

Thèse

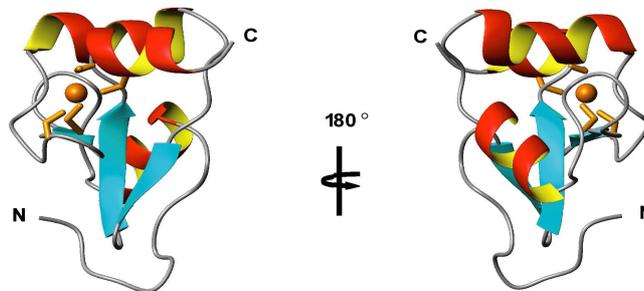
Présentée à l'École Doctorale des Sciences de la Vie et de la Santé
de L'Université de Strasbourg pour obtenir le grade de

Docteur de l'Université de Strasbourg

par

Abdellahi OULD M'HAMED OULD SIDI

*Production, analyse structurale et caractérisation
fonctionnelle d'oncoprotéines E6 issues de différents
papillomavirus*



Soutenue publiquement le 18.06.2009

Membres du jury :

Directeur de Thèse : Dr Gilles Travé

ESBS, Strasbourg

Rapporteur Interne : Prof Etienne WEISS

ESBS, Strasbourg

Rapporteur Externe : Prof Angel ALONSO

DKFZ, Heidelberg

Rapporteur Externe : Dr Nicolas WOLFF

I. Pasteur, Paris

Examineur : Dr Georges ORFANOUDAKIS

ESBS, Strasbourg

Examineur : Dr Yves JACOB

I. Pasteur, Paris

REMERCIEMENTS

*Les recherches qui font l'objet de ce mémoire ont été menées au sein de l'IREBS, FRE 3211 dirigée par **Dr Jean-luc Galzi**, dans l'équipe Oncoprotéines, dirigée par le **Dr. Georges Orfanoudakis** et **Dr Gilles Travé**.*

Qu'il me soit permis tout d'abord de remercier vivement ceux qui m'ont guidé et soutenu pour me permettre de mener à bien ce travail pendant ces années de thèse.

*En premier lieu, je tien à exprimer ma plus profonde gratitude à **Dr Gilles Travé** qui m'a proposé ce sujet de thèse et m'a accompagné tout au long de sa réalisation avec beaucoup d'intérêt. **Gilles**, j'ai apprécié chez toi la qualité d'un grand chercheur plein d'optimisme, ton sens très développé de la rigueur et tes qualités humaines: esprit très ouvert, sympathie couronnée d'une énorme modestie, et soutien. Je te remercie pour la confiance précieuse que tu m'as accordé, pour les conseils, l'assistance morale et l'encadrement sans faille que tu as assuré tout au long de mes années de formation et de recherches. Merci pour m'avoir aidé, soutenu et facilité mon intégration depuis mon arrivée en France. Ma mémoire est encore fraîche, il était minuit à la gare de Strasbourg, le temps faisait 4°C, un record de froid pour moi!, quand tu m'as accueilli généreusement. Ce fut notre première rencontre, et depuis, cette générosité (record elle aussi !) n'a jamais fléchi... Que tu trouves ici le témoignage de ma reconnaissance pour toi et ta petite famille.*

Je remercie l'Association pour la Recherche sur le Cancer (ARC) pour m'avoir accordé une troisième année de financement et la Société Française de Biochimie et Biologie Moléculaire (SFBBM) pour avoir financé ma participation au 33^{ème} congrès de FEBS à Athènes (Grèce).

*Je tiens également à remercier : **Prof Etienne WEISS**, **Prof Angel ALONSO**, **Dr Nicolas WOLFF**, **Dr Yves JACOB** et **Dr Georges ORFANOUDAKIS**, qui ont bien voulu me faire l'honneur de juger ce travail, en acceptant malgré leurs nombreuses activités, d'être rapporteurs et examinateurs de ce travail, et en dépit du retard inopportun de l'envoi de mon document dont je m'excuse au passage.*

*Je voudrais remercier **Dr Katia ZANTER**, ce travail n'aurait pas atteint son ampleur sans toi, merci infiniment pour ta gentillesse. Nos discussions parfois animées ont toujours été constructives et vraies. Tes connaissances, ta participation active, tes critiques et tes conseils, ta patience et ta disponibilité durant tout le temps passé ensemble ont énormément inspiré mon évolution professionnelle et humaine.*

Mes remerciements vont également à **Dr Georges ORFANOUDAKIS, Dr Murielle MASSON, Dr François DERYCKERE, Dr Yves NOMINE**, Je vous suis reconnaissant pour toutes les discussions et les échanges enrichissants que nous avons eus pendant toute la durée de ma thèse. J'ai particulièrement apprécié vos dynamismes, et vos grandes réactivités face à toutes mes interrogations et tous mes problèmes. Merci Murielle particulièrement pour la correction d'une partie de ce manuscrit malgré ton séjour à l'hôpital.

Mes remerciements s'étendent à l'ensemble de mes amis et collègues du laboratoire, les présents comme ceux qui sont déjà partis, pour la convivialité et la bonne humeur. Il me faudrait plusieurs pages pour les citer tous, mais j'aimerais mentionner spécialement : le petit champion « Barcelonais » **Sadek** je te dis simplement « ありがとうございました » pour ta contribution à la préparation de ce manuscrit. Le jeune cristallographe **Dr Sébastien**, l'estimation que j'ai eu pour toi dès le début, n'est que la diffraction de ta présence éclaircissante dans l'équipe. **Charlotte**, merci pour ta contribution bien fructueuse pour nos aventures de thèses. Merci de tout Core+1 à **Anissa**, et je te dis « choukrann » avec **Khaled**, le prince héritier de ma paillasse. **Katja** (bioinfo), **Anne, Xavier, Anne So, Audrey, Rim, Ahmed** et **kambiz** un grand merci pour vos discussions divertissantes qui allègent les moments difficiles dans le labo.

Je remercie infiniment mon professeur à Nouakchott **Prof Ahmedou Ould Houmeida** pour son soutien de très longue durée, depuis mes premières années à l'université, et qui ne fait qu'augmenter au fur et à mesure. Sans lui ne je ne serais jamais arrivé à ce niveau.

Cela va de soi, je remercie indéfectiblement **mes parents, mes sœurs, frères et amis** au pays. Ils ont été présents malgré l'éloignement géographique, pour écarter les doutes, soigner les blessures et partager les joies. Que Dieu vous protège et vous accorde chacun (e) une longue vie. Cette thèse est un peu la vôtre aussi.

Également, je tien à remercier la petite colonie mauritanienne à Strasbourg, **Khaled Ould Kar, Sid'ahmed Chouaib, Moulay ould Hamouni, Khaled bahah, Ahmed Ould Nouman** et **Hindou Mint Zein**. C'était un immense plaisir de me retrouver entre vous.

ICI, TOUT SPECIALEMENT, AVEC AMOUR ET AFFECTION, JE VOUDRAI ASSOCIER A CETTE THÈSE MON EPOUSE **HOUDA MINT BABAH**, SANS L'AIDE DE QUI, JE N'AURAIS PAS PU MENER A BIEN CE TRAVAIL. MON CŒUR ET MOI NE POURRONS JAMAIS TE REMERCIER ASSEZ POUR CE QUE TU REPRESENTES POUR NOUS. MERCI A TOI **HOUDA**. ET BARAKA ALLAHOU VIK.

ABREVIATION

aa amino acid

ADN Acide Désoxyribonucléique

AP1 Activator Protein 1

ARN Acide Ribonucléique

ARNm ARN messenger

ATP Adénosine 5'-triphosphate

BPV Bovine Papillomavirus

Brd4 Bromodomain containing protein 4

CBP CREB Binding Protein

cdk cyclin dependent kinase

CDK Cyclin Dependent Kinase

CDKI Cyclin-Dependent Kinase Inhibitor

CIN Cervical Intraepithelial Neoplasia

CKI cdk Inhibitor

CMH Complexe Majeur d'Histocompatibilité

CR Conserved Region

CRPV Cottontail Rabbit Papillomavirus

Cys cystéine

DBD DNA Binding Domain

DNA DesoxyriboNucleic Acid

DO Densité Optique

DPV Deer Papillomavirus

E Early region

E6-AP E6-Associated Protein

E6BP E6 Binding Protein

EDTA Ethylendiamintetraacetic acid

EGF Epidermal Growth Factor

EV Epidermodysplasia Verruciformis

FIPV Francolinus leucoscepus papillomavirus

GADD Growth Arrest on DNA Damage

GST Glutathione-S-Transferase

hDlg human homolog of the Drosophila melanogaster tumor suppressor

HLA Human Leucocyte Antigen

HPV Human Papillomavirus

H-ras Harvey rat sarcoma viral oncogene homolog

hScrib human homolog of the *D. melanogaster* tumor suppressor Scribble
HSQC Heteronuclear Single Quantum Coherence
hTERT human Telomerase Reverse Transcriptase
IARC International Agency for Research on Cancer
IFN Interféron
IPTG Isopropyl-1- β -D-thiogalactopyranoside
IRF-3 Interferon Regulatory Factor 3
kDa kilo-Dalton
Kpb kilo paire de base
L Late region
LCR Long Control Region
M "molar" = 1 mol/L
MAGI Membrane Associated Guanylate kinase with Inverted domains
MAGUK Membrane Associated Guanylate Kinase
MBP Maltose Binding Protein
MDM2 Murine Double Minute 2
MUPP1 Multi PDZ domain Protein
NF- κ B Nuclear Factor kappa B
NLS Nuclear Localisation Signal
NMR Nuclear Magnetic Resonance
OMS Organisation Mondiale de la Santé
ORF Open Reading Frame
ori origine de réplication
OVPV Ovine Papillomavirus
PAGE Polyacrylamide Gel Electrophoresis
PCR Polymerase Chain Reaction
PDGF Platelet Derived Growth Factor
POL Phase Ouvert de Lecture
pRb protéine du Rétinoblastome
PV Papillomavirus
Rb Retinoblastoma protein
rTEV recombinant Tobacco Etch Virus protease
SDS Sodium dodecyl sulphate
siRNA small interfering RNA
Sp1 a GC box binding protein
SPR Surface Plasmon Resonance
TEV Tobacco Etch Virus

TGF Transforming Growth Factor
TNF Tumour Necrosis Factor
URR Upstream Regulatory Region
VIH Virus de l'Immunodéficience Humaine
YY1 protéine ying-yang 1
 β -IFN interféron bêta

ACIDES AMINES

A Ala Alanine
C Cys Cysteine
D Asp Aspartic acid
E Glu Glutamic acid
F Phe Phenylalanine
G Gly Glycine
H His Histidine
I Ile Isoleucine
K Lys Lysine
L Leu Leucine
M Met Methionine
N Asn Asparagine
P Pro Proline
Q Gln Glutamine
R Arg Arginine
S Ser Serine
T Thr Threonine
V Val Valine
W Trp Tryptophane
Y Tyr Tyrosine

SOMMAIRE

INTRODUCTION GENERALE	1
Les Papillomavirus	1
Classification et phylogénie des Papillomavirus	2
L'évolution de papillomavirus	5
Structure du virus	7
L'organisation génomique des papillomavirus humains	8
Cycle de réplication du virus HPV	10
Pénétration du virus	10
Infection virale productive	11
Infection latente	11
L'intégration de l'ADN viral au génome cellulaire	11
Pouvoirs pathogènes des HPV	13
Histoire naturelle de l'infection à HPV	16
Cancer du col de l'utérus	16
Facteur étiologique principal	18
Facteurs étiologiques secondaires (cofacteurs)	20
Dépistage, traitement et la prophylaxie	22
Les protéines virales	23
La protéine L1	23
La protéine L2	24
La protéine E1	26
La protéine E2	28
La protéine E4	32
La protéine E5	32
Les oncoprotéines E7 et E6	33
La protéine E7	34

L'activité de E7 et ses cibles cellulaires -----	34
Caractérisations biochimiques de E7 -----	38
La protéine E6 -----	41
La protéine E6, ses fonctions et ses cibles -----	41
L'interaction E6, p53 et E6AP -----	42
Les autres cibles de E6 -----	44
E6 et Bak -----	44
E6 et Paxilline -----	44
E6 et les protéines à domaines PDZ -----	44
E6 et E6-BP -----	46
E6 et TNF-1 -----	46
E6 et E6TP-1 -----	46
E6 et c-Myc -----	46
E6 et la télomérase (hTERT) -----	47
E6 et acides nucléiques: ADN et ARN -----	47
Conclusion sur les fonctions complémentaires de E6 et E7 -----	49
Caractérisations biochimiques de E6 -----	49
Production et Purification de E6 -----	50
Tentatives échecs et d'expression de la protéine -----	50
La purification de E6 HPV16 soluble -----	51
La prédiction de la structure secondaire des protéines E6 -----	52
Structure de la protéine E6 -----	53
Présentation du projet de recherche -----	56
RESULTATS -----	58
Première partie : Production et caractérisation de HPV16 E6 -----	58
La purification du domaine E6-N hautement soluble -----	58

Identification d'une interface de dimérisation dans le domaine E6-N -----	60
Production et caractérisation du domaine C-terminal de HPV E6 -----	62
Production et caractérisation de HPV 16 E6 entière -----	64
Interaction de HPV16 E6 avec l'ubiquitine ligase E6-AP -----	67
HPV16 E6, liaison E6-AP et dégradation de p53 -----	67
Deuxième partie : Caractérisation biophysique de la protéine E6 de BPV -----	71
Manuscrit 1 -----	74
Troisième partie : Identification et caractérisation des protéines précoces E6, E7 et E9 issues de papillomavirus aviaires -----	95
Manuscrit 2 -----	98
Quatrième partie: l'auto-association des protéines E6 -----	119
Manuscrit 3 -----	120
DISCUSSION GENERALE -----	141
CONCLUSIONS ET PERSPECTIVES -----	149
MATERIELS ET METHODES -----	154
REFERENCES BIBLIOGRAPHIQUES -----	161

INTRODUCTION GENERALE

INTRODUCTION GENERALE

Les Papillomavirus

Classification et phylogénie des Papillomavirus

Les papillomavirus (PVs) sont des virus extrêmement anciens. Le terme Papillomavirus vient du latin *papilla*, qui signifie bouton et du suffixe *ome* pour signifier le caractère tumoral. La famille des *Papillomaviridae* comporte de très nombreux virus infectants divers espèces, ils sont d'ailleurs identifiés par leur hôte naturel : HPV (humain), BPV (bovin) ou encore DPV (deer, cervidés) et tout récemment CmPV (tortue) (Groff et lancaster, 1985; Herbst *et al.*, 2009). Jusqu'à récemment, les papillomavirus étaient regroupés avec les polyomavirus pour former la famille des *Papovaviridae*. Le terme Papovavirus dérive des deux premières lettres des premiers virus constituant cette famille: Papillomavirus du lapin, polyomavirus de souris et virus vacuolisant du singe connu sous le nom de SV40. Cependant, la différence de taille du virion et de longueur de leur ADN ainsi que l'absence de similitude dans la comparaison des génomes ont amené les chercheurs à diviser les Papovaviridae en deux familles distinctes: les *Polyomaviridae* regroupant les polyomavirus et SV40, et les *Papillomaviridae* couvrant l'ensemble des papillomavirus.

Les dernières années ont révélé un nombre croissant de papillomavirus (PVs), dont les génomes ont été séquencés et répertoriés dans la base de données GenBank, totalisant près de 200 souches de virus. Le génome des PVs est contenu dans une molécule d'ADN circulaire double brin d'environ 8000 paires de bases dont un seul brin est codant, avec des phases ouvertes de lecture (ORF, *open reading frame*) qui se chevauchent. Par définition, les gènes E1 à E7, indispensables à la transcription et à la réplication virale, font partie de la région dite précoce, par opposition aux gènes de capsid L1 et L2 appelés tardifs. La protéine issue du gène L1, entière ou sous forme de fragments, est communément utilisée pour détecter les infections ou pour le typage des PVs. Contrairement à la plupart des virus, il n'existe pas une classification sérotypique pour les PVs. Pour cette raison la taxonomie des PVs repose sur des identités de séquences c'est-à-dire une classification génotypique. Cette classification est traditionnellement basée sur l'étude de gènes structuraux, tout particulièrement le gène L1. Pour qu'un nouveau type de PV soit reconnu, il faut que le génome complet du virus ait été séquencé et que sa séquence L1 présente une divergence de plus de 10 % avec la séquence L1 du type connu le plus proche génétiquement. Une différence de 2 à 10 % définit l'appartenance à un sous-type et une différence de moins de 2 % définit un variant (Kämmer *et al.*, 2002, Cruz *et al.*, 2004 ; de Villiers *et al.*, 2004)

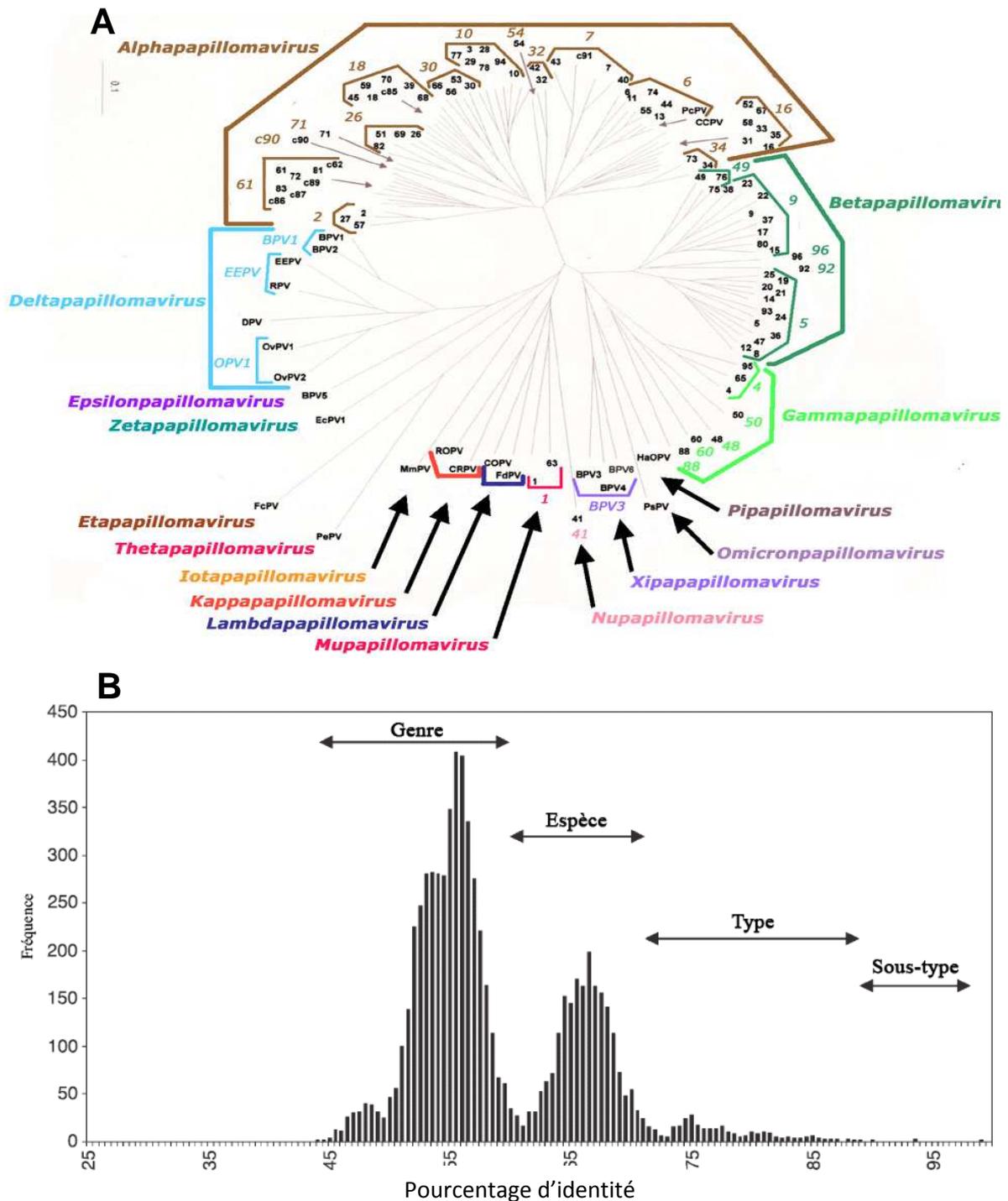


Figure 1 : Phylogénie des papillomavirus. (A) Arbre phylogénétique représentant la distance relative séparant la séquence L1 de 118 papillomavirus. Les différents genres viraux sont indiqués en couleur, le regroupement des souches virales en espèces est indiqué par un trait de couleur plus fin. Les souches virales représentées seulement par un numéro correspondent aux papillomavirus humains (ex : 16 = HPV16). (B) Pourcentage d'identité de la séquence L1 de ces 118 papillomavirus par genre, espèce, type et sous-type (D'après de Villiers, 2004).

Cette classification taxonomique traditionnelle est néanmoins imparfaite. C'est pourquoi l'*International Committee on the Taxonomy of Viruses* (ICTV) a établi une nouvelle classification taxonomique avec trois objectifs:

- cette classification doit établir la relation existant entre les types de PVs.
- elle doit intégrer les termes de « genre » et « d'espèce » communément utilisés dans la classification systématique de tous les organismes biologiques et fréquemment appliqués aux virus.
- elle doit intégrer la relation entre la classification taxonomique et les propriétés biologiques et pathologiques des virus (de Villiers *et al.*, 2004).

La comparaison des séquences du gène L1 identifie clairement 3 niveaux taxonomiques : le genre, l'espèce et le type (Figure 1) qui sont définis ainsi:

- les différents genres partagent moins de 60% de leur séquence d'acides nucléiques.
- les espèces au sein d'un même genre partagent entre 60% et 70% de leur séquence nucléique.
- les traditionnels types des PVs au sein d'une même espèce partagent entre 71% et 89% de leur séquence d'acides nucléiques.

Dans cette nouvelle classification, les définitions de sous-type et de variant restent inchangées. Par exemple, parmi les genres nouvellement déterminés, le genre alpha-papillomavirus est le plus important et compte des HPV muqueux et cutanés. Le genre bêtapapillomavirus inclue les HPV cutanés associés à l'épidermodysplasie verruciforme et le genre gamma compte exclusivement des HPV cutanés. Le genre deltapapillomavirus est également important quantitativement et ne comporte pas de PV infectant l'homme, mais le cerf (DPV), les ovins (OVPV) et les bovins (BPV). Enfin, les autres genres de papillomavirus regroupent les PV infectant d'autres animaux ainsi que des HPV rares (de Villiers 2004).

Les PVs, isolés à partir de différentes espèces de mammifères, d'oiseaux ou de reptiles, sont généralement considérés comme des virus très spécifiques à leurs hôtes. Cependant quelques espèces de papillomavirus bovins (BPV) sont responsables d'infections non spécifiques chez les chevaux ou d'autres espèces relativement éloignées (Chambers *et al.*, 2003). Au-delà de leur spécificité d'espèce (hôte), les PVs présentent aussi une forte

spécificité ou tropisme tissulaire. Les PVs sont strictement épithéliotropes et sont capables d'infecter soit les épithéliums cutanés, soit les épithéliums muqueux.

Au milieu des années 70s, les PVs et particulièrement les papillomavirus humains font l'objet de nombreuses études pertinentes (zur Hausen 1977, Orth *et al.*, 1977). Ces travaux ont confirmé la remarquable pluralité des HPV. Ces virus sont classés le plus souvent en fonction de leur tropisme tissulaire et de leur potentiel oncogénique. Deux grands groupes de HPV se distinguent:

- Les HPV préférentiellement associés aux lésions cutanées. Les HPV de type 1 et 4 par exemple sont fréquemment retrouvés dans les verrues, alors que les HPV de type 5 et 8 sont incriminés dans l'épidermodysplasie verruciforme (une maladie héréditaire rare à transmission autosomique récessive conduisant à long terme à des carcinomes cutanés précoces, souvent considérée comme un modèle de carcinose virale chez l'homme (Orth *et al.*, 2001).

- les HPV infectant les muqueuses anogénitales (col utérin, vulve, vagin, pénis et anus) et oropharyngées. Parmi ces virus, certains sont dits à "bas risque" ou à faible potentiel oncogène : c'est le cas des HPV6 et 11 communément retrouvés dans les condylomes génitaux, alors que d'autres sont dits à "haut risque" : c'est le cas des HPV16 et 18 impliqués dans la carcinogenèse du col utérin. Dans ce dernier groupe, sont aussi inclus des HPV dits à risque intermédiaire: il s'agit des HPV31, 33, 35, 51... fréquemment détectés dans les lésions anogénitales (Van Ranst *et al.*, 1992).

L'évolution des papillomavirus

Plusieurs observations suggèrent que la vitesse du processus d'évolution des PVs est très lente, en raison des circonstances suivantes :

- les PVs sont des virus à ADN et utilisent les mêmes mécanismes enzymatiques que leurs hôtes pour la réplication;

- la réplication elle-même est aussi lente parce qu'elle dépend de la division des cellules épithéliales infectées, avec un taux de division semblable. L'on peut donc soupçonner que la diversification est corrélée entre les virus et leurs hôtes, suivant un processus dit de coévolution (Ong *et al.*, 1992; Chan *et al.*, 1997).

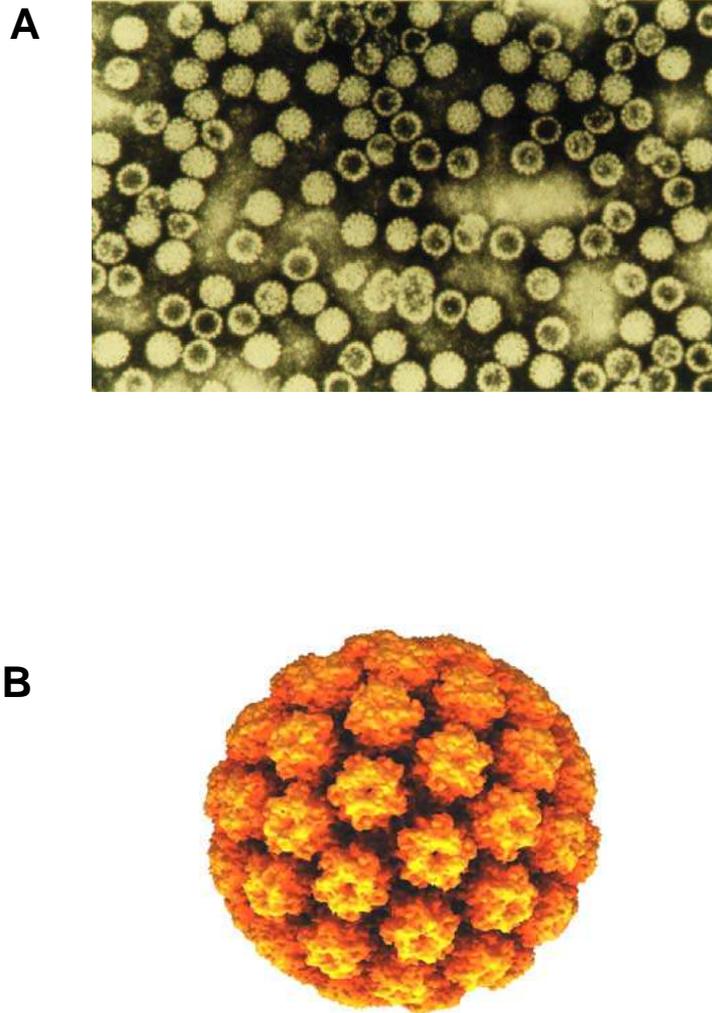


Figure 2: Le papillomavirus. (A) Particules virales, visibles en microscope électronique. **(B)** Reconstruction par cryo-électromicroscopie d'une capsid d'un papillomavirus humain (Modis *et al.*, 2002).

