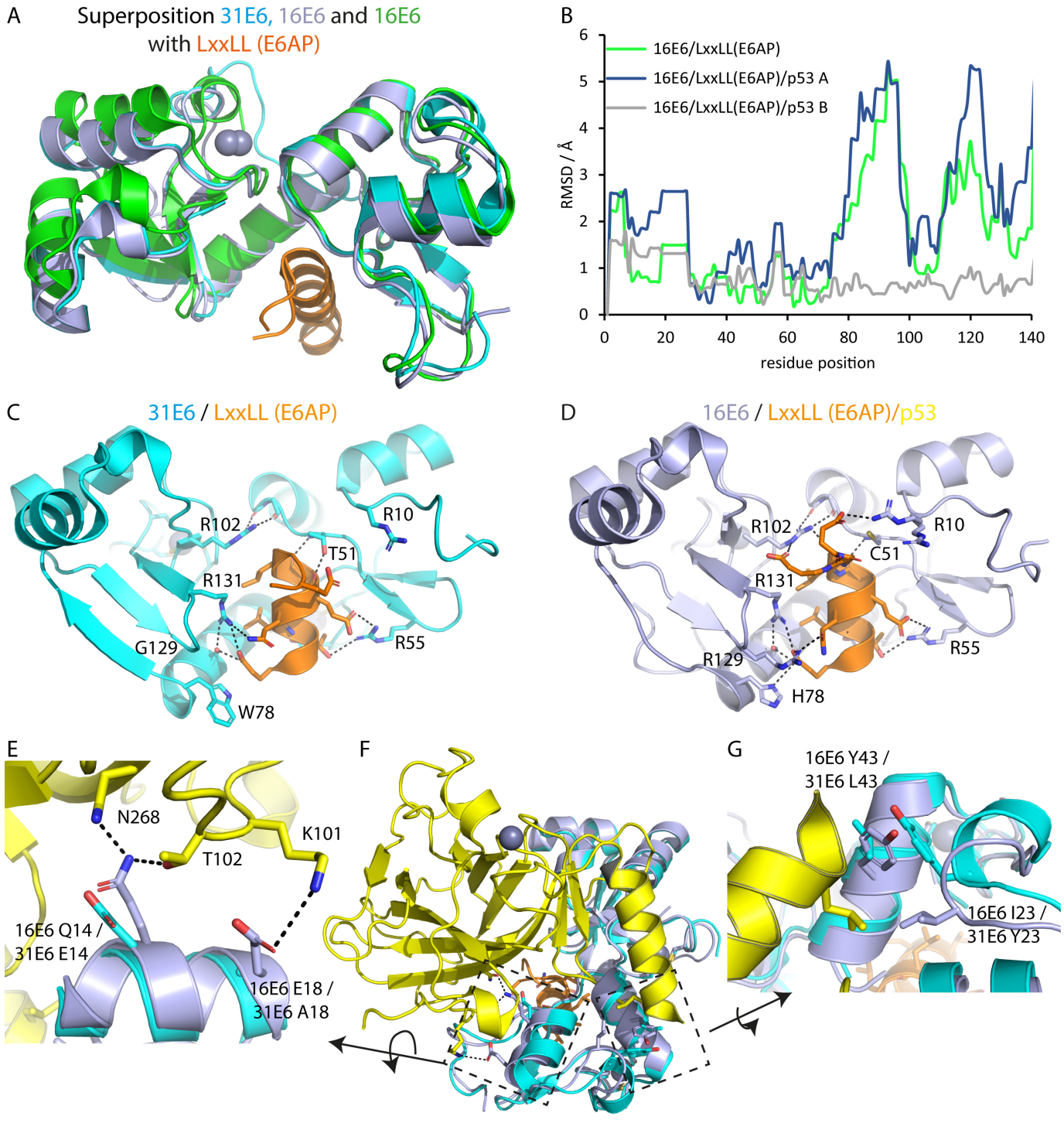
851 **Table 3 conservation of amino acids of 16E6 important for p53 binding (17),**
852 **variations between 16E6 and 31E6 are highlighted in grey/bold, * mutations**
853 **show a significant decrease in binding**