Jusqu'ici, la co-évolution entre les papillomavirus et leurs hôtes a été tacitement admise par la plupart des auteurs, malgré que certains aient également avancé des origines multiples pour certaines souches de PVs (Garcia-Vallvé *et al.*, 2005). Cependant, la théorie de la coévolution présente plusieurs lacunes. Par exemple, la coévolution exige que tous les PVs infectant un hôte donné soient regroupés dans la classification et partageant un ancêtre commun récent. Dans un tel cas, les rapports phylogénétiques entre les virus doivent pouvoir être reconstruits littéralement d'après les rapports phylogénétiques entre leurs hôtes. Or, tel n'est pas le cas: les PVs infectant les primates ne sont pas monophylétiques et ont au moins quatre ancêtres différents (Gottschling *et al.*, 2007 ; Chan *et al.*, 1995). Ces virus ancestraux auraient donné lieu séparément aux différents genres (alpha, bêtas, gamma, mu et nupapillomavirus). Un scénario semblable est observé pour les PVs infectant les artiodactyles, avec au moins deux ancêtres, un pour les delta PVs et un autre pour différents papillomavirus bovins rares (BPV) (Chan *et al.*, 1995). En outre, l'on peut postuler l'existence d'un papillomavirus ancestral commun pour les genres kappa, lambda et mupapillomavirus, qui infectent les lagomorphes, les carnivores et les primates, respectivement. Cependant, l'ancêtre commun le plus récent pour ces trois ordres de mammifères a vécu il y a environ 92 millions d'années, et il est également commun pour beaucoup d'autres ordres, tels que les artiodactyles, les Cétacés, l'ordre des siréniens et encore les périssodactyles (Kumar et Hedges, 1998).

Reste à préciser que les carnivores et les périssodactyles ont partagé un ancêtre commun récent, les cétacés et les artiodactyles ont tous les deux également un ancêtre commun, mais ce n'est pas le cas pour les papillomavirus respectifs qui les infectent (Garcia-Vallvé *et al.*, 2005).

Structure du virus

Les papillomavirus sont des virus non enveloppés de petite taille et à cycle non lytique. Ils possèdent une capsidite d'environ 600 Å de diamètre et contenant 72 capsomères de la protéine structurale L1 en association avec 12 copies de la protéine L2 par virion (Baker *et al.*, 1991). Malgré les différences de séquences pouvant exister entre les différents genres de PVs, l'organisation de la capsidite et du génome leur est commune. L'organisation des capsomères permet la formation d'une capsidite adoptant une structure proche d'une symétrie icosaédrique (Figure 2). La protéine L1 ainsi appelée protéine majeure de la capsidite parce qu'elle est capable à elle seule de reconstituer une capsidite synthétique identique à la

capside naturelle du virus (Kirnbauer *et al.*, 1992; Bishop *et al.*, 2007). La structure atomique de la protéine L1 a été déterminée par cristallographie et un modèle pour la structure de la capsid e a été proposé (Modis *et al.*, 2002). D'après ce modèle, l'assemblage des capsomères se fait grâce à des liaisons ioniques mais aussi par la création de ponts disulfures intermoléculaires entre l'extrémité C-terminale des protéines L1 d'un capsomère avec l'autre capsomère opposé. Cependant, la protéine L2, dite protéine mineure de capsid e, joue un rôle secondaire dans l'assemblage du virion mais aussi un rôle facilitateur pour l'infection de la cellule hôte (Buck *et al.*, 2008).

L'organisation génomique des papillomavirus humains

Le génome des HPV contient une dizaine de phases ouvertes de lecture ORF (Open Reading Fragment) qui sont transcrits en ARN polycistroniques à partir de l'un des deux brins de l'ADN circulaire du virus. Les transcrits viraux suivent un processus complexe de différenciation à travers des mécanismes alternatifs d'épissages avant la traduction (Zheng et Baker., 2006).

Le génome du HPV contient deux promoteurs majeurs. Le promoteur précoce initie la transcription en amont de la région ORF (phase de lecture ouverte) de E6, alors que le promoteur tardif se trouve dans la région ORF de E7 et s'active durant la phase tardive du cycle de vie viral. Par exemple, pour HPV16, les gènes viraux codant pour les protéines précoces E6, E7, E1 et E2 sont transcrits à l'aide du promoteur précoce p97. Durant la phase productive du cycle de vie virale le promoteur tardif p670 est activé ayant pour effet une forte transcription des protéines E1, E2, E4, E5 ainsi que les protéines de la capsid e L1 et L2. La transcription des ORF précoces est régulée par des éléments situés dans la région non codante nommée LCR (*Long Control Region*) à proximité de l'ORF de E6 (Figure 3). La région non codante appelée URR (*Upstream Regulatory Region*) s'étend sur environ 1Kb. C'est une longue région de contrôle jouant un rôle important dans la promotion de la transcription virale. Elle contient en plus du promoteur précoce des séquences régulatrices qui sont situées en cis en amont de ces promoteurs et contiennent des sites de fixation spécifiques de facteurs trans-régulateurs d'origine cellulaire ou virale. Parmi les facteurs cellulaires, l'on retrouve les facteurs AP1, NF1, Sp1 et les récepteurs aux stéroïdes qui activent la transcription des gènes précoces, alors que les facteurs p53, YY1, Oct-1 et les récepteurs de l'acide rétinoïque l'inhibent.

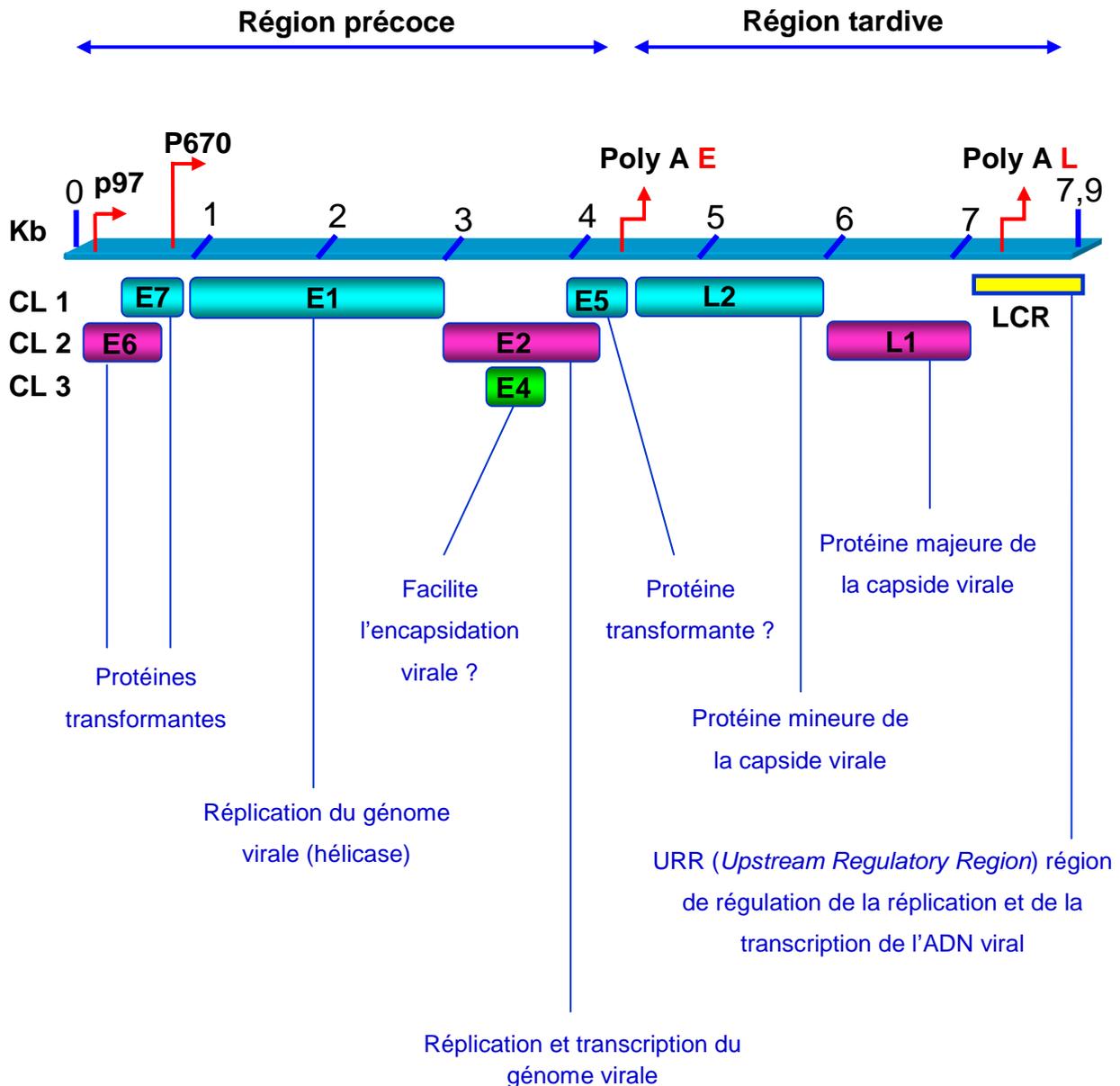


Figure 3: Le génome viral des papillomavirus. Le génome de papillomavirus est fortement conservé à travers les types. Les ORFs E6, E7, E1, E2, E4 et E5 codent les protéines nécessaires à la réplication de l'ADN et l'expression des gènes viraux, les ORFs L1 et L2 codent les protéines virales de la capside. Un seul brin de l'ADN bicaténaire sert à la production d'un ARNm polycistronique. La transcription est contrôlée par des éléments « enhancer » situés dans la région non codante LCR. L'initiation de la transcription au niveau des promoteurs précoces et tardifs dépend de chaque type de virus (ici, l'exemple est le HPV16), mais pour l'ensemble des types, l'usage des promoteurs est lié au degré de différenciation des kératinocytes. (d'après Stanley et al., 2007).

Ces facteurs présentent une spécificité cellulaire ou tissulaire qui contribue au tropisme des papillomavirus (Kanaya *et al.*, 1997; Kisseljev, 2000; Morris *et al.*, 1993).

La région LCR contient également des séquences spécifiques pour la fixation des protéines virales E1 et E2 qui sont impliquées dans l'initiation de la régulation et sont connus aussi comme des régulateurs de la transcription pour E6 et E7.

Cycle de réplication du virus HPV

Pénétration du virus :

Lors d'une infection naturelle, les HPV pénètrent dans l'épithélium spécifiquement au niveau des cellules basales (cellules peu différenciées en division) à la suite d'un microtraumatisme ou d'une lésion tissulaire. Cette zone de transformation est favorable pour l'infection virale, car c'est une zone de fragilité à la fois mécanique et immunologique. La faible densité des cellules de Langerhans au niveau de cet épithélium pourrait être responsable d'une immunodéficience locale, permettant aux HPV d'infecter aisément cette zone de transformation. Cependant, ces virus ne pourront pénétrer dans les cellules basales qu'à la faveur d'une lésion tissulaire ou d'un microtraumatisme. Les molécules mises en jeu lors de l'interaction avec la cellule cible sont encore peu connues. L'adsorption du virus se ferait via une interaction avec des héparines sulfates, interaction qui serait suivie d'une liaison avec les « intégrines » $\alpha 6 \beta 4$ permettant l'internalisation des virions par endocytose au niveau des puits recouverts de clathrine. Une fois dans le cytoplasme, le virus migre activement vers le noyau cellulaire grâce au réseau de microtubules et de filaments d'actine du cytosquelette (Selinka *et al.*, 2002). Pour profiter de la machinerie cellulaire, le virus doit alors passer la membrane nucléaire. La capsid virale étant plus grande que le diamètre des pores nucléaires, elle est démantelée (Nelson *et al.*, 2002). La décapsidation serait facilitée par la rupture des ponts disulfures intra-capsomériques dans l'environnement réducteur de la cellule, et le transport de l'ADN viral dans le noyau se fait grâce aux séquences de localisation nucléaire présentes sur les protéines L1 et L2. Une fois dans le cytoplasme, le virus migre activement vers le noyau cellulaire grâce au réseau de microtubules et de filament d'actine du cytosquelette (Selinka *et al.*, 2002).

L'ADN peut se répliquer sous forme épisomale, dans une infection productive ou bien persister à l'état latent. Enfin, certaines parties de l'ADN viral peuvent s'intégrer au génome de la cellule hôte déclenchant une infection non productive mais pathogénique.

Infection virale productive avec ADN virale extra-chromosomique :

C'est dans les cellules basales que les gènes précoces E1, E2, E6, E7 sont exprimés (Wilson *et al.*, 2002) et le génome viral persiste dans le noyau sous la forme d'éléments extra-chromosomiques (50 à 100 copies par cellule). Le nombre de copies virales peut alors être maintenu quasi constant pendant de nombreuses années (Stubenrauch et Laimins, 1999). Après les divisions cellulaires, les cellules filles infectées quittent la couche basale et entament un processus de différenciation lors de leur migration vers les régions supra-basales. Il est important de souligner que l'expression des gènes « oncogènes » E6 et E7 de HPV est très faible dans les cellules différenciées, car ils sont contrôlés négativement par la protéine E2 (action trans-inhibitrice de E2). Contrairement aux cellules non-infectées où le cycle cellulaire est arrêté dès qu'elles se détachent de la lame basale, les cellules infectées continuent leur différenciation tout en entrant dans la phase de synthèse du cycle cellulaire, ce qui permet une amplification du génome viral (Stubenrauch & Laimins, 1999).

Dans les cellules superficielles de l'épithélium, la protéine E4, bien que codée par un gène précoce, est considérée comme une protéine tardive et exprimée le plus souvent en parallèle avec les protéines L1 et L2. La synthèse des protéines de la capsidite permet l'encapsidation du génome et la production des virions complets. Les HPV ont un cycle non lytique; les virions matures nouvellement formés sont donc libérés dans le milieu extérieur au cours de la desquamation des cellules des couches les plus superficielles de l'épithélium.

Infection latente :

Contrairement à l'infection productive, il arrive parfois que l'ADN viral reste à l'état libre sous forme plasmidique dans la cellule. Une infection latente s'installe donc (Oh *et al.*, 2004), mais, sous l'influence de certains facteurs exogènes ou endogènes (immunodépression locale ou générale) l'évolution vers une infection productive est possible.

L'intégration de l'ADN viral au génome cellulaire :

L'intégration du génome viral au sein du génome de la cellule de l'hôte ne fait pas partie du cycle normal du virus car elle empêche la production de nouveaux virions.

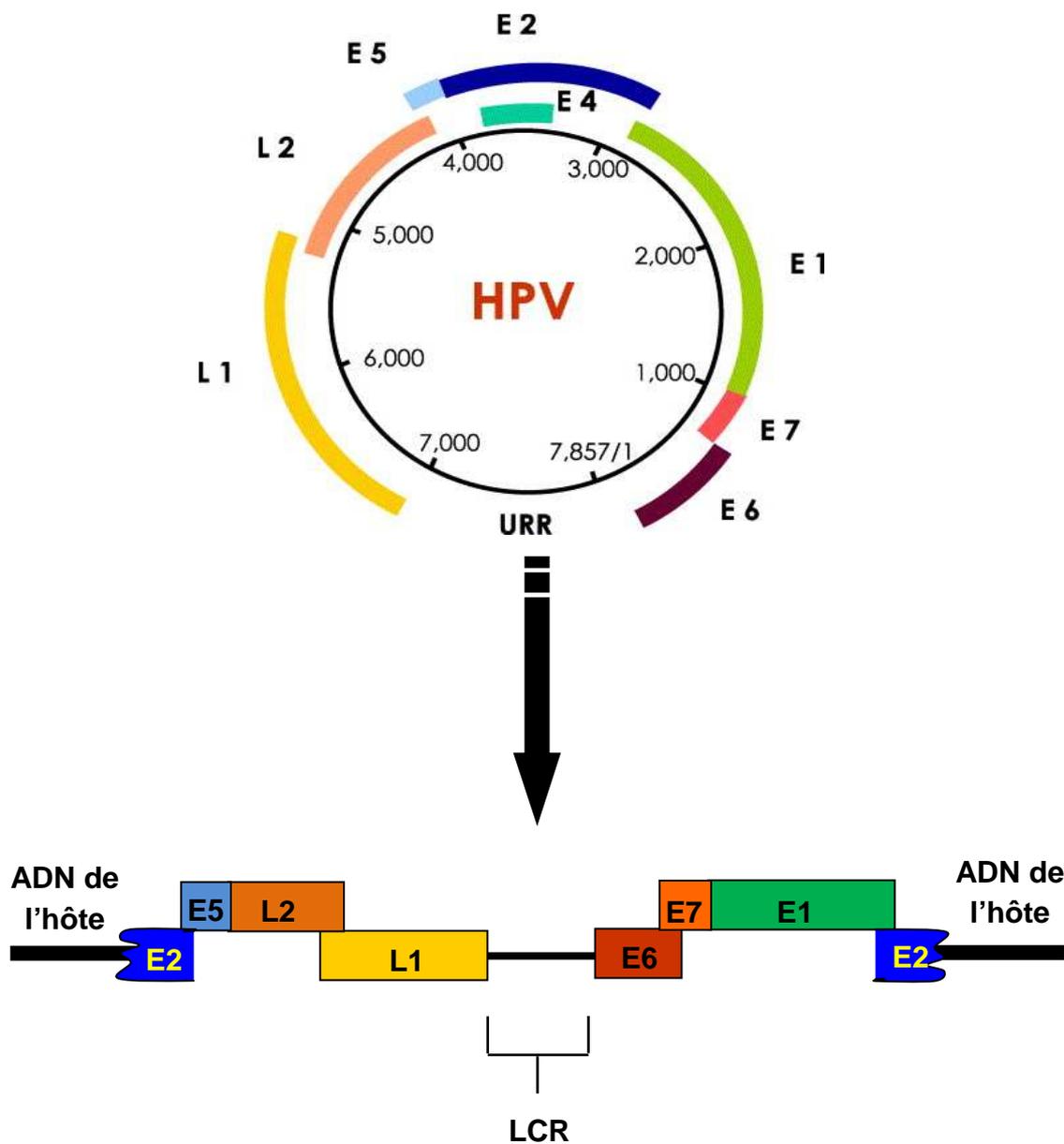


Figure 4: Représentation schématique de l'intégration de l'ADN virale, dans le génome de l'hôte. L'intégration se produit dans la plupart des cellules cancéreuses et semble être un événement décisif pour le développement tumoral. L'ORF de E2 est scindé au cours de l'intégration, entraînant une perte de l'ORF E4. (Adopté d'après: Dell et Gaston, 2001)

L'intégration est précédée par la linéarisation du génome viral suite à une rupture dans la région E1/E2 (Romanczuk et Howley, 1992). La rupture du gène E2 se fait de manière aléatoire mais certains auteurs suggèrent qu'il existe des zones préférentielles d'intégration (Hebner et Laimins, 2006). Cette étape de transformation concerne uniquement les HPV à "haut risque" dont le type 16. C'est un événement important de la dérégulation des gènes E6 et E7 car leur transcription n'est plus réprimée par la protéine E2, le gène E2 étant partiellement scindé au cours de la linéarisation du génome viral (Figure 4). Les transcrits L1 et L2 commencent à disparaître alors que les transcrits de E6 et E7 seront très abondants. Les protéines codées par ces deux gènes possèdent la propriété d'interagir avec des produits de gènes cellulaires impliqués dans le contrôle de la prolifération cellulaire. La synthèse accrue des protéines oncogènes favorise la promotion et la prolifération tumorale (la figure 5 présente un modèle du cycle viral de HPV 16).

Pouvoirs pathogènes des HPV

Il existerait plus de 200 types d'HPV (zur Hausen, 1999). À l'heure actuelle, le génome d'une centaine d'entre eux a été isolé et séquencé. Les HPV infectent les cellules épithéliales de la peau ou des muqueuses, et sont transmis par dissémination de ces cellules lors de la desquamation. Ces infections, le plus souvent responsables de lésions bénignes (verrues cutanées, condylomes ano-génitaux, papillomes laryngés), sont parfois associées à certains cancers. On distingue donc les HPV dits à "bas risque", ne générant que des micro-proliférations cellulaires bénignes, par opposition aux HPV à "haut risque", associés à certaines tumeurs malignes. La plupart des HPV à tropisme cutané sont de bas-risque. Mais, dans certains cas d'infection par les HPV5 ou HPV8, les lésions peuvent évoluer en cancer de la peau au niveau des zones exposées aux UV (Majewski & Jablonska, 2003, Pfister *et al.*, 2003).

Les HPV à tropisme génital sont transmis par voie sexuelle et sont très fréquents puisque plus de 80% des femmes sexuellement actives ont été ou sont infectées par un HPV durant leur vie (Bosch, 2003). Parmi les HPV muqueux, on distingue les HPV à "bas risque", par exemple les types 6 et 11 responsables de verrues génitales ou condylomes acuminés, et des HPV à "haut risque" qui sont associés aux CIN (lésions précancéreuses du col de l'utérus). Parfois, ces lésions peuvent évoluer en carcinomes in situ du col de l'utérus. Actuellement, une quinzaine de types de HPV ont été identifiés comme agents étiologiques du cancer du col de l'utérus les plus fréquemment en cause étant les deux types 16 et 18

(Munoz *et al.*, 2006). Ces HPV oncogènes peuvent également toucher d'autres muqueuses ano-génitales et notamment provoquer des cancers de la zone ano-génitale (vulve, pénis ou anus) (Steenbergen *et al.*, 2005) (Tableau 1).

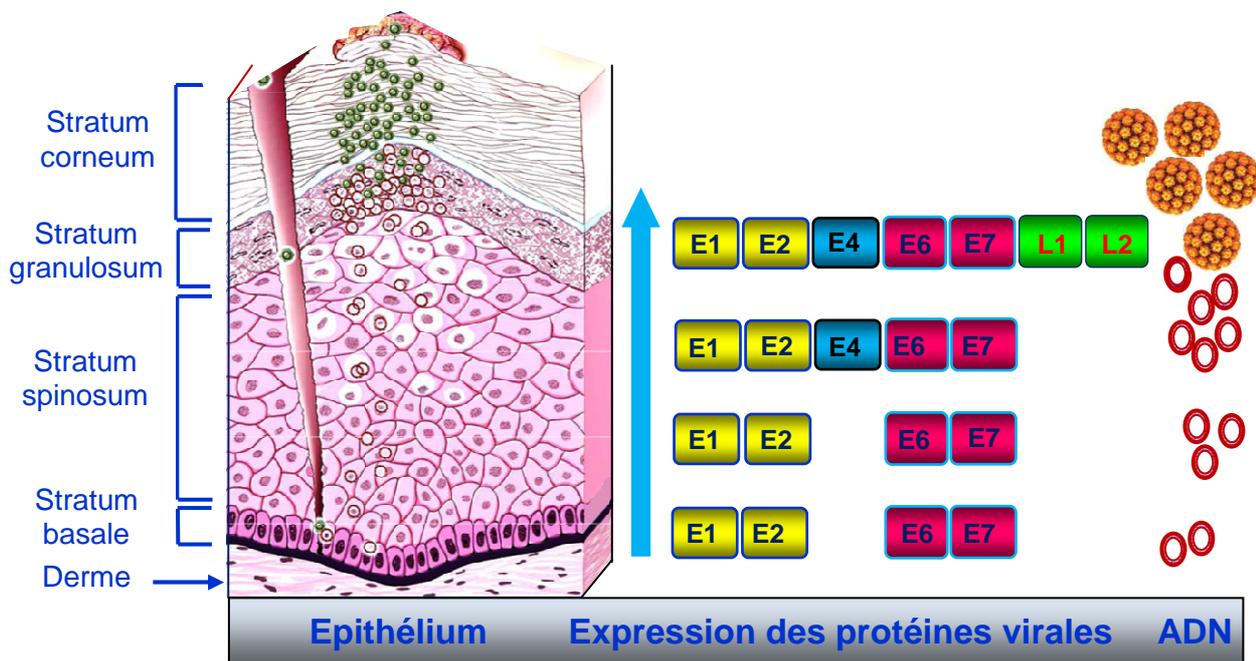


Figure 5: Modèle du cycle viral de HPV16. L'infection se produit à la faveur de microlésions de l'épithélium, ce qui permet au virus de pénétrer dans les cellules de la couche basale. Au cours de la différenciation de la cellule hôte, les protéines virales précoces sont exprimées et l'ADN virale est amplifié. Ce n'est que dans les couches supérieures des kératinocytes, plus différenciées (stratum granulosum), que les protéines de la capsid L1 et L2 sont produites. Ceci conduit à l'assemblage des particules virales et ensuite les virions seront libérés au niveau de la couche cornée.

Type	Pathologies
HPV-2, HPV-27, HPV-57	Verrues épidermiques classiques, verrues génitales infantiles
HPV-26, HPV-51, HPV-69, HPV82	Lésions malignes et bénignes des muqueuses
HPV-53, HPV-30, HPV-56, HPV-66	Lésions malignes et bénignes des muqueuses
HPV-18, HPV-39, HPV-45, HPV-68, HPV-70	Lésions malignes des muqueuses, principalement dans les adénocarcinomes
HPV-7, HPV-40, HPV-43	Condylomes acuminés, CIN et lésions cutanées de bas-risque
HPV-16, HPV-31, HPV-33, HPV-35, HPV-52, HPV-58, HPV-67	Lésions malignes des muqueuses, principalement dans les carcinomes squameux du col de l'utérus
HPV-6, HPV-11, HPV-13, HPV-44, HPV-74	Lésions bénignes des muqueuses. HPV-6 et HPV-11 dans les verrues génitales, condylomes acuminés du col, papillomes du larynx. Certaines de ces lésions peuvent progresser vers des cancers
HPV-34, HPV-73	Lésions des muqueuses de haut-risque
HPV-5, HPV-8, autres	Lésions cutanées malignes et bénignes chez patients immunodéprimés
HPV-4, HPV-65, autres	Lésions cutanées types verrues communes
HPV-1, HPV-63	Lésions cutanées, verrues plantaires
HPV-41	Lésions cutanées (type verrues, kistes épidermiques plantaire ou palmaire)

Type de l'HPV	
Haut risque	16, 18, 31, 34, 33, 35, 45, 51, 52, 56, 58, 66, 68, 70, 73
Bas risque	6, 11, 40, 42, 44, 53, 54, 55, 57, 59, 67, 69, 71, 74,

Tableau 1: Classification des principaux HPV impliqués dans des pathologies humaines. Les HPV à haut-risque, responsables des cancers du col de l'utérus (en rouge), et les HPV potentiellement à haut-risque (en orange) (Munoz *et al.*, 2005; McMurray *et al.*, 2001).

Histoire naturelle de l'infection à HPV lié à l'histoire naturelle du cancer du col de l'utérus

Il est donc désormais clairement établi que le risque de développer un cancer, épidermoïde ou adénocarcinome, du col de l'utérus est étroitement lié à la présence de HPV à "haut risque". L'HPV de type 16 (HPV16) est le plus fréquemment associé aux tumeurs du col de l'utérus (53,3 %), suivi des types 18 (17,2 %), 45 (6,7 %), 33 (2,9 %) et 31 (2,6 %) (Munoz *et al.*, 2004) (Figure 6).

Le cancer épidermoïde du col de l'utérus est la première tumeur solide humaine reconnue comme étant à quasiment 100% viro-induite. Les dysplasies cervicales ou néoplasies intra-épithéliales cervicales (CIN: Cervical Intra-epithelial Neoplasia) sont considérées comme les précurseurs de ce type de cancer. Le degré de dysplasie se mesure d'après la proportion de l'épithélium occupée par les cellules basales / parabasale. Plus la lésion est sévère, plus la maturation de l'épiderme devient défectueuse. On peut distinguer 3 formes de CIN codifiées en trois stades de gravité croissante: le premier stade est l'apparition des CIN1 où seul le tiers inférieur de l'épithélium est modifié, les CIN2 correspondent à des épithéliums dont les deux tiers sont modifiés et enfin les CIN3 sont rencontrés lorsque toute la hauteur de l'épithélium est impliquée dans ces néoplasies (Figure 7). En générale le cancer cervical résulte d'une progression sur plusieurs années (de 10 à 20 ans), du CIN1 au CIN3 puis vers un cancer (Barron et Richart, 1968).

Cancer du col de l'utérus

Le cancer du col de l'utérus est un problème important de santé publique, car il représente une des principales causes de mortalité chez la femme à travers le monde. Il est le deuxième cancer le plus fréquent chez la femme avec près de 500000 nouveaux cas estimés en 2002 (Parkin *et al.*, 2005). Le cancer du col de l'utérus a provoqué en 2005 près de 260 000 décès dont près de 95 % dans les pays en développement, pays dans lesquels ce cancer est la première cause de mortalité dans la population féminine d'après le Centre International de Recherche sur le Cancer de l'OMS.

Il existe une grande inégalité de répartition de l'incidence selon les pays, au bénéfice des pays industrialisés, car 83 % des nouveaux cas de cancers du col de l'utérus surviennent dans les pays en voie de développement. Le taux d'incidence standardisé varie

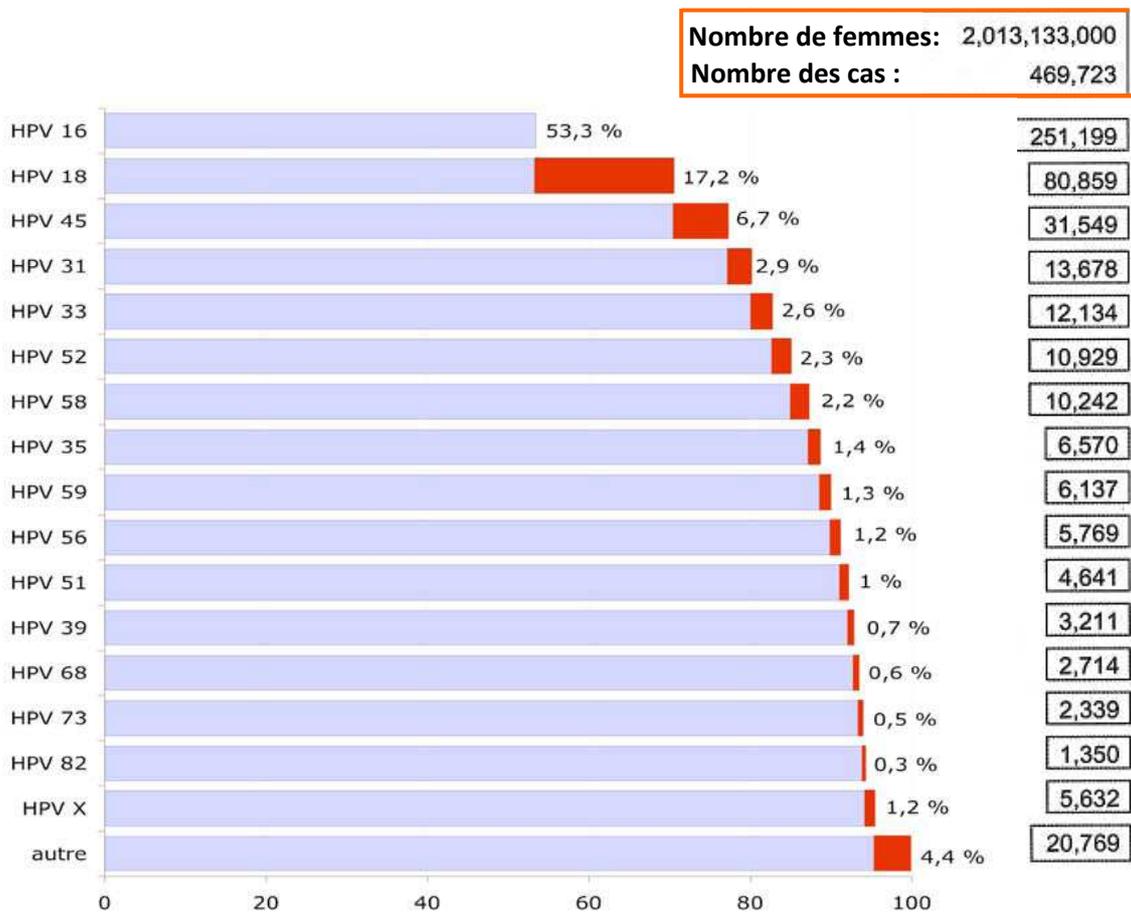


Figure 6: Prévalence des différents types d'HPV à "haut risque" dans le cancer du col de l'utérus. Etude à l'échelle mondiale, toutes les régions du monde confondues (les femmes sont âgées de 15 ans et plus) x : types de HPV non caractérisés (D'après: Munoz *et al.*, 2004)

de 2,1 (en Égypte) à 47,3 (au Zimbabwe) cancers invasifs du col utérin pour 100 000 femmes selon les pays. Toujours dans les pays en voie de développement, le risque d'être atteint d'un cancer du col utérin au cours de la vie est estimé à 4 % contre un risque inférieur à 1 % dans les pays industrialisés. Ces différences sont essentiellement liées à l'absence d'une politique de dépistage systématique du cancer combiné à des coûts excessifs pour les programmes de prévention.

En France le cancer du col de l'utérus arrive en 8ème position de mortalité par cancer chez la population féminine. Les tests de dépistages systématiques pratiqués dans le pays ont eu en effet un effet bénéfique, puisque les taux d'incidence standardisés sont passés de 15,9 pour 100000 à 8,6 pour 100000 de 1978 à 1992 (Weidmann, 1998).

Facteur étiologique principal:

Il est très clair aujourd'hui que l'infection persistante par un HPV à "haut risque" oncogène est la cause principale du cancer du col utérin. Il s'agit du premier cancer à avoir été reconnu par l'OMS comme étant attribuable à 100% à une infection virale. L'ADN de ces HPV a été retrouvé dans 99,7% des échantillons de biopsies des carcinomes du col de l'utérus (Walboomers *et al.*, 1999).

Face à la défense immunitaire, de l'organisme, l'infection par les HPV est transitoire dans la majorité des cas. L'éradication requiert en général 10 à 18 mois (Franco *et al.*, 1999; Moscicki *et al.*, 2001). Les anomalies cervicales survenues après l'infection par les HPV régressent spontanément dans la plupart des cas. Parfois l'infection virale peut persister et l'intégration d'une partie de l'ADN viral dans le génome des cellules hôtes conduit à l'apparition des lésions précancéreuses.

Plusieurs études ont montré que l'intégration du génome HPV dans l'ADN chromosomique de la cellule hôte était très souvent associée aux lésions précancéreuses et aux cancers du col de l'utérus. L'étude des sites d'intégration chromosomiques d'intégration des séquences des HPV a été réalisée sur des cultures continues de kératinocytes néoplasiques (HeLa, Caski...) et sur des cultures à court terme d'échantillon de tumeurs génitales. Il apparaît que certaines régions chromosomiques (8q24.1, 12q14-q15) constituent des sites fréquents d'intégration du génome des HPV dans les tumeurs génitales. Ces régions chromosomiques « fragiles » ont déjà été impliquées dans d'autres processus tumoraux (Wentzensen *et al.*, 2004 ; Couturier *et al.*, 1991). Il apparaît que l'intégration de séquences virales adjacentes à des gènes cellulaires impliqués dans le contrôle de la prolifération cellulaire peut entraîner une activation ou une inactivation de ces gènes.

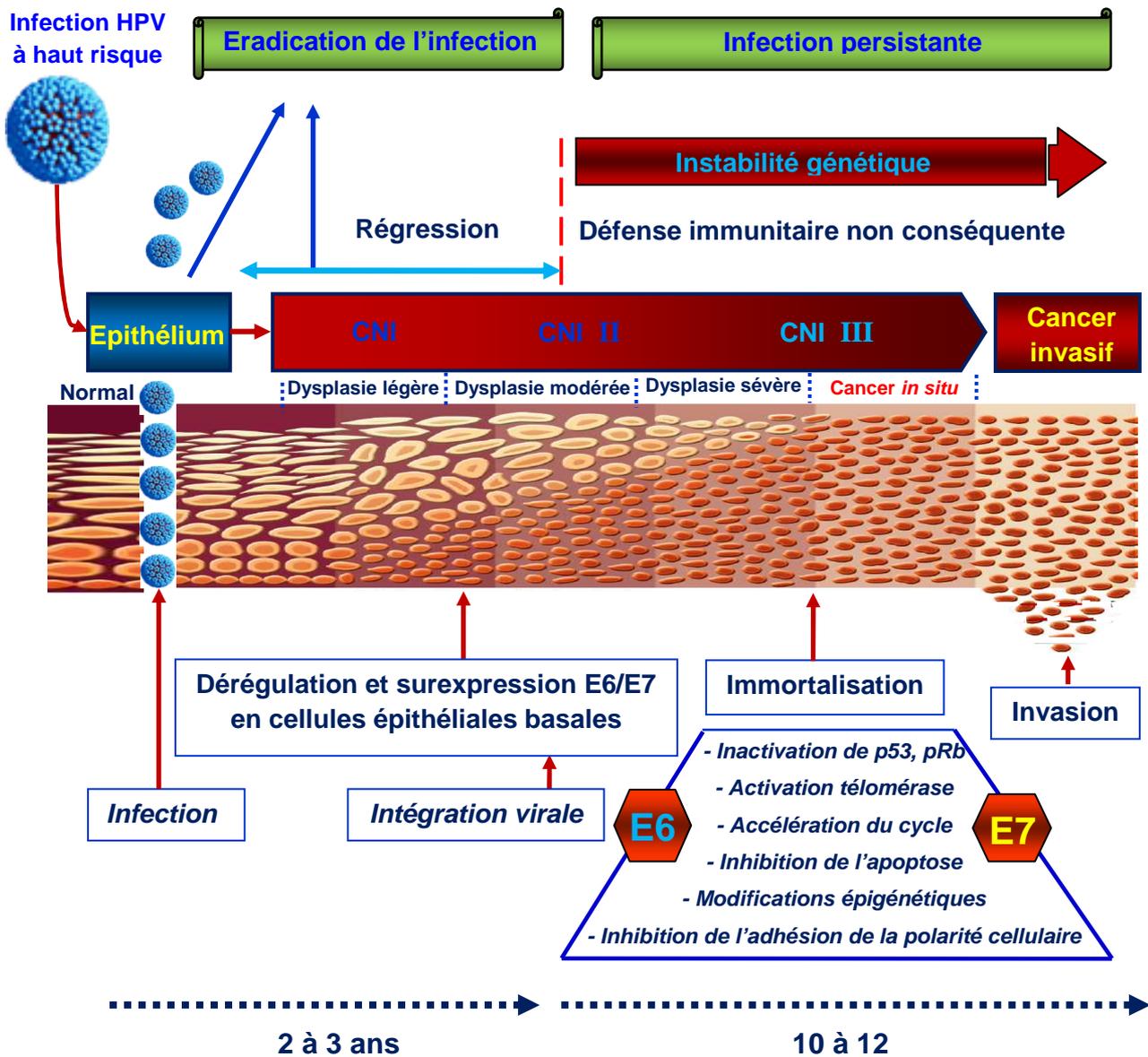


Figure 7: Modèle de l'histoire naturelle d'un HPV à haut risque liée à l'histoire naturelle du cancer du col de l'utérus. Progression en multi-étapes observées dans le développement du cancer du col de l'utérus.

Par exemple, l'intégration de séquences virales dans la région 2p24 au voisinage de l'oncogène myc, s'accompagne, dans les cancers cervicaux, d'une expression d'un gène normalement réprimé dans un épithélium cervical sain (Couturier *et al.*, 1991 ; Sastre-Garau *et al.*, 1995 ; Thorland *et al.*, 2000). Cependant, il semble que le lien entre l'avènement du processus carcinogénèse avec l'intégration du génome virale dans l'ADN des cellules hôtes, soit davantage dû à des modifications structurales du génome viral ayant pour conséquence une dérégulation de l'expression des oncogènes viraux, plutôt qu'à la mutagenèse occasionnée par l'intégration au niveau d'un locus donné (Wentzensen *et al.*, 2004). Toutefois, bien que l'infection par les HPV à "haut risque" soit identifiée comme le responsable direct de la maladie, il existe certains cofacteurs secondaires.

Facteurs étiologiques secondaires (cofacteurs) :

L'incidence de différents facteurs sur la fréquence de survenue des cancers du col de l'utérus a été largement étudiée (Figure 8).

- Le comportement sexuel est considéré comme un facteur aggravant le risque de développer un cancer du col de l'utérus. Le risque est environ trois fois supérieur chez les femmes ayant 10 partenaires différents, comparativement à celles ayant un seul partenaire (la charge virale est peut être en jeu). La précocité du premier rapport sexuel entre également en compte, les femmes ayant eu leur premier rapport sexuel avant l'âge de 16 ans présentant un risque deux fois plus élevé que celles dont le premier rapport a eu lieu après 20 ans. Cette augmentation est probablement due à la plus grande sensibilité du col utérin à l'action de différents carcinogènes pendant l'adolescence (Singer, 1975). De plus, une association entre l'utilisation à long terme d'une contraception orale et le risque de développer un cancer du col a été observée dans la plupart des études menées sur le sujet, mais elle demeure très faible. L'action des hormones contraceptives peut en partie être expliquée par l'existence d'éléments de réponse aux glucocorticoïdes et à la progestérone au niveau de la zone de régulation du génome viral (LCR).

- L'état du système immunitaire au moment de l'infection est très certainement un facteur important puisqu'il a été mis en évidence une augmentation de la persistance des HPV et de la progression des CIN chez les femmes immunodéprimées. Chez les sujets infectés par le VIH, la prévalence des infections à HPV est accrue et corrélée à une augmentation du potentiel oncogène (Matorras *et al.*, 1991).

- Les facteurs génétiques sont aussi des cofacteurs pour le développement du cancer du col de l'utérus. Certaines études ont porté sur l'influence du polymorphisme de la

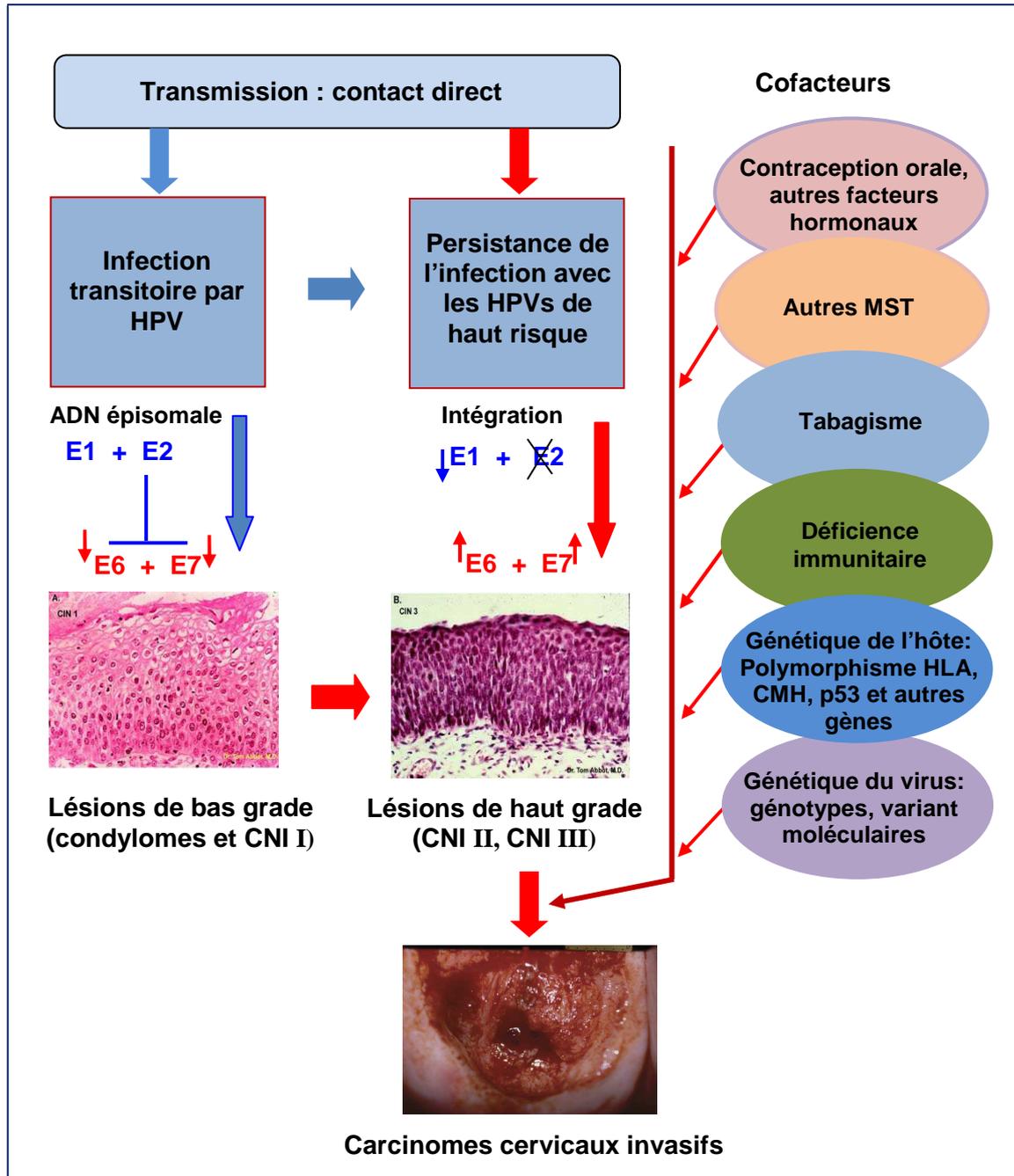


Figure 8: Modèle étiologique du cancer du col de l'utérus. L'infection par un papillomavirus humain de haut risque n'est pas forcément suffisant pour l'apparition des lésions pré-cancéreuses du col de l'utérus. Le développement vers des carcinomes invasifs nécessite la contribution de plusieurs cofacteurs. (D'après: Franco et Harper, 2005; Blachon et Demeret, 2003).

séquence de la protéine suppresseur de tumeur p53 conduisant à la présence soit d'une arginine (Arg) ou d'une proline (Pro) en position 72 (Matlashewski *et al.*, 1987).

La transformation des cellules et l'induction du cancer par les HPV nécessite l'inactivation de p53. Certains auteurs ont montré que l'expression de p53 dans sa forme p53-R72 dans des cellules de rongeurs conduit à un effet deux fois plus efficace pour mener les cellules vers l'apoptose. (Thomas *et al.*, 1996 ; Crook *et al.*, 1994). Ensuite il a été démontré que la forme p53-Arg72 était plus vulnérable à l'inactivation HPV-dépendante que la forme p53-Pro72 (Storey *et al.*, 1998). Cependant, le rapport entre un tel génotype de p53 et la dégradation ou l'inactivation de p53 par les protéines HPV à "haut risque" reste controversé (Koushik *et al.*, 2004; Oliveira *et al.*, 2008).

D'autres prédispositions génétiques (notamment au niveau du système HLA) augmenteraient le risque de développer un cancer induit par HPV. L'immunité cellulaire joue un rôle majeur dans le contrôle de l'infection à HPV et dans l'évolution des lésions dysplasiques, en particulier la présence d'une immunité spécifique vis-à-vis de la protéine E6 du HPV16 (lymphocytes T mémoire) favoriserait la clairance virale (Piersma *et al.*, 2008).

En fin, le rôle de la consommation du tabac est bien décrit comme un cofacteur dans le processus de cancérogènes du col de l'utérus (Kjellberg *et al.*, 2000 ; Hellberg *et al.*, 2005).

Dépistage, traitement et la prophylaxie

L'histoire naturelle du cancer du col de l'utérus laisse suffisamment de temps aux méthodes de dépistage, puisque entre l'infection à HPV et le cancer invasif il s'écoule plus d'une décennie. Le test de dépistage de référence des lésions cancéreuses et précancéreuses du col utérin utilisé en routine est l'examen cytologique (le frottis cervical utérin) et des examens colposcopiques. Récemment, ce test a été complété par de nouvelles méthodes de détection des PVH. Ces méthodes reposent essentiellement sur des techniques d'hybridation moléculaire à l'aide de sondes et/ou amorces spécifiques (hybridation in situ, hybridation liquide, PCR) qui permettent une détection qualitative et même quantitative des génomes viraux et plus récemment de leurs transcrits (PCR quantitative). Le dépistage des lésions précancéreuses de haut grade en particulier, permet en principe de réduire l'incidence du développement du cancer. Cependant, le traitement commence par la surveillance et repose sur la destruction chirurgicale ou l'exérèse de la tumeur, accompagnée d'un suivi rigoureux postopératoire.

Il est donc clair que l'enjeu principal dans la lutte contre le cancer du col de l'utérus pour le moment reste la prévention.

Il n'existe pas de modèle animal adapté pour l'étude de l'efficacité d'un vaccin protecteur contre les HPV, puisque l'infection est strictement spécifique à l'espèce humaine. Des pseudo-particules virales contenant les protéines de capsid des virions, mais ne contenant pas le génome viral, appelées (*virus like particles*) (VLP) et produites par expression et auto-assemblage des protéines constitutives du virus (Touzé *et al.*, 1998), sont aussi immunogènes que les virions natifs (Christensen *et al.*, 1995). Seuls les épitopes conformationnels spécifiques du type viral, contenu dans les VLP composées de la protéine L1 ou L2, sont capables d'induire expérimentalement, chez l'animal, des anticorps neutralisants protecteurs vis-à-vis des infections à HPV de même type. C'est ce qui a amené la société américaine Merck à développer un vaccin prophylactique contenant les VLP-L1 des HPV de type 6, 11, 16 et 18 (Villa *et al.*, 2005) sous le nom de Gardasil™ qui a été commercialisé en juin 2006. Enfin, un autre vaccin (Cervarix^{MC}, immunisant contre les HPV 16 et 18) développé par le laboratoire britannique GlaxoSmithKline (GSK) a été mis sur le marché un an plus tard.

Les protéines virales

La protéine L1

Le produit du gène tardif L1 est une protéine de 55 kDa, son rôle essentiel est l'encapsidation du virus. Elle est exprimée tout comme L2 par le promoteur tardif lors des phases tardives du cycle viral au niveau du cytoplasme des cellules les plus différenciées de l'épithélium. Ces protéines structurales possèdent des signaux de localisation nucléaire (NLS) par lesquels elles sont ensuite transférées dans le noyau où elles s'assemblent pour constituer la capsid virale. La protéine L1 est majoritairement retrouvée dans la capsid, et est capable à elle seule de s'assembler sous forme de particules virales VLPs (*virus-like particles*) (Giroglou *et al.*, 2001; Pinto *et al.*, 2003). La capsid virale mesure environ 55 nm, composé de 70 unités de capsomères arrangées sous une forme icosaédrique et résistante à la chaleur (Bonnez *et al.*, 1994). Cependant la formation d'une capsid virale intacte nécessite la participation de la protéine L2. La stoechiométrie L2/L1 est 1/30, soit 12 copies de L2 par virion.

La structure atomique du virion de papillomavirus n'est pas encore déterminée à cause de la difficulté d'obtenir une quantité suffisante des particules virales pour les analyses structurales. Cependant, par cristallographie, une structure de haute résolution 3.5 Å d'une petite particule contenant 12 pentamères a été résolue. Cette structure révèle les interactions pentamère-pentamère essentielles pour l'assemblage de la particule (Chen *et al.*, 2000). L'assemblage expérimental *in vitro* montre que la partie N-terminus de L1 agit comme un commutateur entre les 72 pentamères du VLP. Un pentamère L1 est constitué de cinq "coudes" d'environ 100 résidus dans le C-terminus, chacun des ces "coudes" se compose de trois hélices- α (h1, h3 et h4). L'interaction entre deux pentamères voisins implique de fortes liaisons hydrophobes entre l'hélice h4 du premier pentamère et les hélices h2 et h3 du deuxième (Bishop *et al.*, 2007) (Figure 9).

Des études controversées suggèrent un rôle de l'intégrine $\alpha 6$ dans l'attachement des particules virales sur les kératinocytes (Evander *et al.*, 1997; Sibbet *et al.*, 2000), mais certains auteurs ont prouvé que des molécules d'héparine sulfate protéoglycanes (HSPGs) sont les récepteurs pour la particule virale de HPV11 (Shafti-Keramat., 2003; Rommel *et al.*, 2005).

La protéine L2

C'est la protéine mineure de la capsid virale (75kDa), et le composant secondaire nécessaire pour la formation de la capsid et son pouvoir infectieuse (Doorbar et Gallimore, 1987; Roden *et al.*, 1996). Il a été proposé que L2 interagisse à la fois avec le génome viral et avec la protéine L1, permettant l'assemblage de l'ADN épisomal dans la nouvelle capsid (El Mehdaoui *et al.*, 2000; Zhou *et al.*, 1994).

Une fois que la particule virale pénètre dans cellule, certains ont suggéré que l'interaction de L2 avec de la β -actine pourrait participer au transport des particules virales vers le noyau (Yang *et al.*, 2003). De même, il a été rapporté que les derniers 40 résidus du C-terminus de L2 interagissent avec un réseau de microtubules via la dynéine, (une protéine motrice associée aux microtubules) qui pourrait transporter les particules virales infectieuses ou bien les particules nouvellement assemblées (Florin *et al.*, 2006).