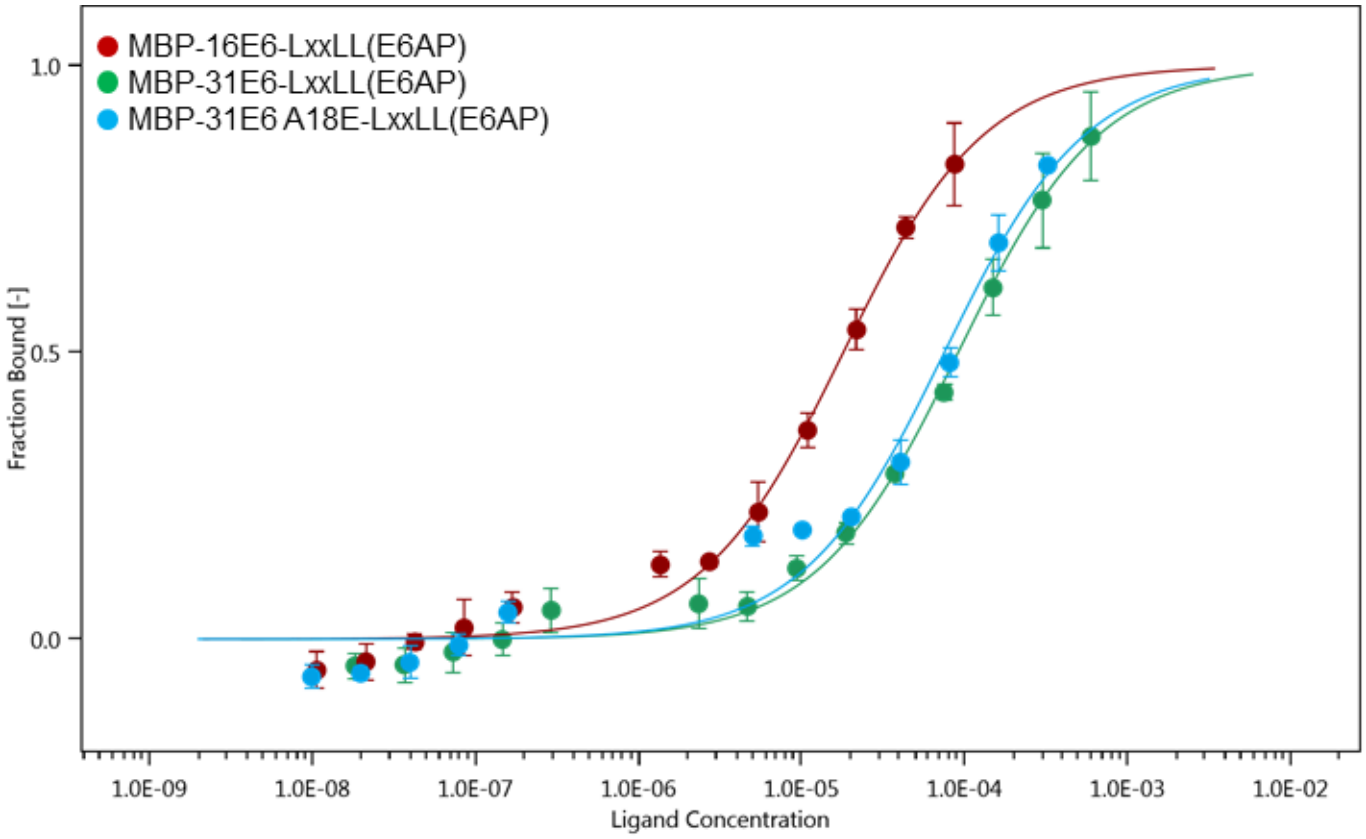
16E6	Q6	E7	R8	Q14	E18*	Y43	D44*	F47*	D49*	L100	P112
31E6	**A6**	E7	R8	**E14**	**A18**	**L43**	D44*	F47*	D49*	L100	P112

854

855







Journal of Virology