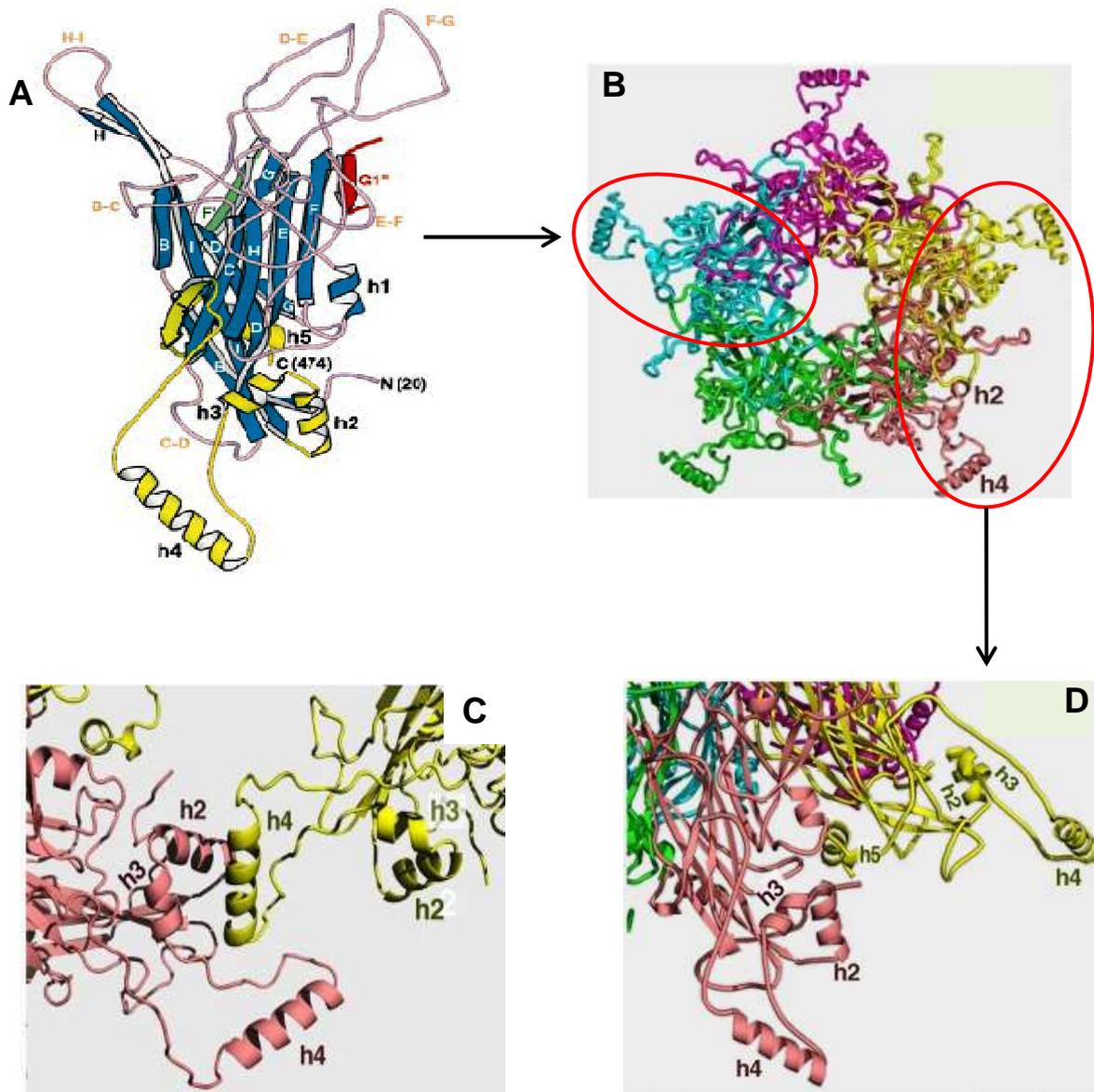


Figure 9: Les éléments structuraux de la protéine L1 et l'assemblage de la particule virale. (A) Représentation schématique d'un monomère de L1. **(B)** Représentation schématique d'un pentamère L1, montrant les projections latérales de chacun des cinq monomères. **(C)** Les interactions entre deux pentamères impliquent une hélice h4 d'un pentamère et les deux hélices h2 de son voisin. **(D)** Vue rapproché des projections latérales de deux monomères, montrant les positions détaillées des hélices h2, h3, h4, et h5 dans le C-terminus de L1 (D'après, Chen *et al.*, 2000; Bishop *et al.*, 2007)

Quelques études sur la vaccination contre les infections aux HPV, suggèrent que l'immunisation avec les protéines L2 puisse fournir une immunité couvrant une plus large variété des souches HPV par rapport à l'immunisation avec L1-VLPs, liée à la présence d'épitopes fortement conservés dans les L2 de différents HPV (Alphs *et al.*, 2008).

La protéine E1

E1 est une protéine d'environ 70 kDa de taille, son niveau d'expression est relativement bas, mais tout à fait essentiel pour la réplication virale. Cette protéine est bien conservée entre les papillomavirus humains et les papillomavirus bovins. Par conséquent la protéine E1 de BPV est souvent considérée comme un modèle pour les HPV E1 (Enemark *et al.*, 2000). E1 présente aussi des homologues de séquences et de fonction avec la grande protéine T-antigène de SV40 et des polyomavirus (Mansky *et al.*, 1997; Enemark *et al.*, 2002; Schuck et Stenlund, 2005).

La E1 protéine joue un rôle critique dans la reconnaissance et la fixation sur l'origine de réplication « ori » dans le génôme virale et par la suite son activation. La structure tridimensionnelle de la protéine E1 montre la présence de trois domaines fonctionnels: un domaine N-terminal, dont la fonction reste encore peu connue, un domaine central pour la fixation de l'ADN, et le domaine C-terminal portant la fonction hélicase/ATPase (Figure 10).

Cette fixation reste néanmoins faible et nécessite toutefois l'aide coopérative de la protéine E2. Cette dernière s'occupe de charger la protéine E1 sur l'ADN en s'y fixant elle-même transitoirement. Chacune de ces protéines s'homodimérise et se fixe sur l'ADN avec une affinité beaucoup plus importante que si E1 se fixait seule sur ces séquences. Il a été montré que la dimérisation de E1 via son domaine de liaison à l'ADN est primordiale pour l'initiation de la réplication. Par la suite, la protéine E2 se détache rapidement après hydrolyse de l'ATP et le recrutement par E1 d'autres protéines chaperonnes telle que Hsp70 et Hsp40 (Liu *et al.*, 1998; Lin *et al.*, 2002). E1 forme alors un hexamère sur l'ADN. Sous cette forme, elle possède une activité ADN-hélicase qui catalyse la séparation des deux brins d'ADN au niveau de l'origine de réplication dans la direction à 3' vers 5' (Fouts *et al.*, 1999; White *et al.*, 2001). L'activité ATPase a été localisée dans la région C-terminale de E1, et fournit l'énergie nécessaire pour l'activité hélicase (Yang *et al.*, 1993). La structure du domaine central ou DBD (DNA binding domain) a été résolue par cristallographie. Elle comprend 4 hélices α et une boucle bien commandée, formant un repliement en continuité sur la surface de la protéine, mais la fixation de l'ADN se fait séparément sur les deux brins.

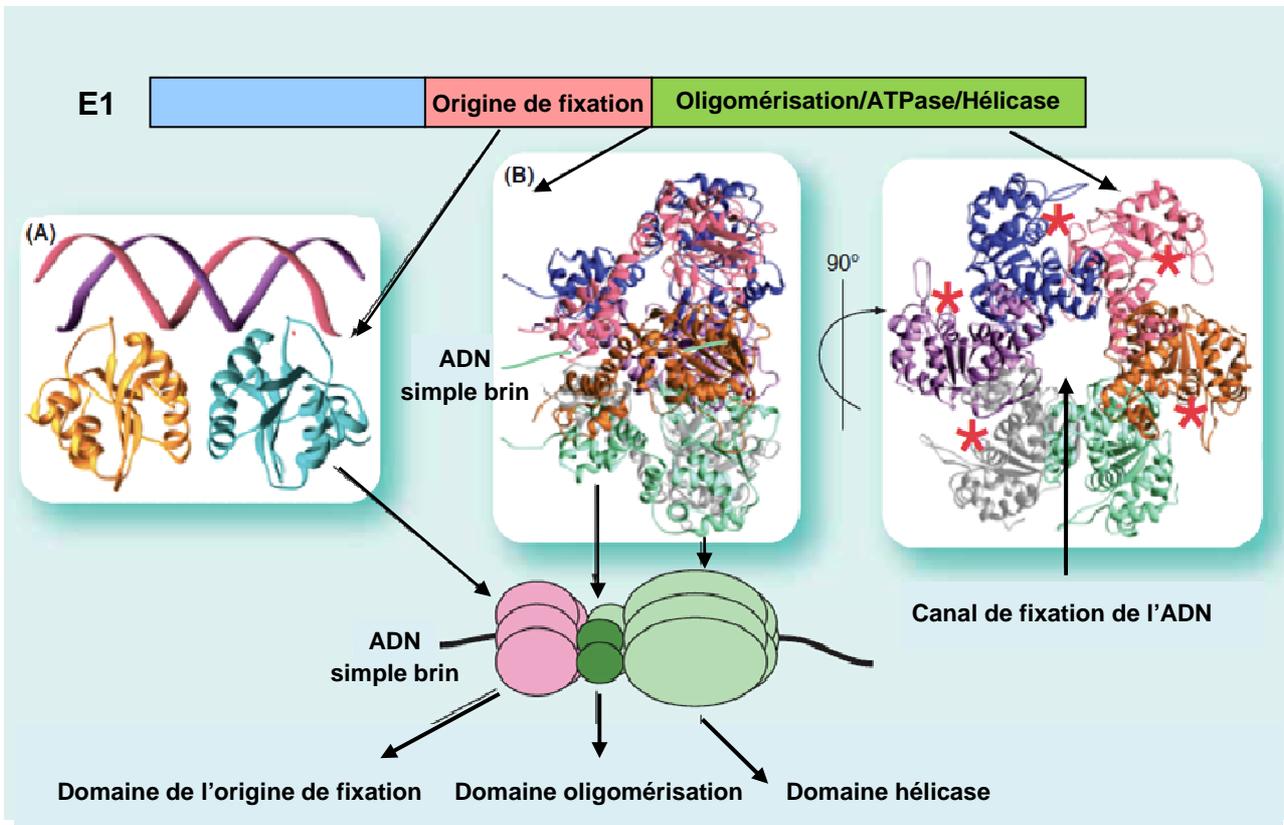


Figure 10: L'architecture des domaines de la protéine E1: En haut, les modèles structuraux, (A) Premier événement d'identification entre un dimère du domaine de fixation de l'ADN OBD (*origin binding domain*) et l'ADN. Les deux sous-unités sont montrées en bleu-vert et or. (B) Modèle montrant deux angles de vues de l'hexamère du domaine E1 oligomérisation/ATP/hélicase lorsqu'il est lié à l'ADN simple brin. Chaque sous-unité est représentée dans une couleur différente. L'oligomérisation se fait par l'intermédiaire de la région N-terminus de ce domaine. Les astérisques rouges indiquent les endroits de fixation de l'ATP entre les sous-unités. En bas, une représentation schématique de la protéine E1 sous forme d'hexamère lié à l'ADN simple brin. (Enemark *et al.*, 2006; D'après la revue Hegde, 2006).

La liaison ADN-hélice α se fait seulement par des interactions non spécifiques, alors que la liaison ADN-boucle est celle qui dirige le contact direct avec les bases spécifiques de fixation de E1. Les résidus qui sont impliqués dans la liaison avec l'ADN sont fortement conservés entre tous les PVs.

Le domaine C-terminal (les résidus 170 – 649) comporte notamment un site de liaison à l'ATP et des domaines avec des activités ATPase-hélicase (Seo *et al.*, 1993). La structure tridimensionnelle du domaine fonctionnel (c'est-à-dire sous forme hexamérique) a été résolue par cristallographie. Elle présente deux hexamères encerclent chacun un simple brin d'ADN qui se glisse dans un canal d'environ 13 Å de diamètre (Figure 10). L'arrangement des protomères présente une forte homologie avec la structure hexamérique de l'antigène-T de SV40 (Enemark et Joshua-Tor, 2006).

La protéine E2

La protéine E2 est une protéine multifonctionnelle de 50 KDa retrouvée dans le noyau cellulaire, le plus souvent sous forme de dimère (Antson *et al.*, 2000). C'est une protéine dont la structure est très conservée au sein des différents PVs. Elle est composée de trois domaines fonctionnels distincts (Figure 11).

Le domaine N-terminal d'environ 200 résidus est responsable de l'initiation de la réplication ou de l'activation de la transcription (McBride *et al.*, 1989); Ce domaine est capable d'interagir avec la protéine Brd4 (Bromodomain containing protein 4) qui s'associe avec les chromosomes du fuseau mitotique. Ceci permet la co-localisation de E2 de BPV 1 et d'HPV 16 avec le fuseau mitotique et l'attachement du génome viral aux chromosomes de l'hôte ce qui est indispensable au maintien des formes épisomales dans les cellules infectées (You *et al.*, 2004).

La structure du domaine N-terminal de transactivation TAD (transactivation domain) révèle l'existence d'un repliement spécifique caractérisé par la présence de deux modules, un module du N-terminal constitué par trois hélices antiparallèles tandis que l'autre module de C-terminal apparaît sous forme de trois feuillets- β . L'orientation des deux modules du E2-TAD semble varier selon le type et la souche virale, mais en dehors de cela, la structure est identique (Antson *et al.*, 2000) (Figure 11). La cristallisation du domaine E2-TAD en complexe avec le domaine hélicase de E1 a donné lieu aux six hétérodimères à symétrie unique.

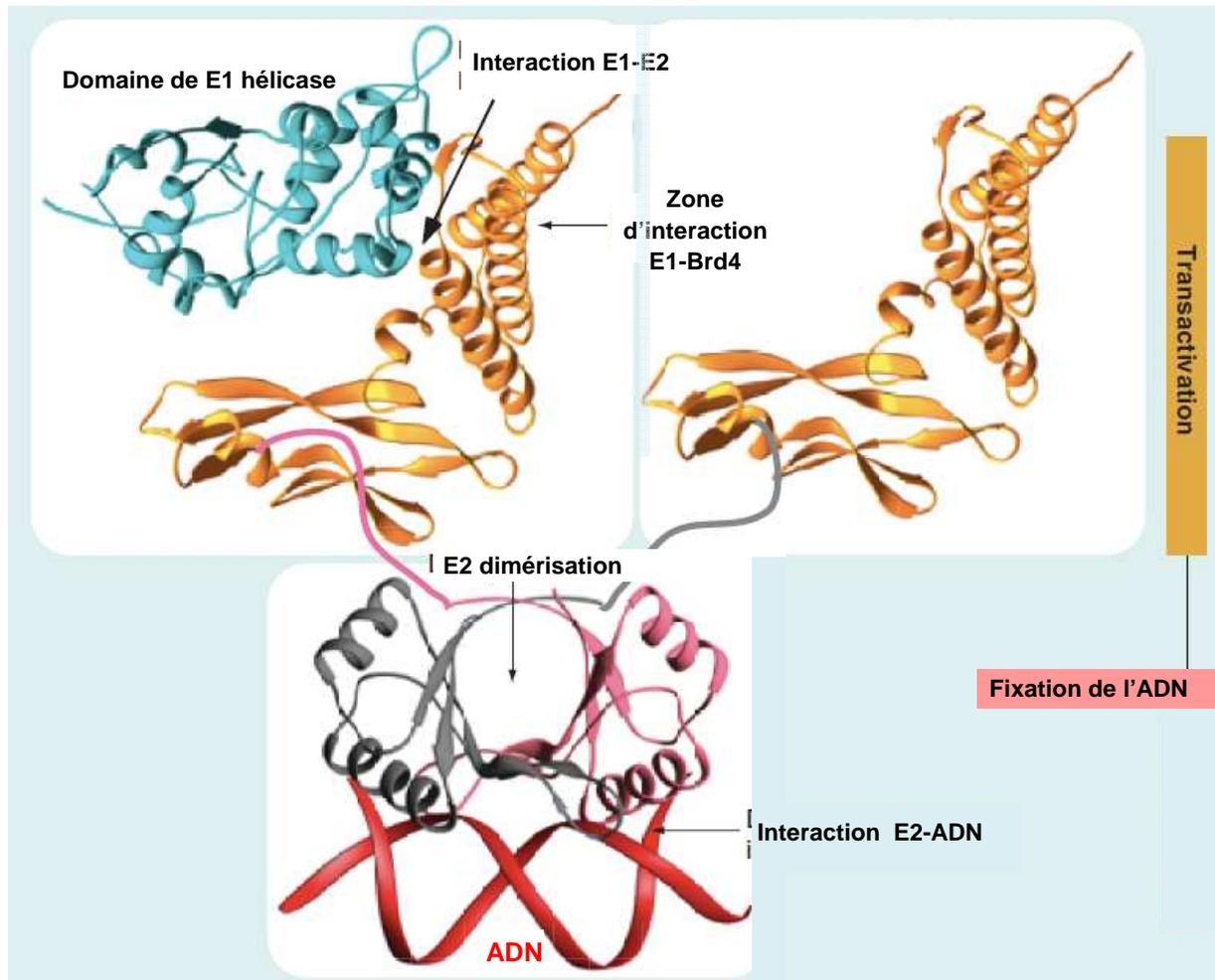


Figure 11: Structure du domaine de fixation de l'ADN et le domaine de transactivation de la protéine E2 en liaison avec l'ADN. Le domaine de fixation de l'ADN (DBD) forme un dimère, les deux sous unités sont représentées avec des couleurs différentes. L'une de sous-unités du domaine de transactivation se lie avec le domaine hélicase de la protéine E1 (Abbate *et al.*, 2004). Des sites potentiels pour l'interaction entre E2 et ses partenaires sont illustrés sur le modèle structural. (D'après: Abbate *et al.*, 2004; Harris et Botchan, 1999; Hegde, 2006)

Bien que les trois hélices de E2-TAD qui font en générale des contacts de dimérisation (en l'absence de E1), conservent leurs structures dans le complexe, le potentiel de dimérisation de cette interface semble perdu dans la structure du complexe. La structure du complexe E2-TAD –E1/hélicase est illustrée comme la lettre « C », avec E1/hélicase à l'intérieur en contact avec une hélice et le linker reliant les modules N-terminus et C-terminus du domaine E2-TAD qui représente le corps de la lettre « C » (Abbate *et al.*, 2004) (Figure 12).

Le domaine central de E2, appelé aussi la région « charnière », de séquence et taille variable selon les types de PVs (McBride *et al.*, 1989) comporterait une séquence du motif PEST (riche en proline, glutamine, serine et thréonine) qui contient des sites de phosphorylation (Penrose et McBride, 2000). La phosphorylation de l'un de ces sites dans le motif PEST est connue pour être un signal d'ubiquitylation et de dégradation de la protéine E2 par la voie de protéasome (Penrose *et al.*, 2004). Bien que cette région soit considérée généralement comme une région flexible non structurée, certaines études ont confirmé par RMN et dichroïsme circulaire l'existence d'une structure secondaire persistante pour une trentaine de résidus couvrant la région PEST. Cette structure comprend une courte hélice α suivie par une hélice polyproline dans l'état non phosphorylé (Garcia-alai *et al.*, 2006). Cependant, il a été montré dans cette étude que la phosphorylation du polypeptide déstabilise les deux hélices (l'hélice- α et l'hélice polyproline) dans la structure secondaire (Garcia-alai *et al.*, 2006).

Le domaine C-terminal de E2 compte environ 90 résidus et assure la dimérisation de la protéine ainsi que sa liaison à l'ADN au niveau du motif ori (McBride *et al.*, 1988). Ce domaine régule la réplication du génome viral en association avec E1, mais inhibe par encombrement stérique la transcription des gènes précoces E6 de HPV génitaux (Dong *et al.*, 1994; Demeret *et al.*, 1997). Les protéines E2 de certaines souches non cutanées de papillomavirus reconnaissent une séquence palindromique consensus de 12 pb (ACCgNNNNcGGT) appelé E2 BS (E2 binding site), les nucléotides indiqués par les lettres minuscules ne sont pas spécifiques pour l'identification du site (Hedge, 2002).

La structure 3D du domaine C-terminal de fixation de l'ADN (DBD) de E2 issu de différents types BPV1 (Hegde *et al.*, 1992), HPV16 (Hegde et Androphy, 1998), HPV18 (Harris et Botchan, 1999), HPV 6 (Hooley *et al.*, 2006) et 31 (Liang *et al.*, 1996) et son complexe avec l'ADN présente un repliement comparable dans tous les cas. La structure commune comporte un homodimère (deux unités) avec 8 feuillets- β antiparallèles (chaque monomère contribue pour moitié de ces feuillets) entourés de l'extérieur de 4 hélices- α . La topologie de chaque sous-unité est la suivante: β 1- α 1- β 2- β 3- α 2- β 4.

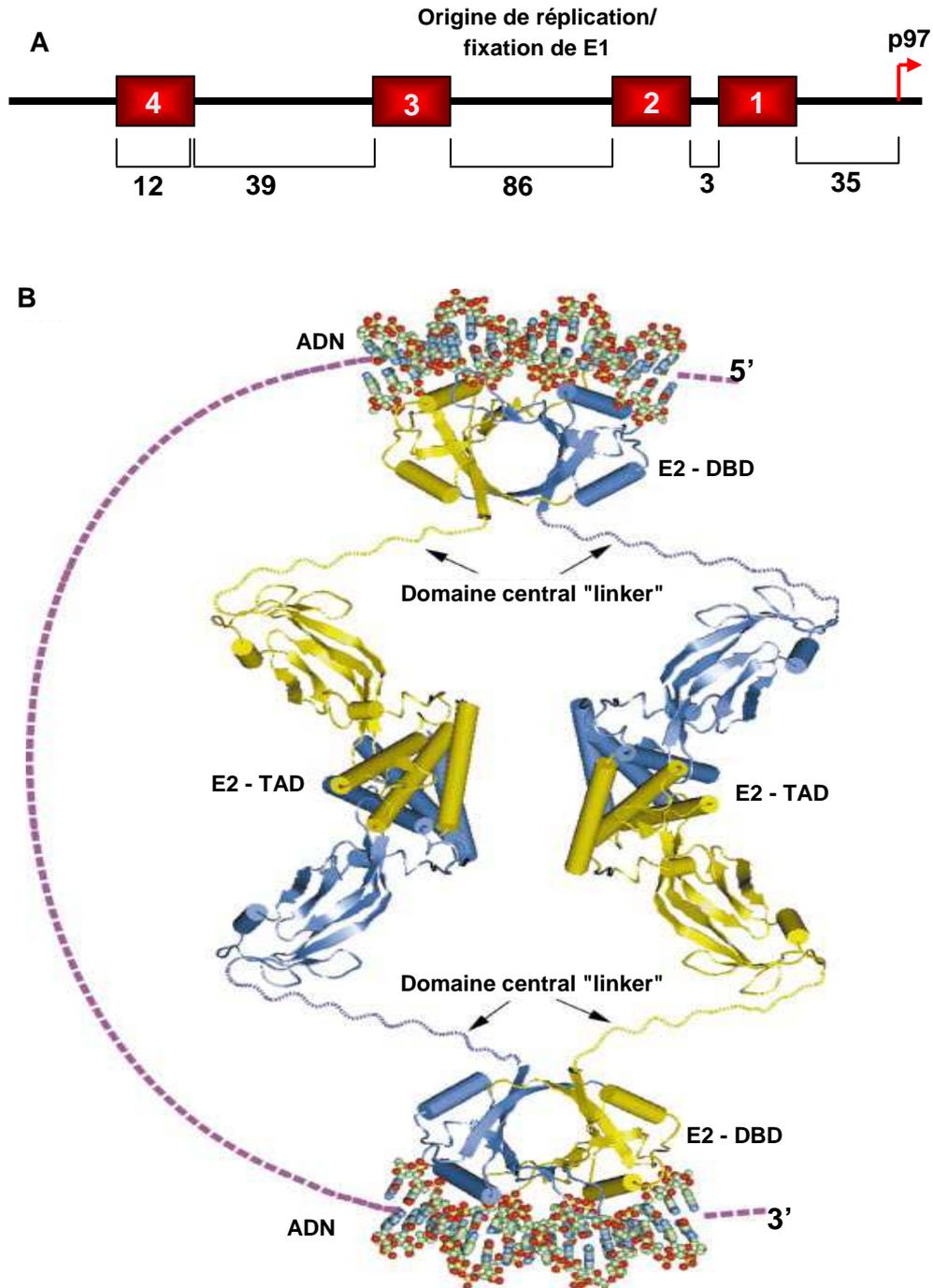


Figure 12: Modèle de la fixation de la protéine E2 dans la région LCR du génome HPV16. Illustrant, **(A)** Représentation schématique montrant les quatre sites de fixation de E2 sur l'ADN, chacun des quatre sites correspond à 12 pb. **(B)** Modèle de formation des boucles d'ADN, montrant des interactions entre les modules d'E2NT et d'E2CT. Le domaine E2 DBD s'associe sous forme dimérique pour lier l'ADN génomique (D'après Antson *et al.*, 2000).

L'une des hélices (hélice $\alpha 1$) est utilisée dans le contact avec l'ADN et appelée «l'hélice de reconnaissance d'ADN». La différence significative entre les structures tridimensionnelles de différents domaines E2-DBD, dépend de l'orientation relative des hélices de reconnaissance, qui varie selon la souche virale (Cicero *et al.*, 2006; Hooley *et al.*, 2006).

En conclusion, la protéine E2 joue un rôle essentiel dans le cycle viral en initiant la réplication et en contrôlant la transcription des oncogènes viraux. Son inactivation est donc une étape clé dans la transformation maligne comme mentionné dans le paragraphe de l'intégration virale.

La protéine E4

Cette protéine de 17 kD est synthétisée par le promoteur tardif qui s'active généralement lorsque les cellules hôtes commencent à se différencier. La région qui code pour cette protéine est l'une des plus variables au sein des PV (Doorbar *et al.*, 1986) ; la protéine E4 est transcrite avec les 5 acides aminés N-terminaux de la protéine E1, car elle est exprimée sous la forme d'une protéine de fusion E1^{E4} (Chow *et al.*, 1987; Wang *et al.*, 2004). Il existe deux autres formes mineures de 10 et 16 kDa qui sont obtenues par modifications post-transcriptionnelles et par clivage protéolytique (Brown *et al.*, 1996). L'ARNm codant pour E4 est très abondant lors de son expression. Il a été suggéré que la protéine HPV1 E4 représente entre 20 et 25% de l'ensemble des protéines virales exprimés dans la cellule hôte (Doorbar *et al.*, 1986). Cependant, le rôle de cette protéine n'est pas encore complètement élucidé. Pour les papillomavirus à "haut risque", E4 faciliterait l'encapsidation du génome viral et à travers sa capacité de liaison à la cytokératine elle favoriserait la déstabilisation du cytosquelette (Doorbar *et al.*, 1991; Wang *et al.*, 2004). E4 semble aussi être capable d'induire un arrêt du cycle cellulaire en phase G2 (Longworth et Laimins, 2004; Knight *et al.*, 2004).

La protéine E5

Le gène E5 de papillomavirus code pour une petite protéine hautement hydrophobe d'environ 83 résidus (Bubb *et al.*, 1988) localisée principalement dans la membrane des endosomes, l'appareil de Golgi (Oetke *et al.*, 2000) et occasionnellement dans la

membranes plasmique (Conrad *et al.*, 1993). La protéine E5 semble constituée de 2 domaines fonctionnels. Le domaine N-terminal est responsable de la localisation au niveau membranaire et des compartiments cytosoliques (Schlegel *et al.*, 1986).

Le domaine C-terminal possède une activité transformante chez le BPV1 (Burkhardt *et al.*, 1987; Sparkowski *et al.*, 1996), mais cette activité est très faible pour la protéine E5 des HPVs.

La protéine E5 ne présente qu'une faible activité de transformation (Valle et Banks, 1995) et n'est pas nécessaire pour les ultimes étapes de la carcinogenèse, parce que le gène codant pour E5 est souvent perdu lors de l'intégration sans que cela affecte la cellule hôte. De plus, il n'existe pas d'ORF E5 dans le génome de certains papillomavirus. Toutefois, E5 augmente la capacité proliférative des kératinocytes humains, et leur réponse au facteur de croissance épidermique EGF (Epidermal Growth Factor). D'autres études ont montré qu'en présence d'EGF, l'expression de E5 entraîne une augmentation du niveau de phosphorylation de son récepteur EGFR (Epidermal Growth Factor receptor) (Straight *et al.*, 1993). Récemment, il a été aussi démontré que les protéines E5 et E6 de HPV à "bas risque" et à "haut risque" coopèrent ensemble pour induire *in vitro* la formation de koilocytes en cellules cervicales humaines. E5 et E6 induisent également la koilocytose dans des kératinocytes humains, mais pas en cellules COS de primates. La délétion des 20 derniers acides aminés C-terminaux de E5 inhibe complètement le koilocytosis, tandis qu'une délétion partielle de 10 résidus maintient approximativement 50% de son activité. Ces données suggèrent que les deux oncoprotéines E5 et E6 coopèrent dans le cycle de vie viral (Genther *et al.*, 2003). De plus, la protéine E5 d'HPV16, comme celle de BPV1, joue un rôle dans la diminution de la réponse immunitaire en inhibant l'expression des molécules du CMH-I (complexe majeur d'histocompatibilité de type I). En effet, E5 inhibe l'acidification de l'appareil de Golgi, ce qui permet la séquestration des molécules du CMH de classe I et empêche la présentation des épitopes viraux à la surface cellulaire (Ashrafi *et al.*, 2006).

Les oncoprotéines E7 et E6

Le cycle de vie du HPV dépend largement de la machinerie cellulaire de l'hôte pour sa réplication. Une des « stratégies » visées par les protéines précoces est d'assurer la continuité de la division cellulaire. Cependant, un grand nombre des facteurs de réplication de son hôte sont sous le contrôle de facteurs de transcription tels que pRB, p53, ... Les protéines E7 et E6 interagissent avec plusieurs de ces facteurs et entraînent par la suite leur

dégradation, déjouant le système de régulation et menant à l'immortalisation des cellules infectées. Ces phénomènes sont à la base de leur pouvoir oncogène. Dans les paragraphes suivants, je détaillerai, pour ces deux oncoprotéines, leurs caractéristiques biochimiques et leurs rôles ainsi que leurs cibles cellulaires.

La protéine E7

La protéine E7 est codée par un OFR qui se trouve immédiatement en aval du gène E6. Tout comme E6, E7 est capable d'induire la transformation des lignées cellulaires de fibroblastes de rongeurs (Vousden *et al.*, 1988). E7 est une phosphoprotéine localisée principalement dans le noyau cellulaire. Son action passe par la formation de complexes avec des protéines impliquées dans la différenciation ou encore dans la prolifération comme par exemple la pRb (Dyson *et al.*, 1989; Phelps *et al.*, 1992).

L'activité de E7 et ses cibles cellulaires

Il a été mentionné que E7 est capable de lier le facteur de transcription pRb par la région CR2 qui contient le motif (LXCXE). La dérégulation du cycle cellulaire par interaction avec pRb est une propriété commune à d'autres oncoprotéines virales telles que, E1A de l'adénovirus ou l'antigène T du virus SV40 (Munger *et al.*, 2001) ce qui souligne l'importance de cette interaction pour la progression du cycle viral. Il est important de rappeler brièvement le rôle de la protéine pRb dans le cycle cellulaire.

La protéine pRb est le produit du gène suppresseur de tumeur RB1, situé sur le chromosome 13, et appelé le gène de susceptibilité au rétinoblastome. La pRb sous forme hypophosphorylée se lie au facteur de transcription E2F (protéine cellulaire indispensable sous forme active à l'entrée de la cellule en phase S et, à l'activation de certains gènes impliqués dans la réplication de l'ADN), et ce pendant les phases G0 et G1 du cycle cellulaire, permettant de maintenir E2F sous forme de répresseur translationnel. Au moment de l'entrée de la cellule en phase S, la protéine pRB est phosphorylée (au niveau de certains résidus sérine) par les complexes cyclines-cdk, en particulier les complexes cyclines D-cdk4 et cycline E-cdk2, permettant ainsi la libération du facteur E2F qui peut alors activer les gènes nécessaires à la prolifération cellulaire (Dell et Gaston, 2001; Funk *et al.*, 1997).

Une fois exprimée dans la cellule de l'hôte, la protéine E7 vient perturber le contrôle de E2F, cette perturbation se traduit par la liaison entre E7 et pRb qui va conduire à la dégradation de pRb via le protéasome 26S. Le mécanisme moléculaire de cette dégradation est controversé et peu connu. La dégradation de pRb a pour effet de libérer le domaine d'activation transcriptionnelle de l'hétérodimère E2F/DP (Morris *et al.*, 1993; Melillo *et al.*, 1994) et par conséquent permettre l'entrée en phase S. En effet, dans la cellule, l'interaction des « *pocket domain proteins* » pRB, p107 ou p130 avec E2F/DP masque le domaine d'activation transcriptionnel de l'hétérodimère E2F/DP ce qui bloque strictement l'entrée en phase S (Weintraub *et al.*, 1995; McLaughlin-Drubin *et al.*, 2008). Il a été décrit que l'oncoprotéine E7 est responsable, en inactivant pRb, de l'augmentation de la transcription du gène codant pour un inhibiteur de cdk, p16 (inhibe particulièrement cdk4 et cdk6) impliqué dans l'arrêt du cycle cellulaire en G1. (Ben-Saadon *et al.*, 2004). Néanmoins, les cellules continuent à proliférer malgré le niveau d'expression élevé de p16. Les lignées cellulaires dérivées de cancers du col de l'utérus sont résistantes aux signaux d'arrêt de croissance du TGF (Transforming Growth Factor, une cytokine, sensée limiter la croissance des cellules épithéliales) (Polyak, 1996). Ceci suggère que la capacité de E7 à inactiver p21^{CIP1} et p27^{KIP1} (des inhibiteurs de kinases dépendantes de cyclines) lève l'inhibition exercée par ces celles-ci sur E2F, et puisse contribuer à l'acquisition de la résistance au TGF- β (pour revue, Munger *et al.*, 2001).

La liaison entre E7 et les protéines à "pocket domain", pRb, p107 et p130 est médiée par le motif LXCXE de la région CR2, mais les deux régions du N-terminus (CR1 et CR2) sont nécessaires pour la dégradation de pRb. La protéine E7 à "haut risque" lie mieux pRb que celle à "bas risque". Des mutations conçues dans le motif LXCXE ont affecté la liaison de E7/p53 et le pouvoir d'immortalisation de E7 (Flores *et al.*, 2000). D'autre part, des séquences supplémentaires dans la région du C-terminus de E7 sont nécessaires pour la rupture des complexes E2F/pRb (Helt et Galloway, 2001). Des expériences de substitution de l'extrémité C-terminale de E7 par des régions structurellement semblables provenant de E6, montrent que les protéines E7 mutantes sont incapables de rompre les complexes E2F/pRb mais conservent leur capacité transformante (Mavromatis *et al.*, 1997; Braspenning *et al.*, 1998). Ces données suggèrent que la capacité à rompre le complexe E2F/pRb et l'activité transformante de E7 ne sont pas nécessairement liées. À l'inverse, de nombreux mutants non transformants de HPV16 E7 conservent leur potentiel à interagir avec pRb et à activer efficacement la transcription des promoteurs sensibles à E2F (Banks *et al.*, 1990; Jewers *et al.*, 1992). Il demeure que la liaison entre E7 et les protéines à « pocket domain » entraîne la déstabilisation de ces dernières. Étant donné que les inhibiteurs du protéasome

S26 interfèrent avec la dégradation de pRb entraînée par E7, et que HPV16 E7 peut interagir avec la sous-unité S4 du protéasome S26. Il a été proposé que E7 pouvait induire la dégradation de pRb en le ciblant directement vers le protéasome (Berezutskaya et Bagchi, 1997), mais toutefois, un mutant de E7 déficient dans la liaison à S4 continue de déstabiliser efficacement pRb (Gonzalez *et al.*, 2001). Ceci laisse suggérer d'autres mécanismes d'action.

Dans le cas de HPV 16, il a récemment été démontré que cette dégradation impliquerait la formation d'un complexe entre E7 et l'ubiquitine ligase cullin 2 (Huh *et al.*, 2007).

La protéine E7 peut interagir et activer le facteur de transcription induit par INF, IRF-1 (interferon regulatory factor), le régulateur transcriptionnel de l'interféron β (Park *et al.*, 2000; Um *et al.*, 2002). E7 interagit également avec l'histone déacétylase Mi2 β , le facteur de transcription MMP2, et AP-1 qui contribuent à la suppression de l'activité transcriptionnelle de IRF-1 (Antinore *et al.*, 1996; Park *et al.*, 2000). E7 est capable également de déréguler au niveau des cellules transformées, des voies métaboliques essentielles. Elle interagit avec l'enzyme glycolytique pyruvate kinase (PK) M2, et l' α -glucosidase, deux enzymes impliquées dans le métabolisme du glycogène, ces interactions contribuent à la perte des stocks intracellulaires de glycogène (Zwerschke et Jansen-Durr, 2000).

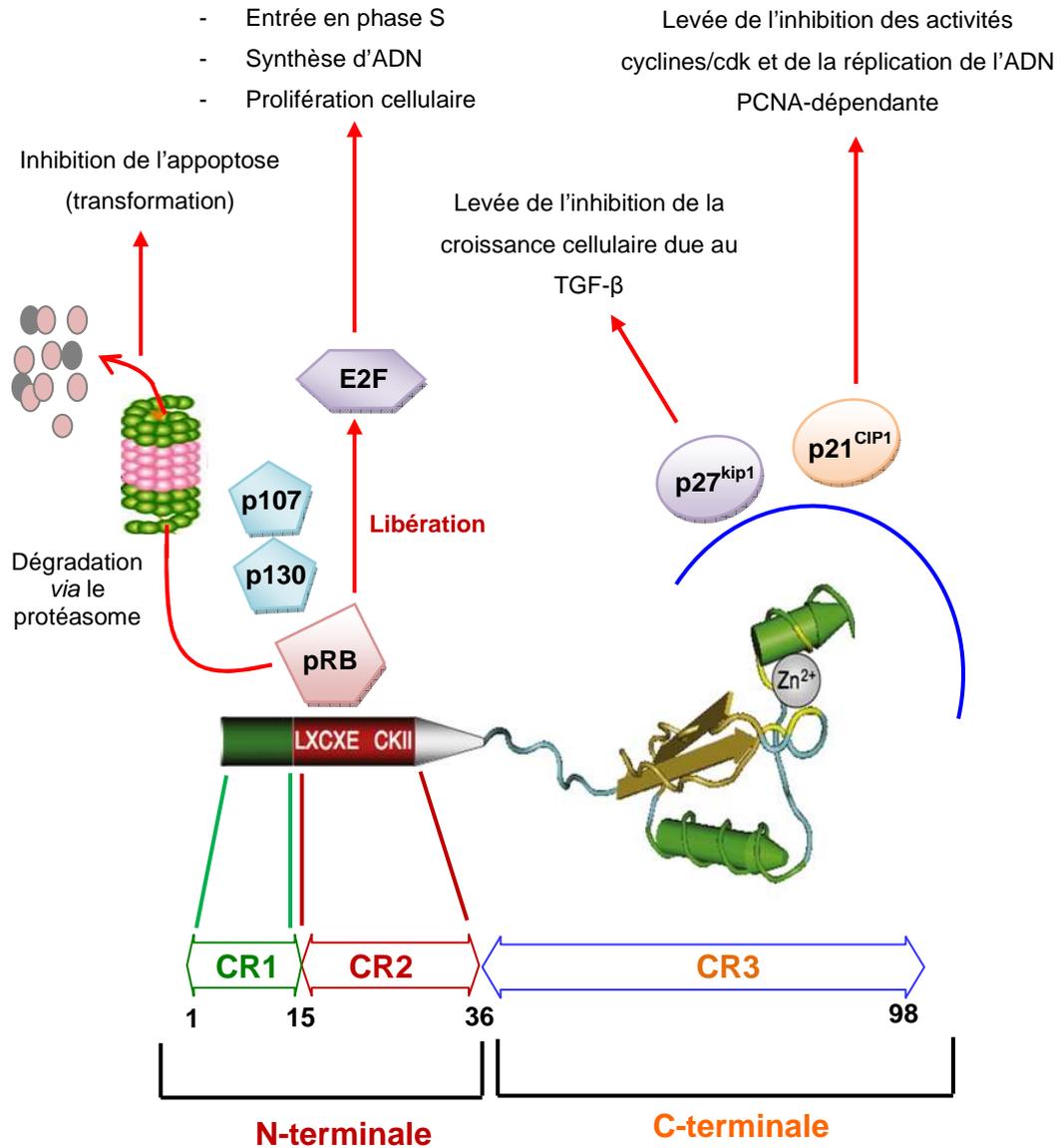


Figure 13: Représentation schématique de la protéine HPV16 E7. La protéine HPV16 E7 comprend un domaine N-terminal et un domaine C-terminal. Les régions CR1, CR2 et CR3 sont représentées sous forme de flèches (D'après McLaughlin-Drubin et Munger, 2008). La région CR2 comprend le motif LXCXE. Cette figure illustre également l'interaction de la protéine E7 avec certaines de ces cibles cellulaires, ainsi que leur rôle dans la transformation et l'immortalisation de la cellule hôte.

Caractérisations biochimiques de E7

HPV E7 est un petit polypeptide plutôt acide et composé d'environ 100 acides aminés. L'oncoprotéine E7 est constituée de deux domaines principaux, le domaine N-terminal et le domaine C-terminal. Le domaine N-terminal contient deux régions très conservées, CR1 et CR2, qui sont essentielles aux propriétés transformantes de E7. Le domaine C-terminal est constitué de la région CR3 qui contient deux motifs en doigt de zinc (Barbosa *et al.*, 1990) impliqués dans la dimérisation/multimérisation de la protéine E7 (Clemens *et al.*, 1995).

Le domaine N-terminal de la protéine E7 partage une petite région de similarité de séquence avec la protéine E1A d'Adénovirus et de l'antigène-T des polyomavirus, qui partagent la remarquable capacité de modifier la régulation transcriptionnelle normale des certains gènes essentiels dans le cycle cellulaire (Figge *et al.*, 1988; Phelps *et al.*, 1988). Les deux site de liaison au zinc se retrouvent sur le domaine C-terminal, sous la forme de deux motifs de Cys-X-X-Cys séparés de 29 résidus (Barbosa *et al.*, 1990). La région CR1 présente une homologie entre les HPV à "haut risque" et les HPV à "bas risque". Elle a été identifiée comme un site de liaison de p600 (Huh *et al.*, 2005; DeMasi *et al.*, 2005). L'autre région CR2, très conservée parmi les HPV, est extrêmement importante parce qu'elle contient un motif consensus Leu-X-Cys-X-Glu (LXCXE). Il s'agit d'un motif de liaison aux régions appelées "pocket domains" présentes dans la protéine pRb et les protéines p107 et p130 apparentées à pRb (Dyson *et al.*, 1989; Lee *et al.*, 1998; Münger *et al.*, 2001) (Figure 13). La région CR2 possède également un site de phosphorylation par la caséine kinase II, qui est essentiel puisque de la phosphorylation d'E7 dépend son activité transformante (Firzlaff *et al.*, 1989; Barbosa *et al.*, 1990; Münger *et al.*, 2001). La région CR3 de E7 intervient dans l'inactivation des inhibiteurs de kinases dépendants des cyclines (CDKI) p27 et p21. Elle interagit également avec un grand nombre de facteurs de transcription comme Mi2 β (Brehm *et al.*, 1999) et TBP (TATA banding protein) (Massimi *et al.*, 1997).

La structure tridimensionnelle à haute résolution de E7 a été résolue pour la région CR3 (région structurée) par cristallographie (Lui *et al.*, 2006) et en solution par RMN (Ohlenschläger *et al.*, 2006). Le domaine N-terminal, non replié, n'est pas présent dans la structure proposée de E7. le domaine C-terminale (CR3) est un dimère obligatoire, de « topologie » $\beta 1$ - $\beta 2$ - $\alpha 1$ - $\beta 3$ - $\alpha 2$ pour chaque protomère. La dimérisation se fait à travers les deux hélices $\alpha 1$ / $\alpha 1'$ et un feuillet- β intramoléculaire (Ohlenschläger *et al.*, 2006) (Figure 14).

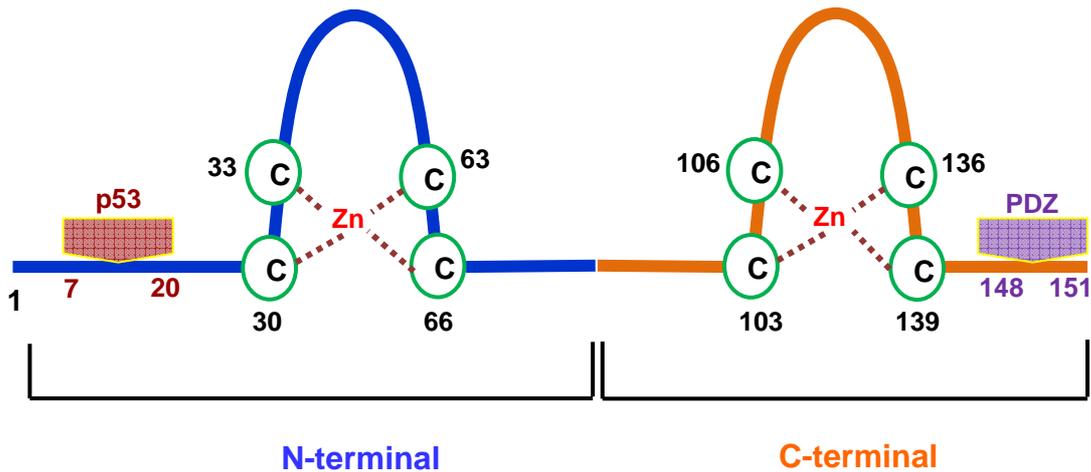


Figure 15: Représentation schématique de la protéine HPV16 E6. Les deux domaines liant le zinc, N-terminal (bleu) et C-terminal (orange). Les zones d'interaction avec les principales cibles cellulaires de E6 sont également mentionnées.

HPV16Rct	CYSLYGTTLLEQQYNKPLCDLLIRRCINCOKPLCPPEKQRLHLDKQRFHNI. RGRWTGRCMSCCRSSR. . . T. . RRETQL
HPV18Rct	SDSVYGGDTLEKLTNTGLYNLIRCLRCQKPLNPAEKLRLHNEKRRFHNI. AGHYRQCHSCCNRAQRRLQRRRETQV
HPV34ct	NQSVYGRLENLTKQLCNILIRCGKCKPLCPLEKQRHVDEKRRFHQI. ADQWTGRCTQCWRP. SATVV
HPV26ct	TCSVYGATLEALTKKSLCNLLIRCHRCQMPGLGPEKQRIVDEKRRFHEI. AGQWGLCTNCRPRRQ. TETQV
HPV53ct	NCSVYGASLEALTKKSLSDLIRCYRCQHPFTPBEKQLHCDYKRRFHKI. SHMWTGSCCLTCWRHTTA. TBSAV
HPV40ct	RYAAAYAPTVEEBTGLTILQVHIRCCCKHRPLSPVEKTNHIVKKTQPFKL. KDSWTOYCLHCWKKCKMEKQR. . . SETLC
HPV27ct	HYSYCGDVTVEETGTGIPIPQLPMRCYIICHKPLSWEEKBALLVGNKRFHNI. SGRWTGHCMQCGSTCTAPDPA. . SRTLH
HPV11Rct	NYAAAYAPTVEEBTGNEDILKVLIRCYLCKPLCRIEKLHILGKARFIKL. NNQWKGRCCLHCWTTTCMEDLLP.
HPV72ct	TYSYCGPTVEQHTGKSLAELVIRCHACCKPLSCQEKVEYQVQTGIHFHFI. SGLWTGRCCQCRGACTARNQP.
HPV10ct	DYSYVEGVEERTKQSIYTTQLIRCYMCKPLVREKDRHRNERRRLHKI. SGYWRGSCHEYCWSRCTVRIPO.
HPV32ct	DRSAFWHTVEQHTGLLLEEQIIRCAICQKPLSPSEKDHHTYNGRRHFFI. LNRWTGRCTQCEE.
HPV12ct	QQTVLGRDIELATGKSIFFDLKIRCTCLSFLDITIEKLDSCGRQLPFHKV. ADRWKGICRQCKHLYLNDR.
HPV20ct	ESTVLRDIEQVYTKSVFDDVRCYTCMKPLDSIEKLDICGHRKRPYLV. RGSWKGICRLCEHFQ.
HPV17ct	EQSVSRELEEI SHKPIGEIPIRCRFLKLDLLEKLDTCYRHQQPHKV. RRRWVGLCRHCGSIG.
HPV23ct	QLTVYGREISQEQRPVIGQICIRCYVCLKSLDLIEKLDICSFNQPFHKV. RNNHWGRCRHKETE.
HPV1Act	QESYEVPEIEELDRPFLQIELRCTCIKLSVAEKLEVVSNGERVHRV. RNRLLAKCSLCLRYAI.
HPV48ct	VCTAKSHLLTGLVKEELSDINIRCOHCYSFLDYLEKLYHLYNDVDFLLI. RGTWRGVCNRCISHEGE.
HPV16Rnt	dPQERPRKLP. . QLCTELQTTIHDIILECVYCRQQLLRREEVYDFAFRDLQIVYR. DGNPYAVCDKCLAFYSEISEYRHY.
HPV18Rnt	dPTRRPFYKLP. . DLCTELNTSLQDIEITCVYCKTVLELTVPEFAPKDLFVYVR. DSIPHAACHKCIDFYSRIRELRHY.
HPV34nt	nPEERPFYKLP. . ALCEEVNIISHIELEDCVYCEQLYRCEVYDFIFRDLQVYVR. KGGPLGVCPQCLLFPYSEKVRQYRHY.
HPV26nt	dPRERFRTLH. . ELCESLNTTLQNLQVQCVCYKBTQWADVYVFAICDLRVYVR. DRSPYAACRCRCVIFYSKITEYRHY.
HPV53nt	nTEERFRTLH. . QLCEVVKPPLLELQVCFCKEALTASEVYNFAYTDLRVYVR. DGYYPGVCKFCLLFPYSEKVRKLEY.
HPV40nt	rCGSQARTLY. . ELCDQCNTLPTLQIDCVFCKTVLKTAEVLAFAFRRELYVVR. DDFPHAACPRCLDLHGKVNQYRNF.
HPV27nt	eENPCFRNIF. . LLCKQYGLSLEDLRLLCVYCRRLSDADVLAFAIKELSVVVR. KGFPPGACCKCLIAAGELRQYRHW.
HPV11Rnt	dASTSATSID. . QLCKTFNLSLHTLQIQCVFCRNALTAETIYAYAYENLEVVVR. DNFPFAACACCLELQQRINQYRHF.
HPV72nt	ngLHNPTNIW. . LLCKEIEVDLEDLITCIPCKNELTTEELLAIAIKELQIVVR. DNWPFQVPCAPCLARATRVRELYW.
HPV10nt	MGAQEFRNIL. . LLCRNCGIPLEDLRLCCIFCTEQLTAAELAFALRELYLVVR. AGVPYQACARCLLQGIIVRRLKYW.
HPV32nt	eASSQPSTLY. . QLCKDFGLTLRNLQICCIWCKHHLTSAEAYAYHFKDLRVVVK. KGFPPYAACAFCLFYSSEVCLBHY.
HPV12nt	nTPELPTTIK. . ELADLLDIPLVDCVPCNFCGFVLDPLVECDFDEKQLTLIWK. GHFVTACCRSCCAATAYEFNEFY.
HPV20nt	lEPPLPATIC. . GLAKLEIPLDCLIPCNFCGNFLTHLEVCFDEKRLTIWK. DHLVFAACCRVCCSAYATAYEFNQFY.
HPV17nt	MDRPKPQTVR. . ELADTLICIPVDILLPCFNCNRFLLAYIELVAFDLKGLQIWIWTEEDFVFACCSSCAYATAQYEFKQFY.
HPV23nt	MDSTRFLTQV. . QLSDELTVPVVDLLPFCFCSRFLTYLELREFDYKHLQIWIWTEEDFVFACCSGCAYASAQFIEIQFY.
HPV1Ant	.MATPIRTVR. . QLSSELCIPYIDVLLPCNFCNWFYLSNAEKLLFDHFDLHLVVR. DNLVFPCCQGCARTVSLLEFVLY.
HPV48nt	MEPQPTDLD. . SYCKYFNISPFDELVKCIPCKFVSIVDLASPHMKRLSVIWR. DNTPFACCTKCLRLTALYEKDNFP.

Figure 16: Alignements des séquences de la protéine E6 de différents types de HPV: Le domaine C-terminal (en haut) d'E6C et le domaine N-terminal (en bas). L'alignement montre les cystéines conservées dans tous les HPV (Nominé *et al.*, 2006).

La protéine E6

Le gène E6 est l'un des premiers gènes exprimés lors d'une infection par HPV, et joue un rôle capital dans l'immortalisation de cellules infectées. L'oncoprotéine E6 ne possède pas d'activité enzymatique intrinsèque mais exerce ses fonctions en interagissant avec de nombreuses protéines cellulaires.

La protéine E6 des HPV est une petite protéine de 151 acides aminés (19kDa), contenant deux domaines E6-N et E6-C. Les deux domaines contiennent chacun deux motifs de cystéines hautement conservés (Cys-x-x-Cys), qui sont impliquées directement dans la liaison du zinc (Cole et Danos, 1987; Barbosa *et al.*, 1989; Vousden *et al.*, 1989; Grossman et Laimins, 1989; Kanda *et al.*, 1991) (Figures 15, 16).

La protéine E6, ses fonctions et ses cibles

Durant l'infection par HPV, E6 joue de multiples rôles interférant avec de nombreuses voies de signalisation cellulaire, avec pour résultat net la mise en place d'un environnement favorable à la réplication de l'ADN viral et la neutralisation concomitante des contrôles de surveillance qui sont activés dans les cellules infectées, obligées de relancer la réplication de l'ADN.

C'est au milieu des années 1980 que la protéine E6 est identifiée comme une oncoprotéine virale, impliquée dans le processus de carcinogenèse (Androphy *et al.*, 1987; Banks *et al.*, 1987). E6 induit différentes modifications dans la cellule hôte. En synergie avec l'oncoprotéine E7, E6 des HPV à "haut risque" est capable d'immortaliser des cellules épithéliales et des fibroblastes humains (Munger *et al.*, 1989; Rapp et Chen, 1998), alors que la protéine E6 des HPV à "bas risque" n'en est pas capable. Ceci semble notamment lié à leur différence d'affinité pour la p53, E6 des HPV à "haut risque" ayant une affinité plus importante pour la p53 (Werness *et al.*, 1990; Li et Coffino, 1996). E6 interagit avec un grand nombre de protéines cellulaires. A ce jour, plus d'une quarantaine de cibles ont été découvertes (Figure 17).

La plus importante cible de E6, et la première à avoir été identifiée est la protéine pro-apoptotique et suppresseur de tumeur p53.

L'interaction E6, p53 et E6AP

La protéine p53 est une phosphoprotéine de 393 acides aminés. Codée par un gène localisé sur le chromosome 17, elle constitue le principal obstacle à la réplication virale, elle peut entraîner l'arrêt du cycle cellulaire ou l'apoptose des cellules infectées. p53 est constituée de 4 régions qui représentent: le domaine de transactivation transcriptionnelle, le domaine riche en proline, le domaine central, le domaine d'oligomérisation et le domaine C-terminal basique. Cette protéine est activée par différents stimuli (stress cellulaire, lésions de l'ADN, ...). L'expression de p53 a pour effet le blocage du cycle cellulaire en phase G1 et G2 permettant la réparation de l'ADN, la réparation directe de l'ADN endommagé, ou bien l'induction de l'apoptose (mort cellulaire programmée) lorsque la réparation n'est plus possible.

Pour dégrader p53, E6 recrute E6AP (E6-associated protein). E6AP est une protéine E3 ubiquitine-ligase (Schwarz *et al.*, 1998) qui accepte l'ubiquitine d'une enzyme E2 de conjugaison (E2 ubiquitin-conjugating enzyme) sous la forme d'un thioester puis transfère directement (grâce à son domaine C-terminale HECT) cette ubiquitine sur p53, générant sa polyubiquitination (Schwarz *et al.*, 1998). La polyubiquitination de p53 semble être la voie principale de sa dégradation par le protéasome S26 dans les cellules exprimant E6 (Figure 18). La dégradation de p53 dépend largement de la force de l'interaction E6-E6AP et entraîne bien évidemment la perte de l'activité de p53.

Cette perte d'activité de p53 a pour conséquence la perte de la capacité de la cellule à induire l'apoptose et à réparer les lésions de l'ADN, générant une instabilité génétique et donc une augmentation du risque de transformation maligne.

Néanmoins, dans les cellules HPV-positives exprimant E6, p53 n'est pas totalement dégradée (Cooper *et al.*, 1993, Lie *et al.*, 1999), ce qui implique l'existence de mécanismes contrôlant l'activité de E6. L'une des hypothèses avancées est l'existence de polypeptides de E6 résultant d'épissages alternatifs (E6*). Ces formes de E6 ont la capacité d'interagir avec le complexe E6-E6AP de la sorte qui bloque le recrutement de p53 et donc inhibe sa dégradation (Pim *et al.*, 1997). Outre la voie de dégradation de p53 dépendante de E6AP, E6 peut interférer par d'autres voies avec l'activité de p53. Il a été démontré que E6 interagirait de manière compétitive avec le coactivateur CBP/p300 (CREB Binding Protein/p300) sur le site de liaison à p53 provoquant la suppression de l'activité transcriptionnelle de p53 (Zimmermann *et al.*, 1999).

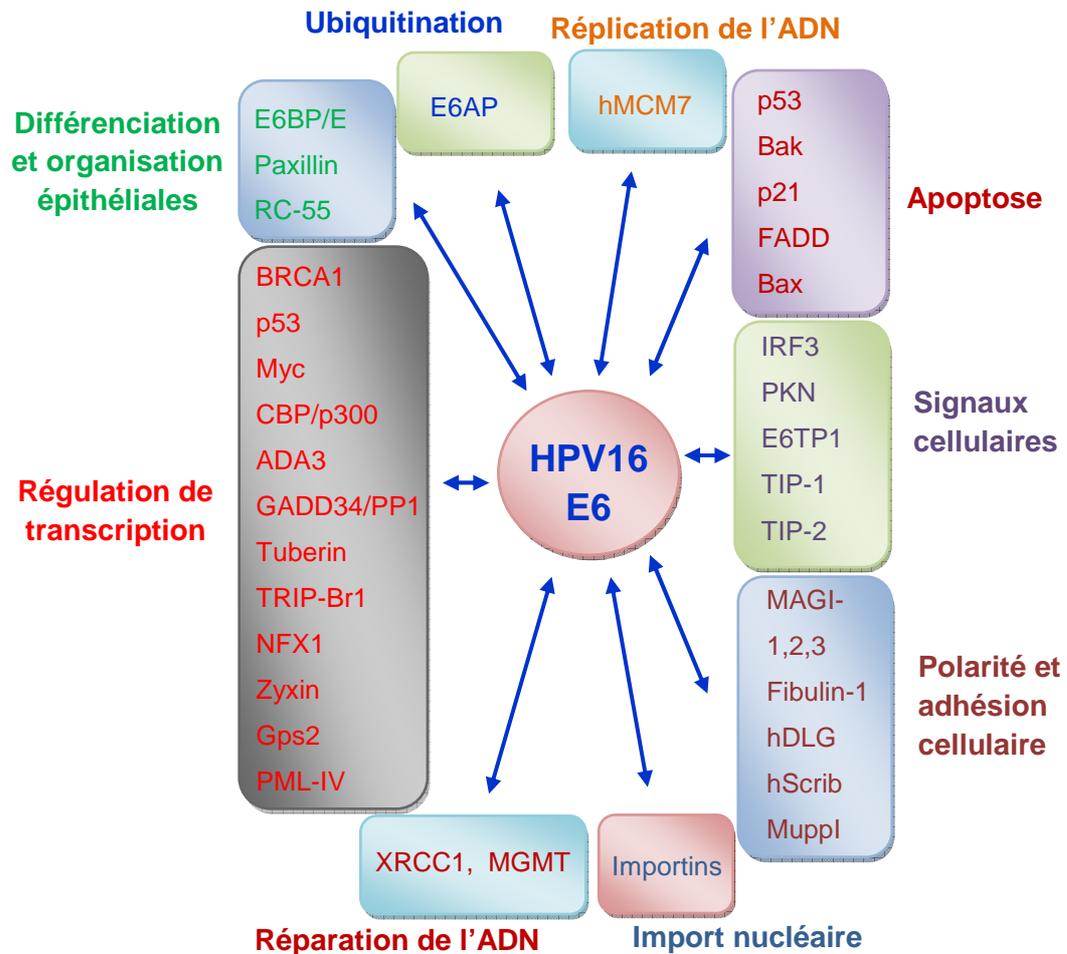


Figure 17: Protéines cellulaires, ciblées par l'oncoprotéine HPV16 E6
(D'après la revue, Tungteakkhun et Duerksen-Hughes, 2007).

Les autres cibles de E6

- E6 et Bak

Bak est une protéine pro-apoptotique impliquée dans un point critique du contrôle cellulaire (la mort cellulaire physiologique programmée). La protéine E6 des HPV anogénitaux à "haut risque" est capable de lier la protéine pro-apoptotique Bak et stimule sa dégradation par la voie de protéasome (Thomas et Banks, 1999). Il a été suggéré que l'induction de cette dégradation passe par l'interaction E6/E6AP (Thomas et Banks, 1998). L'interaction entre E6 et Bak indique que les HPV anogénitaux et certaines souches cutanées sont capables d'inhiber l'apoptose par au moins deux voies, p53-dépendante et p53-indépendante, cette dernière impliquant le Bak (Jackson *et al.*, 2000).

- E6 et Paxilline

La paxilline est une protéine multi-domaines localisée dans les sites d'attachement des cellules à la matrice extracellulaire par les adhésions focales. L'adhésion focale forme une structure reliant la matrice extracellulaire et le cytosquelette d'actine et un important site pour la transduction du signal. Seules les protéines E6 de BPV à "haut risque" sont capables d'interagir avec la paxilline mais sans promouvoir sa dégradation. Cette sélectivité suggère une corrélation entre liaison de E6/Paxilline et le pouvoir transformant de E6 (Tong et Howley, 1997). E6 lie la paxilline *via* un motif consensus (LDXLLXL). L'interaction E6/Paxilline est responsable la rupture de cytosquelette d'actine, ce qui est critique pour les fonctions de la cellules et caractéristique pour plusieurs cellules transformées. (Vande Pol *et al.*, 1998).

- E6 et les protéines à domaines PDZ

Le domaine C-terminal de E6 à "haut risque" présente une séquence consensus d'acides aminés X[T/S]X[L/V] très conservé, qui semble jouer un rôle important pour la transformation des cellules (Crook *et al.*, 1991; Kiyono *et al.*, 1997). Cette séquence est un motif de liaison aux protéines à domaines PDZ (Doyle *et al.*, 1996). Ces protéines font partie de différentes familles fonctionnelles impliquées dans la formation de canaux ioniques, la signalisation et l'adhésion. Les domaines PDZ (Psd-95, Dlg, et ZO1) sont des régions d'interaction protéine-protéine constituées de 80 à 90 acides aminés comprenant six feuillets- β et deux hélices- α , compactés en une structure globulaire.

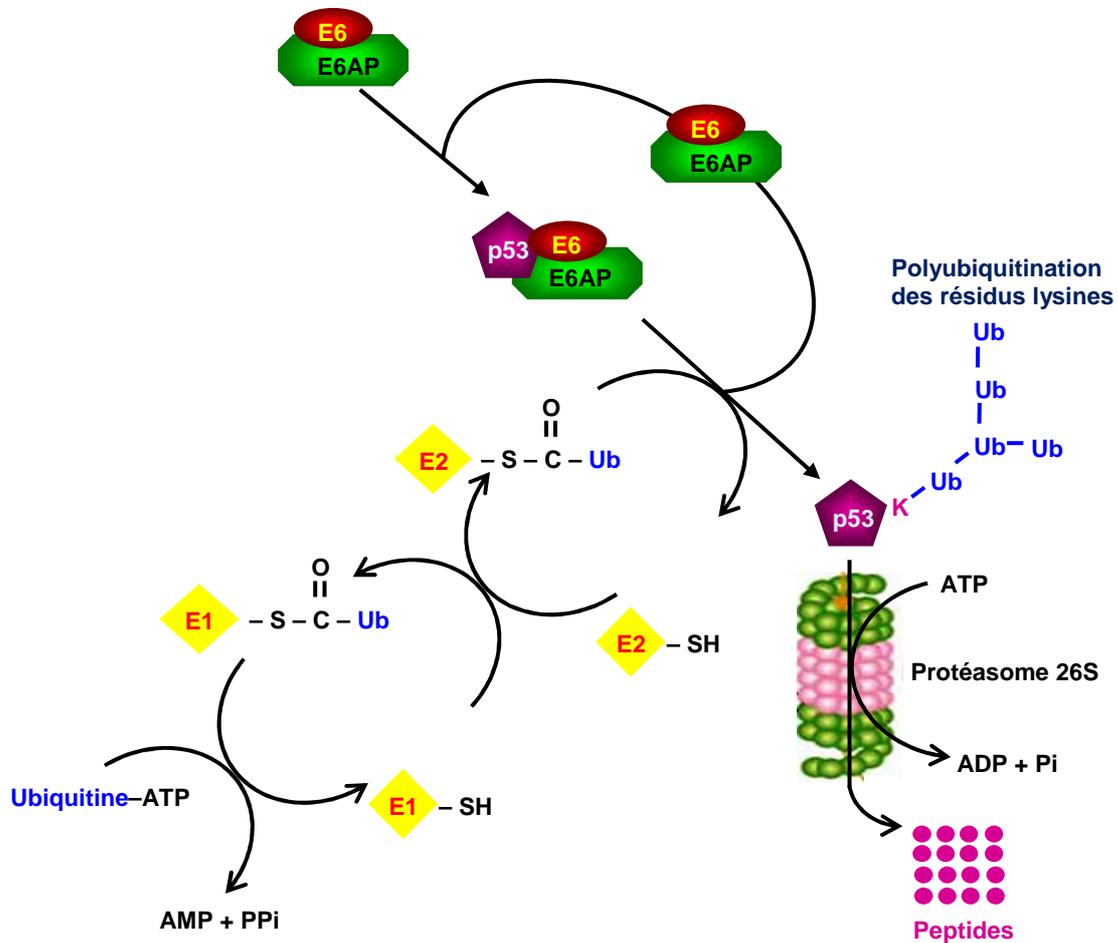


Figure 18: représentation schématique de la dégradation de p53 médiée par l'oncoprotéine E6. La dégradation de p53 nécessite la formation d'un complexe trimérique entre E6, p53 et l'ubiquitine ligase E6AP, qui permet la poly-ubiquitination de p53. Ce processus fait également intervenir deux autres protéines : une protéine activatrice de l'ubiquitine E1 et une enzyme de conjugaison de l'ubiquitine E2. La protéine p53 poly-ubiquitinée est dégradée par le protéasome 26s. La poly-ubiquitination de p53 ainsi que sa dégradation sont des processus ATP-dépendantes (d'après Howley *et al.*, 2006).

Ainsi, E6 lie des domaines PDZ des protéines MUPP-1 (Lee *et al.*, 2000), hDLG (kiyono *et al.*, 1997), hSCRIB (Nakagawa et Huibregtse, 2000) et MAGI-1, 2 et 3 (Thomas *et al.*, 2002; Zhang *et al.*, 2007), déclenchant le plus souvent leur dégradation via le protéasome. Ces interactions sont importantes pour la progression tumorale (Nguyen *et al.*, 2003).

- **E6 et E6-BP**

Les E6 de HPV génitaux à "haut risque" ainsi que la E6 du BPV1 semblent également interagir avec la protéine E6-BP (E6 binding protein) ou ERC-55 (Chen *et al.*, 1995). E6-BP est une protéine liant le calcium, localisée au niveau du réticulum endoplasmique (Chen *et al.*, 1995). La liaison E6-E6BP pourrait inhiber le stade tardif de différenciation des cellules épithéliales (Sherman *et al.*, 2002).

- **E6 et TNF-1**

Le TNF (*tumor necrosis factor*) est capable d'induire l'apoptose dans les cellules infectées par les virus. La protéine HPV16 E6 semble inhiber certains mécanismes de la mort cellulaire en se fixant au niveau du TNF-R1 (récepteur pour le TNF); qui se trouve ainsi privé de ses interactions avec ses partenaires cellulaires et cesse alors de transmettre le signal pro-apoptique (Filippova *et al.*, 2002). Cette interaction interfère notamment avec la capacité de TRADD (*TNF-R1-associated death domain*) d'interagir avec FADD (*Fas-associated death domain*) empêchant ainsi le recrutement de la pro-caspase-8, qui déclenche la cascade protéolytique des caspases conduisant à l'apoptose (Filippova *et al.*, 2009).

- **E6 et E6TP-1**

E6 est capable d'induire la dégradation de E6TP-1 (E6 targeted protein 1 : homologue aux protéines GAP (GTPase activating protein) de la famille Rap) inhibant ainsi son rôle de régulateur négatif sur la multiplication cellulaire. La dégradation de E6TP-1 est E6AP-dépendante et implique le protéasome (Gao *et al.*, 2002). Cette interaction semble également exclusive pour les E6 à "haut risque".

- **E6 et c-Myc**

Une autre cible de E6 est la protéine c-Myc, un facteur de transcription impliqué dans la prolifération cellulaire et présentant aussi un rôle pro-apoptique. E6 stimule sa dégradation via E6AP (Gross-Mesilaty *et al.*, 1998). Or, en parallèle, E7 inhibe la régulation transcripionnelle négative exercée par TGF- β sur c-Myc. Les activités des deux

oncoprotéines semblent donc entrer en contradiction. Cependant, il a été démontré par ailleurs que la dérégulation de l'expression de c-Myc dans les cellules épithéliales différenciées conduit vers l'apoptose (Askew *et al.*, 1991). on peut donc penser que la dégradation de c-Myc par E6 contribue à maintenir un équilibre entre le maintien de la prolifération par E7 et écartement de l'apoptose par E6.

- **E6 et la télomérase (hTERT)**

La télomérase est une enzyme responsable de la réplication de télomères (les éléments de l'ADN situés aux extrémités des chromosomes). Elle est composée d'une sous-unité d'ARN (utilisée comme matrice de réplication) et de la sous-unité protéique hTERT (enzyme portant l'activité transcriptase reverse). Il a été démontré que E6 active la sous-unité catalytique hTERT (Klingelutz *et al.*, 1996). Des mutants de E6 qui augmentent l'activité de la télomérase sont capables d'induire l'immortalisation cellulaire, alors que les mutants qui ne l'augmentent pas perdent la capacité d'immortaliser les cellules (Greider, 1999). Curieusement, des mutants qui ne dégradent pas p53 mais qui gardent leur capacité d'immortalisation, agissent tout de même sur la télomérase en augmentant son activité. Ceci semble confirmer l'importance de l'activation de la télomérase qui apparaît, de façon générale, comme une étape clé dans la transformation des cellules malignes et le développement du cancer chez l'homme (Holt *et al.*, 1996; Plug-DeMaggio *et al.*, 2004).

- **E6 et acides nucléiques: ADN (jonctions de Holliday) et ARN (transcrits E6-E7)**

En sus des nombreuses interactions engagées par E6 avec des cibles protéiques, il a été montré que E6 est capable, dans des expériences menées *in vitro*, de se lier à des motifs structuraux particuliers présents au niveau des jonctions de Holliday, ADN à quatre brins impliqué notamment dans les processus de recombinaison génétique (Ristriani *et al.*, 2001). Seules les protéines E6 à "haut risques" sont capables de lier ces structures d'ADN. Les hypothèses qui ont été avancées explicitent que les jonctions de Holliday sont des ligands naturels de E6 *in vivo*, et, que l'interaction entre les deux partenaires interfère avec les évènements de recombinaisons dans la cellule. Contrairement aux protéines E6 à "bas risque", les E6 à "haut risque" présentent une similarité de séquences avec des endonucléases VII. Cette similarité suggère une ressemblance de la fonction qui pourrait expliquer interaction E6/ADN-jonctions de Holliday (Ristriani *et al.*, 2001, Nominé *et al.*, 2006).

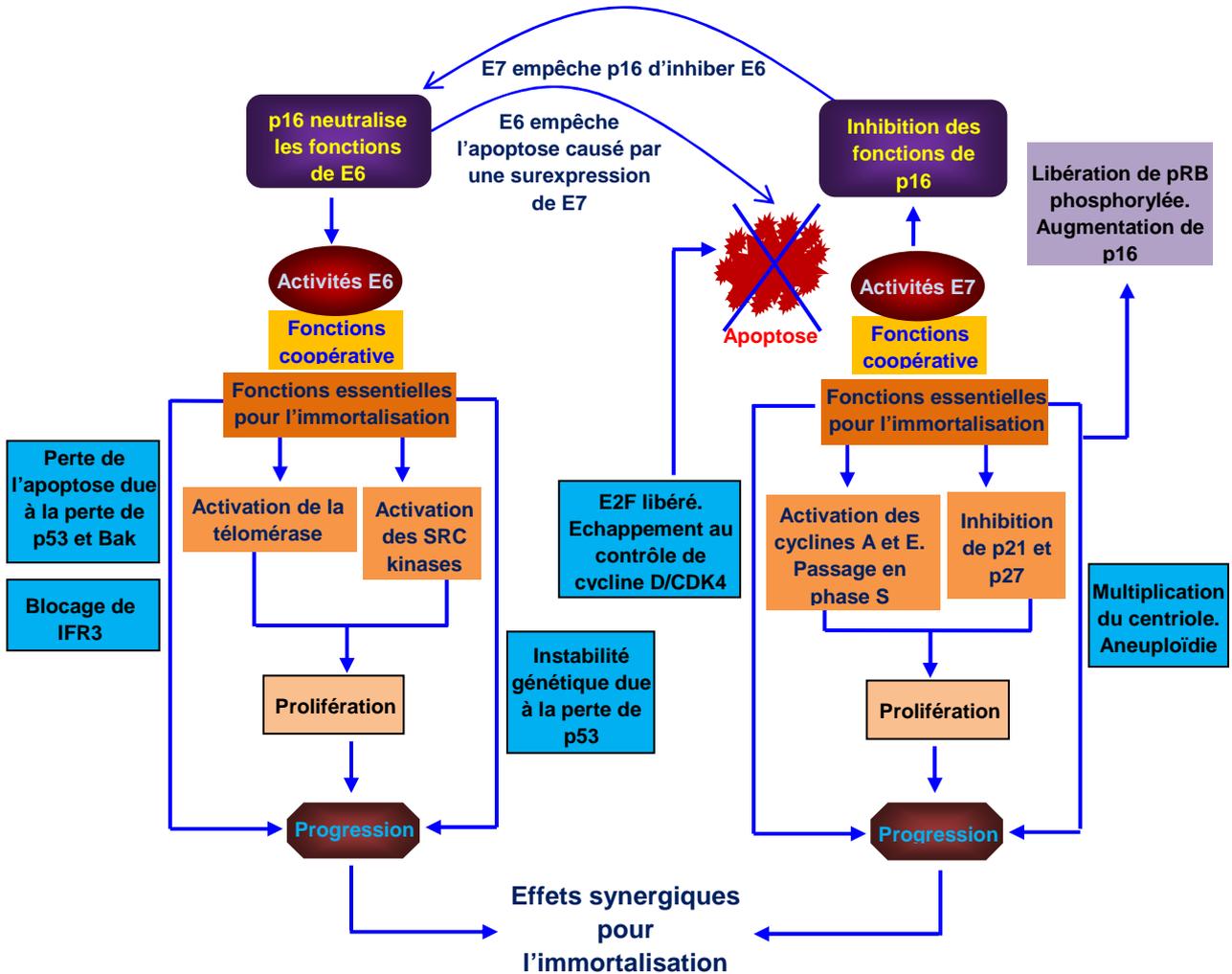


Figure 19: Fonctions des oncoprotéines E6 et E7 et leurs contribution dans les étapes conduisant à la transformation et l'immortalisation. E6 active la télomérase et les SRC kinases et inhibe p53 et BAK. E7 inhibe pRb qui libère E2F dans les cellules hôtes permettant l'augmentation de l'expression de p16. E7 inhibe également l'activité de p16. E7 semble activer les cyclines A et E et inhiber les CKI p21 et p27. E6 et E7 agissent en synergie pour l'immortalisation et la transformation (d'après zur Hausen, 2002).

Il a été récemment démontré que E6 interagit *in vitro* avec des transcrits de E6-E7 issues de HPV16 (HPV16 E6-E7 pré-ARNm), et que cette interaction pourrait jouer un rôle dans la régulation post-transcriptionnel durant l'infection virale (Bodaghi *et al.*, 2009). La délétion d'une partie N-terminale (102 résidus) n'affecte pas l'interaction E6/ARN, alors que le domaine N-terminal seul est déficient pour cette interaction. Ce qui prouve que les 48 résidus C-terminaux de E6 sont responsables de l'interaction protéine/ARNm (Bodaghi *et al.*, 2009). Cette localisation corrèle avec ce qui a été attribué pour l'interaction E6/ADN-jonctions de Holliday (Nominé *et al.*, 2003; Nominé *et al.*, 2006).

Conclusion sur les fonctions complémentaires des oncoprotéines E6 et E7

Plusieurs fonctions ont été décrites pour E6 et E7. La protéine E6 dégrade p53 et Bak pour procurer la résistance à l'apoptose et augmente aussi l'instabilité chromosomique. En outre, l'activation de la télomérase et l'inhibition de la dégradation des kinases de famille SRC par E6 semblent accomplir des fonctions importantes dans la stimulation de croissance. En effet, les deux oncoprotéines E6 et E7 sont capables indépendamment d'immortaliser des cellules humaines, mais, leur expression combinée conduit à des effets complémentaires et synergétiques qui augmentent leur efficacité de transformation. Ceci peut être expliqué par le suivant: E6 semble être altéré par p16, tandis que E7 inhibe l'action de p16 en activant directement les cyclins A et E. En contre partie, E6, empêche l'apoptose induit par E7 en dégradant les protéines p53 et Bak (zur hausen, 2002) (Figure 19).

Caractérisations biochimiques de E6

Dans les cellules, la protéine E6 est retrouvée majoritairement dans le noyau, mais elle se localise aussi dans le cytoplasme (Masson *et al.*, 2003). Il a été démontré chez certaines souches de HPV à "haut risque" la présence d'un isoforme de E6 résultant d'un épissage alternatif de l'ARNm (Doorbar *et al.*, 1990). Cette chimère appelée E6* est une E6 tronquée qui code pour les résidus 1 – 43 dans le cas de HPV-18 (Pim et Banks, 1999). Elle est capable de lier le complexe E6-E6AP et ainsi bloquer l'interaction avec p53 (Pim *et al.*, 1997) et donc la dégradation de p53.

La séquence de E6 était connue depuis les années 1980, mais la caractérisation biochimique de la protéine restait très difficile. Or, les études poursuivies ces vingt dernières

années ont permis progressivement d'élucider ce vaste champ de recherche en allant de la production vers la description structurale de la protéine.

Production et Purification de E6

Échecs et tentatives d'expression de la protéine

Cette petite protéine avait longtemps posé de sérieux problèmes d'expression et de purification. E6 était en effet systématiquement insoluble ou hautement agrégée, indépendamment du système d'expression utilisé (Imai *et al.*, 1989; Androphy *et al.*, 1987; Grossman *et al.*, 1989).

Les premiers essais ont commencé en 1987. Certains auteurs avaient testé l'expression de la protéine E6 de HPV16 (recombinante) dans des souches thermorésistantes d'*Escherichia coli* (induites à 42°C). Les agrégats insolubles de la protéine étaient extraits en milieu dénaturant en présence d'urée (Androphy *et al.*, 1987). Les mêmes résultats ont été observés suite à l'expression de E6 en utilisant le système d'expression de baculovirus infectant des cellules d'insectes. En effet, la protéine a été également associée à la fraction insoluble dans les cellules (Grossman *et al.*, 1989). Exprimée en bactérie (*E. coli*), la protéine E6 a été isolée à partir des débris cellulaires sous forme de corps d'inclusions insolubles. Dans certains cas, l'utilisation des agents dénaturants (guanidine-HCL, Urée) a été la seule solution pour solubiliser la protéine, mettant en doute certaines analyses fonctionnelles effectuées avec ces échantillons (Imai *et al.*, 1989). Plusieurs approches similaires ont été ensuite testées pour l'expression et la purification de la protéine E6 (Stacey *et al.*, 1992; Park *et al.*, 1993). Mais la protéine E6 restait impurifiable sous forme soluble. Cependant, la production de E6 en fusion avec des autres protéines telles que la GST (*Glutathion S-Transferase*) et la MBP (*Maltose Binding Protein*), connues pour leurs effets stabilisants (Kapust et Waugh, 1999; Fox *et al.*, 2000), constitua une ouverture remarquable. La fusion GST-E6 exprimée dans des bactéries (*E. coli*), a pu être produite après des cycles de dénaturation/renaturation sous forme stable avec concentration relativement importante (0,03 – 1 mg/ml). Toutefois, le clivage entre les deux protéines conduisait à la précipitation immédiate de E6 (Lechner et Laimins, 1994).

Des échantillons de protéine HPV16 E6 munie d'un his-tag ont également pu être purifiés à partir des cultures bactériennes sous forme soluble et active biologiquement, sans avoir recours aux méthodes de dénaturation/renaturation. Ce résultat apparaissait comme un

succès, mais curieusement, les méthodes et protocoles utilisés pour cette purification n'ont pas été détaillés (Daniels *et al.*, 1997), et ces travaux n'ont apparemment donné lieu à aucune suite.

La purification de E6 HPV16 soluble:

Le terme « corps d'inclusions solubles » a été introduit par notre laboratoire lorsque des constructions de E6 en fusion avec GST ou MBP, couplés avec l'optimisation des conditions pour l'expression et la purification ont permis de solubiliser la protéine E6 (Nominé *et al.*, 2001a). Toutefois, les préparations de protéines de fusion MBP-E6 contenaient une majorité d'agrégats solubles, de taille relativement homogène et contenant des agglomérats de protéines E6 biologiquement inactives solubilisés par MBP (les "corps d'inclusions solubles"), et une partie monodisperse contenant des protéines E6 bien repliées (Nominé *et al.*, 2001a).

Ensuite, notre laboratoire a montré que la présence de nombreux résidus cystéines dans la protéine E6 (14 cystéines pour HPV16) pourrait conduire à la formation de ponts disulfures intermoléculaires (Ristriani *et al.*, 2000). L'analyse et l'alignement des séquences pour les différentes protéines HPV ont révélé que la protéine E6 du HPV16 contient six cystéines non conservés (Figure 16). Ces six cystéines ont été substituées par des résidus serines, donnant lieu à un nouveau mutant de E6 actif biologiquement et soluble jusqu'à une concentration-limite d'environ 20 μ M (Nominé *et al.*, 2001b; Nominé *et al.*, 2003). Au-dessus de cette concentration, l'agrégation de la protéine (mutée, nommée E6 6C/6S), conduisait lentement vers la formation de structures en formes de fibres (Zanier *et al.*, données non publiées).

Les mêmes auteurs ont montré que les deux domaines de E6 (E6-N et E6-C) sont capables de se replier séparément (Nominé *et al.*, 2003). Le domaine C-terminal nommé E6-C 4C/4S (moitié C-terminale de E6 4C/4S) a été purifié avec succès sous forme de monomère stable à une concentration très élevée (1 mM). Sa structure en solution a été résolue par la RMN (Résonance Magnétique Nucléaire) (Nominé *et al.*, 2003; Nominé *et al.*, 2005). Au contraire, le domaine E6-N 2C/2S précipitait au cours de différentes étapes de purification, et ne supportait pas les hautes concentrations (Nominé *et al.*, 2005).

En parallèle, l'équipe de Michael Cordingley (Boehringer-Ingelheim, Canada), a également mené des travaux importants sur la caractérisation biochimique de la protéine E6. Cette équipe a isolé un mutant de la protéine E6 (Δ 143–151) qui conservait sa capacité de dégradation de p53 et était purifiable à l'état monomère à basse concentration (Lipari *et al.*,

2001). Cette construction exprimée en fusion avec la GST gardait la solubilité après clivage de la GST sur une résine de glutathion (Lipari *et al.*, 2001). Cette équipe est également parvenue à isoler et purifier des échantillons largement solubles d'un domaine N-terminal de E6 (2–77 acides aminés). Ce domaine N-terminal concentré (> 100 μ M) apparaissait sous forme d'un dimère (avec un constante de dissociation de l'ordre de 0,11 mM à 0,1 M NaCl). La structure secondaire de chaque construction a été calculée par des mesures de dichroïsme circulaire montrant: 26% des hélices- α et 27% de feuillets- β pour la E6 (Δ 143–151) et 25% hélices- α et 29% de feuillets- β pour le domaine E6 N-terminal (2 – 77) (Lipari *et al.*, 2001).

La prédiction de la structure secondaire des protéines E6

En l'absence de données structurales sur les protéines E6, en 1996, une prédiction de la structure secondaire de deux oncoprotéines E6 et E7 a été proposée. Des logiciels d'alignement de séquences ainsi que les méthodes communément utilisées pour la prédiction des structures (Chou-Fasman, SAPIENS, PHD et GOR I et II) ont été utilisés pour analyser et valider la prédiction sur l'ensemble de 57 protéines E6. Les résultats obtenus étaient les suivants :

- 32 résidus de E6 hautement conservés dans les deux domaines présentaient un pourcentage de 90% de conservation.
- Une similarité de séquence entre les deux domaines E6-N et E6-C, certains résidus conservant leurs positions dans les deux domaines, en particulier les cystéines impliquées dans la liaison du zinc.
- La première région N-terminale de E6 présentait un motif de $\alpha/\beta/\alpha/\beta$, (α 1– β 1– α 2– β 2), et un motif similaire était prédit pour la région C-terminale (α 3– β 4– α 4– β 5). Un feuillet- β (β 3) reliait les deux motifs.
- Chaque région (ou domaines) était initiée par les premières hélices- α (α 1 et α 3), alors que les deux feuillets- β (β 1 et β 4) se localisaient sur le premier motif (Cys-x-x-Cys). La deuxième paire α - β suivait la même localisation pour le deuxième motif liant le zinc de chaque domaine (Ullman *et al.*, 1996). Certains auteurs ont prétendu avoir expérimentalement obtenu des résultats confirmant cette prédiction (Lipari *et al.*, 2001). Mais, de récents travaux ayant mené à la résolution de la structure du domaine E6-C de la protéine HPV16 E6, ont montré que chaque domaine présentait une autre représentation des éléments de la structure secondaires (β 1– α 1– α 2– β 2 – β 3) (Nominé *et al.*, 2006).

Structure de la protéine E6

À ce jour, la seule structure connue pour la protéine E6 est celle d'un mutant du domaine C-terminal de HPV 16 E6 (E6-C 4C/4S), résolue par notre laboratoire. Sur la base de l'homologie entre le domaine E6-C et le domaine E6-N, un modèle a été proposé pour la protéine entière sous forme d'un pseudo-dimère de E6-N/E6-C (Nominé *et al.*, 2006). La région repliée du domaine E6-C 4C/4S commence par le résidu tyrosine 81 et termine avec l'arginine 141 (Nominé *et al.*, 2003). La structure tridimensionnelle du domaine révèle la présence de 3 feuillets- β en parallèles (S1, S2, S3) avec deux hélices- α courtes (H1, H2). Le site de fixation du zinc est entouré d'un coté d'une longue boucle (L1) reliant H1 à H2, et de l'autre coté d'une courte hélice- α en C-terminal du domaine (Figure 20). L'attachement du zinc vers le noyau principal du repliement est soutenu par plusieurs résidus enfouis dans le repliement: isoleucine 101, lysine 115, leucine 110, et l'arginine 135. Les quatre cystéines conservées 103, 106, 136 et 139 sont impliquées directement dans la liaison aux ions du zinc (Zn^{2+}) (la structure est accessible dans la banque de données PDB sous la référence: 2FK4) (Nominé *et al.*, 2006).

En se basant sur les données structurales du domaine E6-C, un modèle probable du repliement de E6-N a été proposé, dans lequel les résidus prévus pour être enfouis au cœur du repliement représentent des hydrophobes strictement conservés. Les domaines E6-N et E6-C présentent une certaine homologie malgré de légères différences. Certains résidus hydrophobes exposés sur la surface de E6-C sont conservés dans les deux domaines. L'hypothèse plausible avancée par Nominé *et al.* (2006) suggère que cette région hydrophobe représente l'interface d'interaction entre les deux domaines. Cette surface serait moins exposée dans le contexte de la protéine complète. Les auteurs proposaient donc un modèle représentant un arrangement possible des deux domaines dans la protéine E6 entière, basé sur l'enfouissement des surfaces hydrophobes des deux domaines ainsi que la complémentarité de leurs structures secondaires. Ces deux critères n'étaient remplis que dans le cas où les deux domaines étaient placés symétriquement d'une façon pseudodimérique (Nominé *et al.*, 2006). Dans ce modèle, les deux feuillets- β S1 de E6-N et de E6C, sont arrangés d'une façon antiparallèle donnant lieu aux six feuillets- β avec les deux hélices H1 enveloppant les feuillets. Les sites de fixation du zinc sont localisés dans les deux cotés opposés du pseudo-dimère (Nominé *et al.*, 2006) (Figure 21).

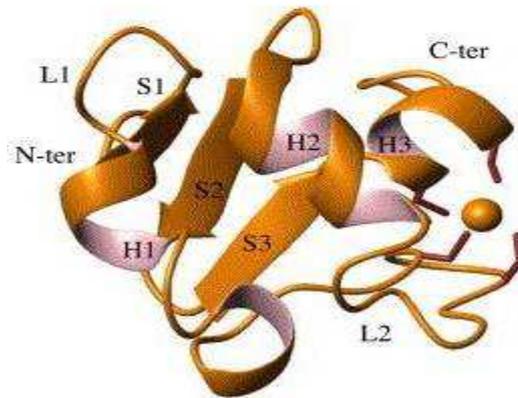


Figure 20: Structure 3D du domaine C-terminal (aa 87-141) de la protéine HPV16 E6. Les feuillets- β (S1, S2, S3), les deux hélix- α (H1, H2), la liaison avec le zinc implique une boucle très longue (L2) qui relie H1 et H2 à un court hélix- α (H3) dans l'extrémité C-terminale (C-ter) (Nominé *et al.*, 2006).

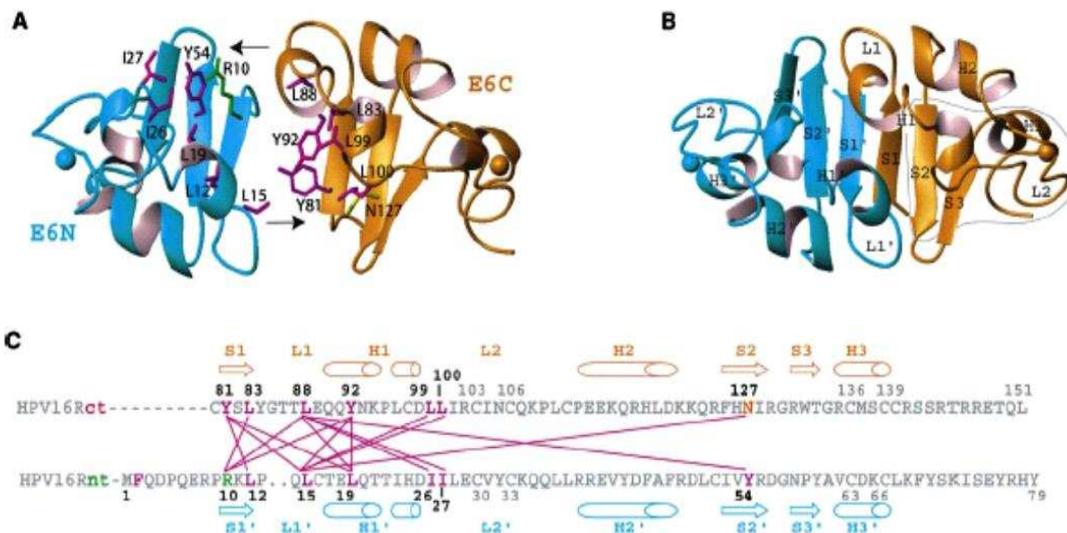


Figure 21: Modèle pseudodimérique de E6: (A) Modèle de E6-N et structure du E6-C, les faces hydrophobes sont situées dans l'interface. **(B)** Modèle de l'arrangement du pseudodimère entre les deux domaines EN-N et E6C. **(C)** Carte détaillant les résidus hydrophobes de la région de l'interface entre les deux domaines (Nominé *et al.*, 2006).

Présentation du projet de recherche

Les principaux enjeux de mon travail de thèse se résument en l'optimisation et la production d'échantillons de la protéine E6 de différents types de papillomavirus pour permettre la résolution de la structure tridimensionnelle de ces protéines. Dans ce travail, des analyses fonctionnelles focalisées avec les différentes constructions de E6 se sont également proposées.

La première partie consiste à compléter l'étude structurale de l'oncoprotéine E6 de HPV16. J'ai commencé d'abord à travailler sur les conditions d'expression et de purification afin d'adopter des protocoles optimisés pour la production de E6. Ensuite, j'ai cloné et purifié des constructions de domaines de E6 fusionnées avec MBP, aussi bien que la protéine entière. Ces constructions étaient conçues sur la base de mutagenèses dirigées contre des résidus responsables de l'agrégation, la dimérisation du domaine N-terminal ou bien des résidus impliqués dans la liaison E6/E6AP. Les démarches de purification m'ont permis de juger la qualité des échantillons du point de vue biochimique, mais les analyses biophysiques telles que la RMN qui a été réalisée par la spectroscopiste de l'équipe, Katia Zanier, nous ont permis de vérifier le repliement des polypeptides produits. Par RMN, nous avons également effectué des expériences de titration afin de cartographier les résidus de la protéine qui sont impliqués dans la dimérisation ou la liaison E6/E6AP. D'autre part, Charlotte Boulade-Ladame, doctorante dans l'équipe, a fait les tests *in vitro* et *in vivo* (dégradation de p53, interaction E6 avec E6AP) pour certains mutants que j'ai produits.

Dans la deuxième partie, notre étude s'est consacrée à la protéine E6 de papillomavirus bovin (BPV-1 6). J'ai utilisé les mêmes protocoles d'expression et de purification, ainsi que les méthodes biophysiques que nous avons décrites dans la première partie. À cela s'ajoute l'étude de l'affinité entre BPV-1 E6 et son principal ligand (la paxilline) que j'ai étudié par Biacore. Le résultat de Biacore m'a permis de constater une forte affinité entre les deux partenaires, et de développer par la suite différentes approches de co-expression et co-purification visant à solubiliser la protéine BPV-1 E6.

La troisième partie du travail concerne l'identification de nouvelles protéines précoces E6, E7 et E9 issues d'un nouveau papillomavirus aviaire (FIPV). J'ai cloné les séquences ORF codantes pour ces protéines dans des vecteurs d'expression chez *E. coli*. J'ai purifié aussi la protéine E6 "monodomaine" de FIPV ainsi que la partie repliée et structurée de la protéine E7. Ensuite, j'ai cloné les trois protéines dans des vecteurs

eucaryotes en fusion avec des tags permettant leur localisation cellulaire. En collaboration avec François Deryckère, les trois protéines E6, E7 et E9 ont été exprimées dans différentes souches cellulaires, ce qui a permis de vérifier leur localisation cellulaire.

Enfin, une dernière partie est consacrée à l'étude de l'oligomérisation et l'agrégation des protéines E6 de différents types de HPV ainsi que la protéine E6 de BPV-1. Ma participation dans cette partie se limite à la production des échantillons.

Ces quatre parties sont présentées successivement sous la forme de chapitres indépendants dans la section résultats.

RESULTATS

RESULTATS

Première partie : Production et caractérisation de la protéine E6 de HPV 16

La purification du domaine E6-N hautement soluble

Notre équipe avait précédemment tenté de produire une forme du domaine HPV16 E6N comportant deux cystéines mutées en sérine (E6-N 2C/2S), qui s'était avérée insuffisamment soluble. Au cours de ma thèse, nous avons pu montrer que le domaine E6-N "sauvage" présentait une meilleure solubilité.

Le domaine E6-N (E6-N "sauvage") est exprimé en fusion avec l'extrémité C-terminale de la protéine MBP (comme protéine porteuse). Les conditions de l'expression et la purification ont fait l'objet de multiples améliorations et d'optimisations (voir la section: matériels et méthodes). Nous avons purifié la fusion MBP-E6-N sous forme majoritairement monodisperse. Après une longue étape d'ultracentrifugation (15 heures à 38000 rpm), nous avons remarqué que la précipitation de la protéine est négligeable par rapport à ce que nous observions d'habitude pour nos échantillons de E6 6C/6S et E6-N 2C/2S. Le surnageant d'ultracentrifugation est soumis à la digestion par la protéase TEV pour séparer MBP de E6-N. Ensuite l'échantillon est injecté sur une colonne de gel-filtration; les deux composantes (MBP et E6) sont séparées selon leurs tailles. Enfin, les fractions correspondantes au domaine E6-N sont collectées. Le domaine E6-N "sauvage" purifié est très soluble, et nous avons vu qu'il supporte des concentrations de l'ordre de 300 μ M, relativement élevées pendant quelques semaines (à 8 °C) sans détecter au cun signe de précipitation. Les étapes de cette purification sont analysées sur un gel de Tris-Tricine-SDS utilisé pour séparer les protéines de petite taille (Figure: 22 – A).

Après avoir surmonté la difficulté de produire des échantillons solubles de E6-N, nous avons ensuite exprimé la fusion MBP-E6-N dans un milieu minimum (M9) permettant le marquage isotopique des protéines recombinantes, destinées à des études de RMN. Par conséquent, nous avons réussi à obtenir d'excellents échantillons marqués (15 N et 13 C) suffisamment concentrés (300 μ M) pour l'étude RMN (Figure: 22 – B).

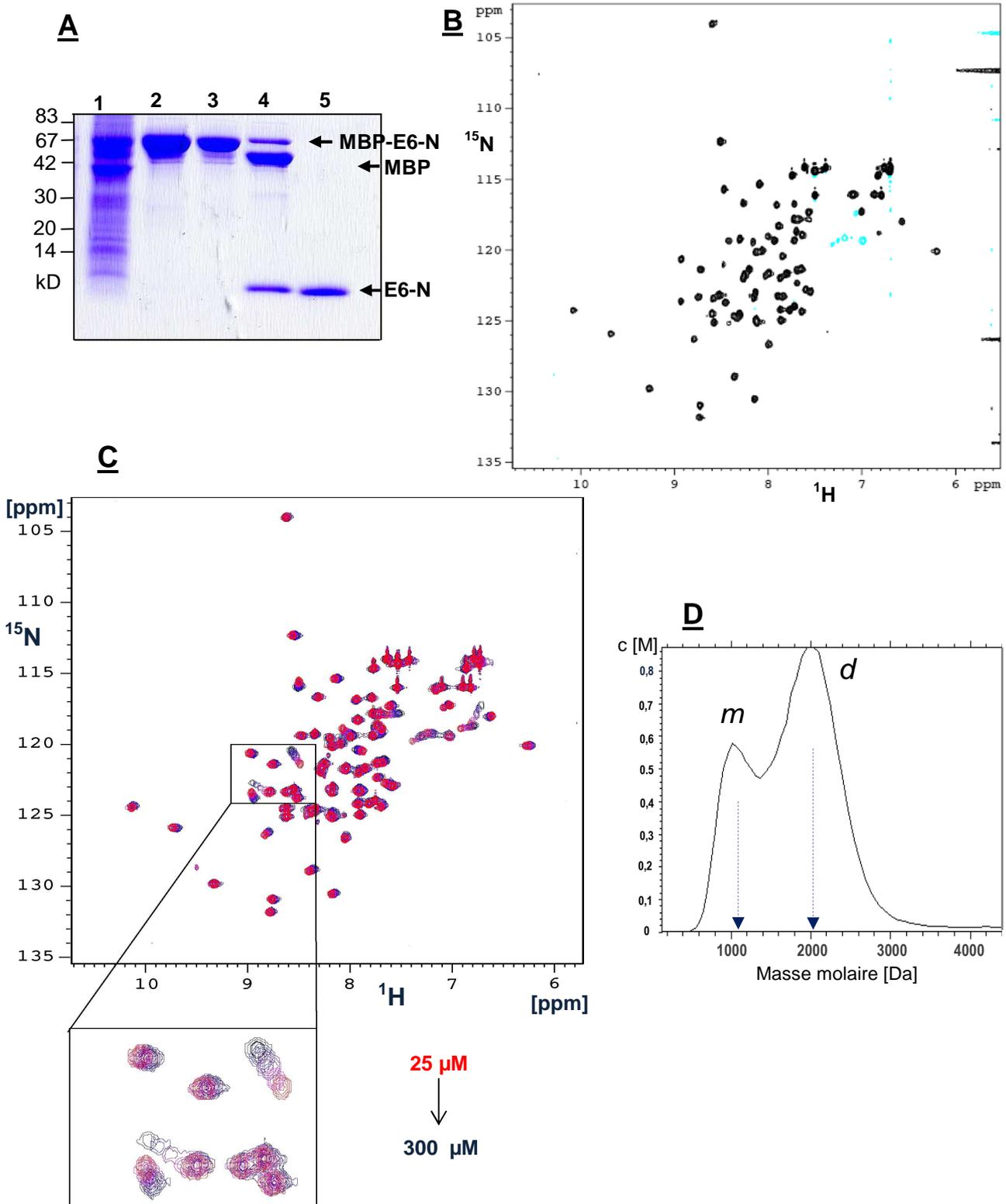


Figure 22. Purification et caractérisation du domaine E6-N: (A) gel Trice-Tricine montrant les principales étapes de purification. 1: MBP-E6 exprimée chez *E. coli* (BL21) est passée sur une colonne d'amylose. 2: l'élué de MBP-E6. 3: le surnageant de l'élué après ultracentrifugation. 4: coupure avec la TEV. 5: le domaine E6-N purifié et isolé après une gel-filtration. (B) le spectre RMN-HSQC (^{15}N , ^1H) d'un échantillon de E6-N (300 μM), le nombre de pics correspond au nombre exact des résidus du domaine (à l'exception de 3 prolines présents dans la séquence). (C) des variations de fréquences (^{15}N , ^1H) liées à la différence de concentration, la superposition de deux spectres HSQC du domaine E6-N (en rouge E6-N 25 μM et en bleu E6-N 300 μM). (D) le profil de sédimentation d'un échantillon de E6-N analysé par ultracentrifugation analytique, le domaine forme un dimère transitoire. *m*: monomère, *d*: dimère.

Toutes les expériences RMN permettant l'attribution des fréquences ^1H , ^{13}C et ^{15}N ainsi que la collecte des NOEs interprotons ont pu être effectuées et exploitées, aboutissant à la résolution de la structure du domaine E6-N (Zanier *et al.*, en préparation).

Identification d'une interface de dimérisation dans le domaine E6-N

Le spectre HSQC du domaine E6-N est d'une bonne qualité, montrant que le domaine isolé est stable et replié. Cependant, nous avons observé par RMN des variations de fréquence (^1H , ^{15}N) dépendantes de la concentration des échantillons (de 25 μM à 300 μM). Ces effets, qui concernent seulement quelques résidus regroupés sur une surface bien délimitée, mettent en évidence l'interface probable de dimérisation du domaine (Figure: 22 – C). La dimérisation du domaine E6-N a déjà été rapporté dans la littérature (Lipari *et al.*, 2001).

Pour approfondir la connaissance sur ce phénomène, nous avons analysé l'échantillon de E6-N par ultracentrifugation analytique. Ces mesures révèlent que le domaine E6-N forme un dimère transitoire avec une constante de dissociation (K_d) estimée à 34 μM (Figure: 22 – D). En plus des résultats obtenus par la RMN, nous nous sommes intéressés à la corrélation entre le phénomène de dimérisation et la présence de certains résidus initialement définis comme surface probable de dimérisation de E6-N. L'état actuel de l'attribution des séquences de E6-N a permis de repérer ces résidus, qui se sont avérés conservés dans les HPV génitaux à "haut risque". Les résidus concernés sont les suivants: H24, E41, Y43, D44, F47, R48.

Nous avons effectué une série de mutations dirigées sur les résidus de cette région dans le contexte du domaine E6-N de HPV16. Les résidus sont mutés ponctuellement pour analyser leurs contributions individuelles à la formation du dimère. Nous avons donc produit, sous forme de fusions à la MBP, puis purifié l'ensemble de ces mutants donnant les constructions:

- E6-N H24R, E41R, Y43R, F47R, F47E, R48A, F47E-Y43R, F47R-Y43E

Certains mutants semblent perdre la solubilité et précipitent lorsqu'ils sont séparés de la protéine solubilisante MBP. C'est par exemple le cas de E6-N E41R, ce qui suggère que le résidu E41, situé au début d'une hélice, est important pour le repliement du domaine. Les autres mutants de E6-N présentent un comportement similaire à celui de E6-N « sauvage » durant la purification. Parfois, la solubilité de E6-N est clairement augmentée, par exemple pour le mutant E6-N F47R. Le repliement de chaque mutant (de E6-N) est confirmé par leur spectre RMN HSQC. Nous avons ensuite vérifié par RMN une éventuelle variation de fréquences (^1H , ^{15}N) liée à la dimérisation de l'échantillon.

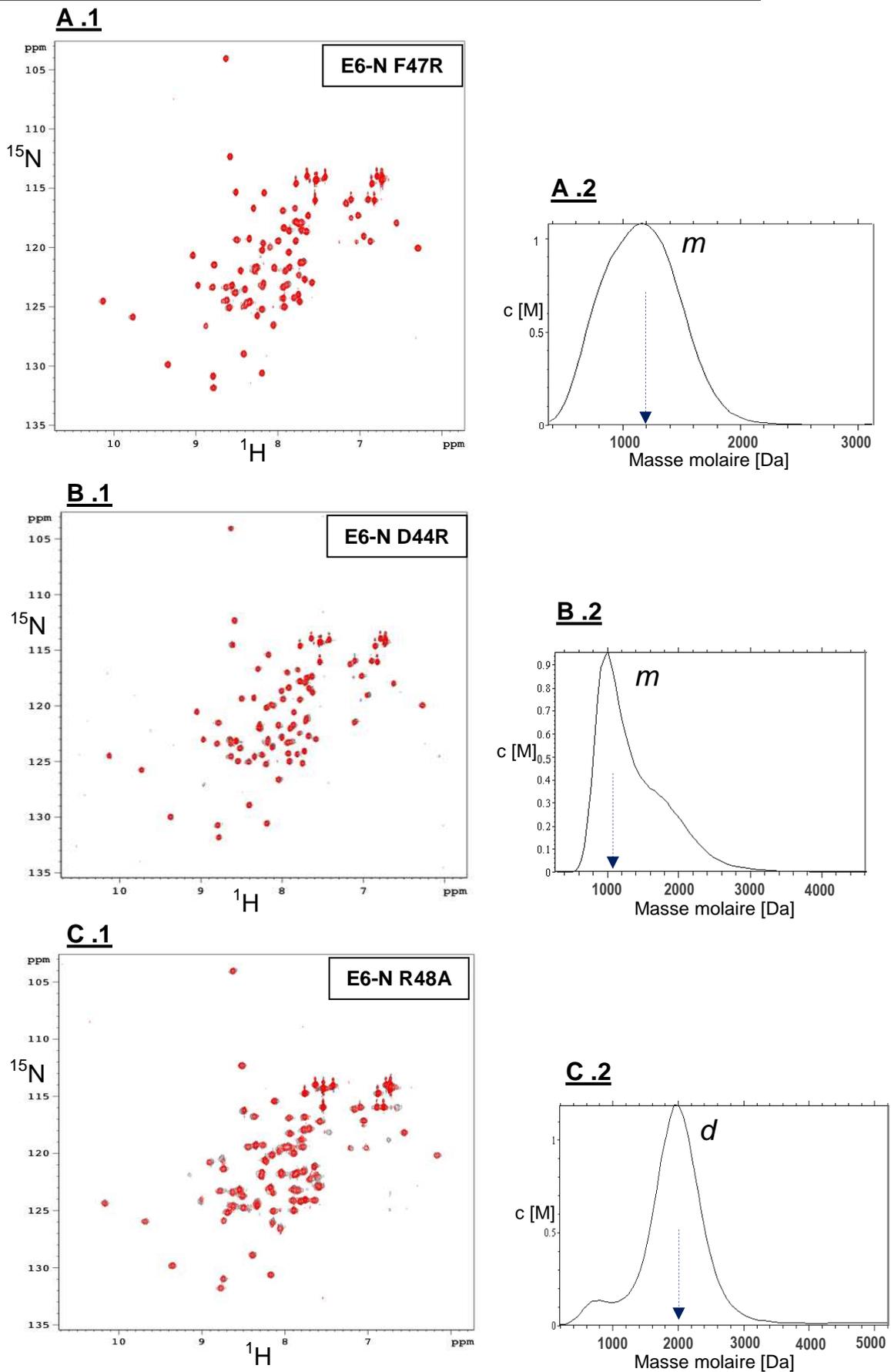


Figure 23. Des mutants monomères du domaine E6-N: (A .1, B .1 et C .1) La superposition des spectres HSQC montrant les variations de fréquences (^{15}N , ^1H) liées à la différence de concentration (en rouge E6-N 25 μM et en bleu E6-N 300 μM) pour chaque mutant. **(A .2, B .2 et C.2)** les profils de sédimentation des mutants analysés par ultracentrifugation analytique. *m*: monomère, *d*: dimère.

Les mesures réalisées avec deux concentrations: 25 μM et 300 μM . La superposition des spectres obtenus pour chaque concentration révèle que:

- les mutants E6-N F47R, E6-N D44R et E6-N H24R présentent les mêmes spectres indépendamment de la concentration (Figure: 23 – A.1 et B.1).
- remarquablement le mutant E6-N R48A présente une large différence entre ses deux spectres (Figure: 23 – C.1).

De plus, les profils de sédimentation, obtenus par l'ultracentrifugation analytique, montrent que certaines mutations favorisent un état monomérique de E6-N. Ces résultats sont concordants avec ceux obtenus par RMN. Ils montrent sans ambiguïté que les mutants E6-N F47R, E6-N D44R et E6-N H24R sont monomériques alors que le mutant E6-N R48A forme un dimère renforcé par rapport au domaine sauvage (Figure: 23 – A.2, B.2, C.2).

Les résultats qui vont suivre suggèrent que la dimérisation du domaine E6-N serait l'un des facteurs induisant l'agrégation de la protéine entière.

Production et caractérisation du domaine C-terminal de HPV E6

La structure du domaine C-terminal de HPV E6 avait été publiée en 2003 par notre laboratoire. Cependant, cette structure était celle du domaine E6-C 4C/4S comportant quatre mutations de cystéines non conservées en sérine. Le mutant E6-C 4C/4S était uniquement la seule forme soluble du domaine. Bien que la substitution de quatre cystéines non conservées dans l'extrémité C-terminal n'affecte pas la fonction de la protéine E6 complète (Nominé *et al.*, 2006), nous avons cherché à limiter le nombre des mutations dans la construction finale du domaine.

En optimisant les protocoles d'expression et de purification, nous avons réussi à obtenir des échantillons très solubles et très concentrés du domaine E6-C sous sa forme sauvage, exempte de mutations. Nous avons obtenu un spectre HSQC d'un échantillon monomérique de E6-C (300 μM) qui présente une dispersion homogène des pics, comparable à celle observée pour le spectre E6-C 4C/4S (Figure: 24).

Alors que les deux domaines sauvages (E6-N et E6-C) sont séparément très solubles, quand ils sont combinés dans la protéine entière, la protéine E6 sauvage est complètement agrégée et difficile à produire.

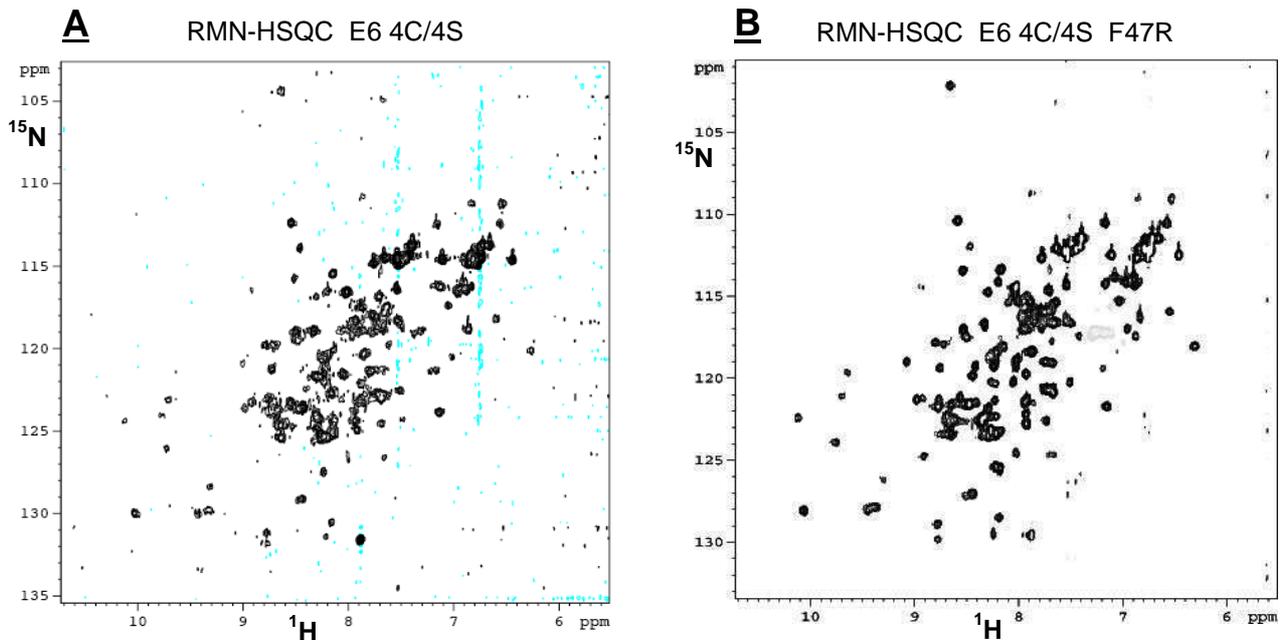


Figure 25. Les premiers spectres RMN-HSQC de la protéine E6 entière. (A) HSQC (^{15}N , ^1H) de la protéine E6 4C/4S (50 μM), les pics qui se répartissent sont rares. Le regroupement des pics au milieu du spectre signifie un mauvais repliement. (B) HSQC (^{15}N , ^1H) de la protéine E6 4C/4S F47R (350 μM), une meilleure distribution des pics représentant les aminoacides au sein du spectre. Quelques pics ont les mêmes fréquences dans les deux spectres de E6-N et E6-C 4C/4S. Vu la qualité du spectre et la concentration relativement importante des échantillons E6 4C/4S F47R, cette construction ouvre la voie vers plus de mesures RMN conduisant à la résolution de la structure de E6.

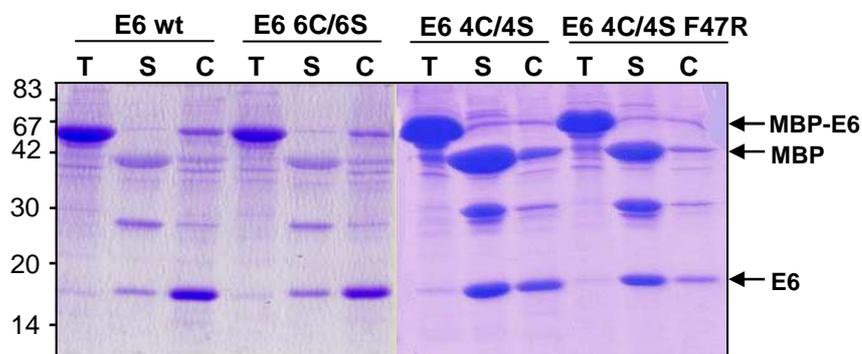


Figure 26. La solubilité de différentes constructions de E6: Les échantillons sont obtenus après une mini-purification en batch de MBP-E6. Les surnageants de la lyse des culots bactériens de 20 ml, sont incubés avec une résine amylose (300 μl), l'élution de MBP-E6 est digérée par la protéase TEV. Les concentrations de différents échantillons sont ajustées, et les produits de digestion de la TEV sont remis à une température de 8 $^{\circ}\text{C}$ pendant 15h. Ensuite les échantillons sont centrifugés pendant 30 min (13000 rpm) pour séparer les agrégats et les parties solubles. Les fractions (surnageant (S) et culot (C)) ainsi que la fusion MBP-E6 non digérée (fraction totale (T)) sont déposées et migrées sur un gel SDS-PAGE 15%. La fraction (T) représente l'élution de MBP-E6 non digérée. (S) le surnageant du produit de digestion contenant la partie soluble de E6. (C) le culot composé essentiellement des agrégats de E6. La comparaison entre les fractions solubles de E6 de différentes constructions, est significative d'une amélioration de la solubilité de E6. Le mutant E6 4C/4S F47R est le plus soluble parmi les autres constructions de E6 « sauvage », E6 6C/6S et E6 4C/4S.

Production et caractérisation de mutants hautement solubles de HPV E6 entière

La solubilité de la protéine E6 entière constituait le problème majeur limitant la caractérisation biophysique de la protéine. Le mutant E6 6C/6S a été purifié dans notre laboratoire, mais les échantillons les plus solubles commençaient à s'agréger à partir d'un certain seuil de concentration situé aux alentours de 10 à 20 μM . Ce seuil de concentration est inférieur à celui nécessaire pour l'étude par RMN.

Pour remédier à ce problème, nous avons d'abord construit un premier mutant combinant la forme "sauvage" du domaine N-terminal avec un domaine C-terminal 4C/4S. Ce mutant, nommé E6 4C/4S, présente une solubilité améliorée par rapport aux autres mutants précédents de E6. Le seuil de concentrabilité est repoussé (50 à 100 μM). Cependant, un premier spectre HSQC a montré que la qualité d'échantillon n'était pas optimale pour continuer avec les autres mesures nécessaires pour l'étude de la structure (Figure: 25 – A).

Dans un deuxième temps, nous avons conçu une série de nouvelles constructions de E6 combinant la forme 4C/4S du domaine C-terminal avec des formes du domaine N-terminal mutées ponctuellement dans la région de dimérisation. Plusieurs de ces mutants se sont avérés hautement solubles. Particulièrement, le mutant E6 4C/4S F47R obtenu a été le meilleur candidat à la résolution de structure 3D de E6 entière. Nous en avons obtenu un échantillon suffisamment concentré (350 μM). Par des expériences de RMN-HSQC, nous avons vérifié que tous les mutants conservent le repliement correct de la protéine.

Le spectre HSQC du mutant E6 4C/4S F47R est le meilleur spectre enregistré pour une protéine E6 complète. Les pics dénombrés correspondent quasiment au nombre d'acides aminés dans la protéine (Figure: 25 – B).

La figure 26 montre un test comparatif de solubilité pour quelques constructions qui ont constitué les étapes clés vers une meilleure construction en termes de solubilité. Les protéines E6 fusionnées avec MBP sont obtenues par mini-purification en batch sur une résine amylose (voir matériels et méthodes). Les échantillons sont restés pendant 15 heures à 8 °C en présence de la TEV, ensuite centrifugés pour séparer les deux parties agrégée et soluble. Les fractions (totale, surnageant et culot) sont déposées et migrées sur un gel SDS-PAGE 15%. La comparaison entre les fractions solubles de E6 (surnageant) de différentes constructions, révèle clairement que le mutant E6 4C/4S F47R est le plus soluble parmi les autres constructions antérieures (E6 « sauvage », E6 6C/6S et E6 4C/4S) (Figure: 26).

Le nombre de cystéines mutées restant important dans cette nouvelle construction (E6 4C/4S F47R), nous avons cherché à le réduire au minimum. Nous avons construit les différents mutants possibles, contenant des combinaisons allant d'une simple mutation (Cys – Ser) à un triplet de cystéines mutées.

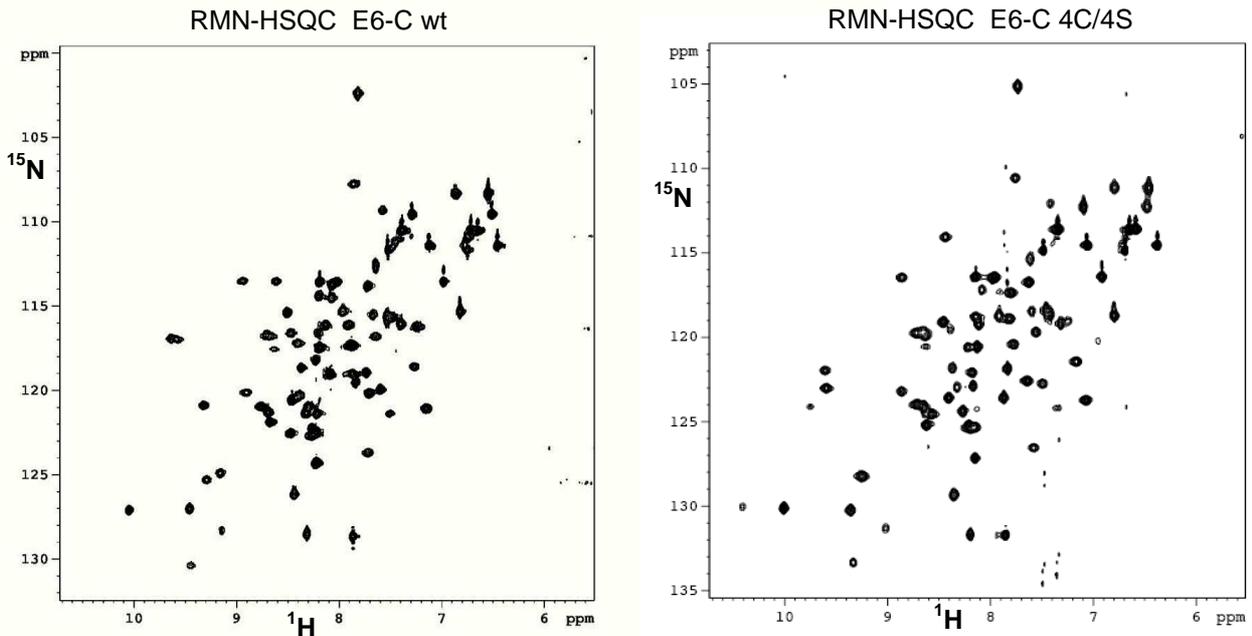


Figure 24. Le spectre RMN-HSQC (^{15}N , ^1H) du domaine E6-C de HPV16: à gauche le spectre de E6-C « sauvage » comparé avec celui de E6-C 4C/4S (à droite). Les deux spectres sont superposables.

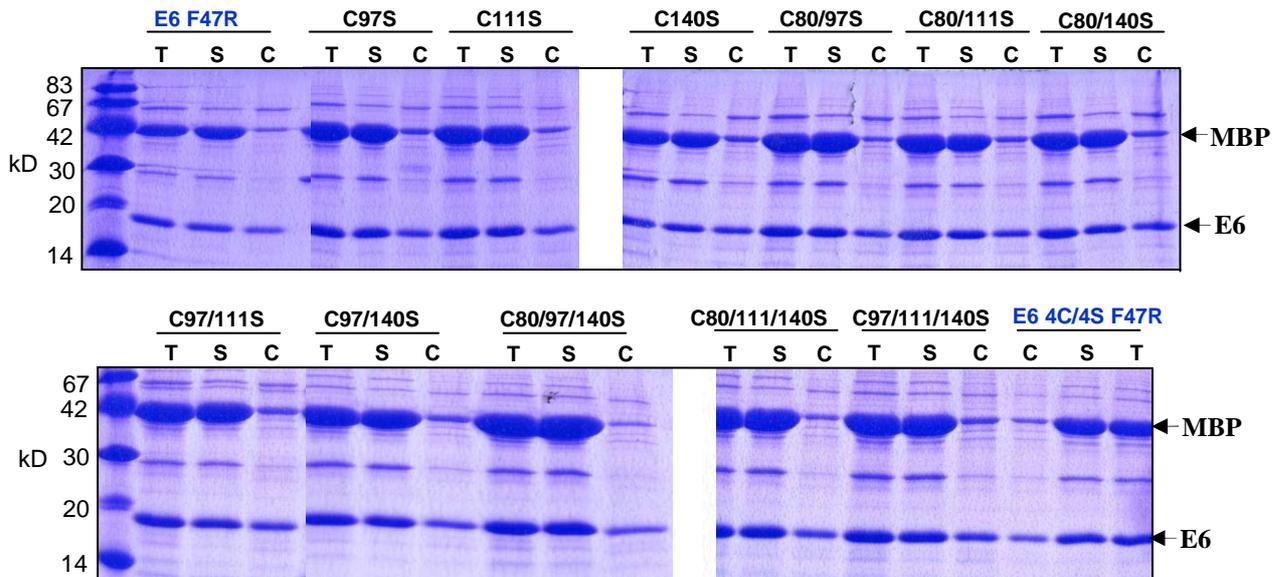


Figure 27. Tests de solubilité de différents mutants incluant un nombre minimum de cystéines mutées comparés avec la construction de E6 4C/4S: Les échantillons sont obtenus après une mini-purification en batch de MBP-E6. Révélation sur un gel SDS-PAGE 15%. L'expression de MBP-E6 (20 ml) a été faite dans les bactéries BL21, après induction, les cultures sont incubées à 20 °C pendant 13 heures. Les constructions sont mini-purifiées dans un petit volume d'une résine amylose. La fusion MBP-E6 ainsi éluée, est soumise à un clivage par la protéase TEV. Après la mise à niveau de la concentration de chaque échantillon, les produits de digestion de la TEV sont remis à une température de 8 °C pendant 15h. Ensuite les échantillons sont centrifugés pendant 30 min (13000 rpm) pour séparer les parties agrégées et solubles. Les fractions (totale (T), surnageant (S) et culot (C)) sont déposées et migrées sur le gel. La fraction (T) représente le produit de digestion qui contient en plus de la MBP, la fraction totale de E6 (agrégée et soluble). (S) le surnageant contenant la partie soluble de E6. (C) le culot composé essentiellement des agrégats de E6. La comparaison entre les fractions solubles de E6 de différentes constructions, indique que la mutation des quatre cystéines (4C/4S) procure l'effet le plus important en terme de solubilité de E6.

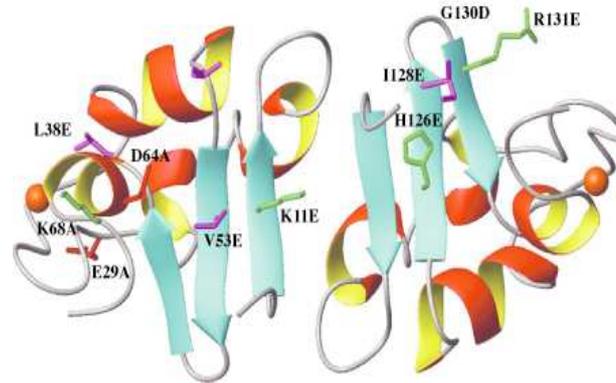
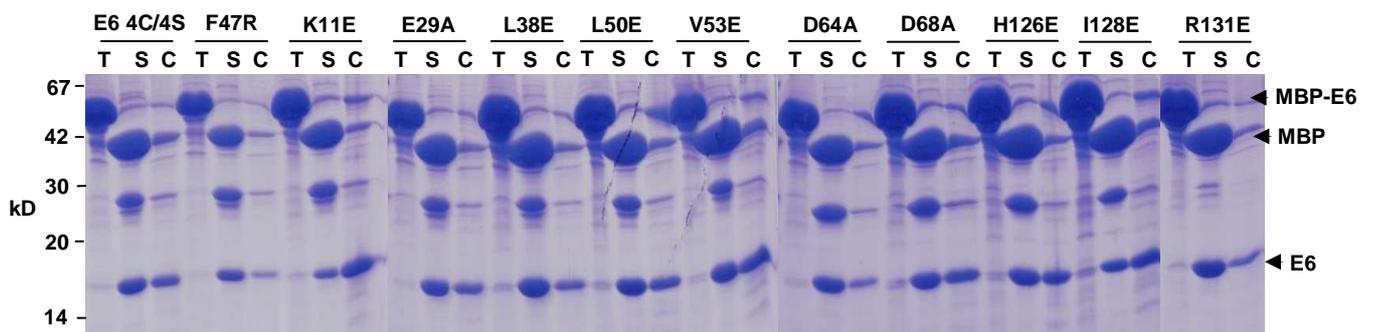
A**B**

Figure 28. Les résidus potentiellement impliqués dans d'interaction entre E6 et le motif LXXLL: (A) les résidus identifiés par titration RMN-HSQC sont représentés sur le modèle de E6. **(B)** tests de solubilité des différents mutants issus d'une série de mutagenèse dirigée sur ces résidus. Les mini-purifications et procédures utilisés dans les tests sont déjà décrits (figure 26). Les mutations ponctuelles sont introduites dans le contexte du mutant E6 4C/4S connu pour sa solubilité. Les deux premiers échantillons (E64C/4S et E6 4C/4S F47R) sont utilisés comme témoins, pour comparer les solubilités des mutants.

Les positions des résidus cystéines mutées sont les suivantes: Cys 80, Cys 97, Cys 111 et Cys 140. La méthode de crible que nous avons utilisé est un test de solubilité sur des mini-purifications. Le même test décrit dans le paragraphe précédent, montre que les différents mutants quelque soit le nombre ou les positions de cystéines mutées apparaissent moins soluble que E6 4C/4S F47R. A la lumière de ce résultat nous considérons qu'il est nécessaire de garder les quatre mutations de cystéines (Figure: 27).

Interaction de HPV16 E6 avec l'ubiquitine ligase E6-AP

Des mesures RMN-HSQC de titration du mutant E6 4C/4S avec un peptide (ESSENTLQELLGEER) provenant de E6-AP (qui contient le motif consensus LXXLL) montrent que certains résidus voient leur environnement chimique changer en présence du peptide. Sur la base des variations des fréquences (^1H , ^{15}N) observées, une série de résidus de E6 possiblement impliqués dans l'interaction est identifiée (Figure: 28 – A). Nous avons ensuite procédé par mutagenèse dirigée au changement ponctuel de chacun de ces résidus. Dans premier temps une vérification de la solubilité des mutants était primordiale avant d'entamer une étude fonctionnelle. Les tests de solubilité sur des protéines E6 mini-purifiées (E6 4C/4S K11E, E6 4C/4S E29A, E6 4C/4S L38E, E6 4C/4S L50E, E6 4C/4S V53E, E6 4C/4S D64A, E6 4C/4S D68A, E6 4C/4S H128E et E6 R131E) révèlent que les différents mutants gardent plus ou moins la solubilité du mutant de départ E6 4C/4S (Figure: 28 – B).

HPV16 E6, liaison E6-AP et dégradation de p53

À l'issue de ces études structurales, nous avons élargi notre étude afin d'étudier les particularités fonctionnelles des différentes constructions que nous avons conçues dans le contexte de E6 complète. Ce volet fonctionnel a été réalisé en collaboration avec une doctorante de notre équipe, Charlotte Boulade-Ladame. Le but était de tester la capacité de l'ensemble des mutants à interagir avec le motif LXXLL, avec E6AP entière ainsi que l'implication de chaque mutant dans le phénomène de dégradation de p53. Nous avons utilisé les deux types de mutants:

- ceux résultant de la substitution des résidus de la région responsable de la dimérisation du domaine N-terminal.
- l'autre série de mutants visant la région « cible » d'interaction potentielle avec les motifs LXXLL précédemment identifiée par titration RMN-HSQC.

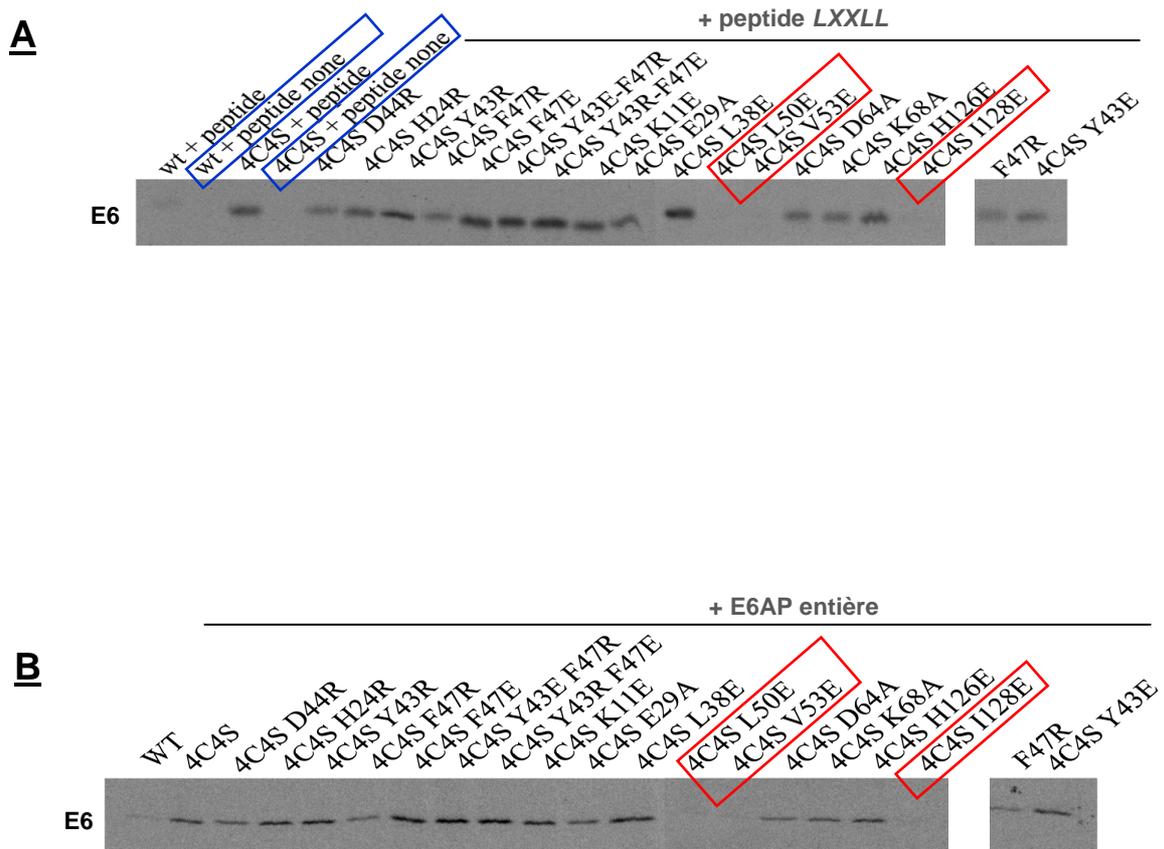


Figure 29. Etude fonctionnelle de deux types de mutants de E6: GST-pull down pour l'interaction *in vitro* entre E6 et E6AP. Les mutants E64C/4S L50, V53 et I128 (encadrées rouges) ne lient pas le peptide qui contient le motif LXXLL (**B**) ni la protéine E6AP entière (**B**).

Pour les premiers mutants, nous montrons par des expériences GST pull-down, que certaines mutations dirigés sur F47, sur F47 et Y43 (F47E, F47R, F47R-Y43E et F47E-Y43R), n'altèrent pas l'affinité et l'interaction *in vitro* entre E6 et son ligand E6AP ; bien que ces mutations abrogent la dégradation de p53 *in vitro* et *in vivo* (Figure: 29, 30). L'interaction E6/E6AP a eu lieu de la même façon si on utilise le GST-peptide LXXLL ou bien la GST-E6AP (Figure: 29).

Pour la deuxième série de mutants, les mutations L50E, V53E et I128E empêchent l'interaction E6/E6AP, et ne dégradent pas p53 *in vitro*, mais curieusement n'affectent pas la capacité de E6 à dégrader p53 *in vivo*. Le mutant E6 4C/4S K11E a un comportement similaire aux mutants de F47: il ne dégrade pas p53 *in vitro* ni *in vivo*, tout en restant capable de lier E6AP (Figure: 29, 30).

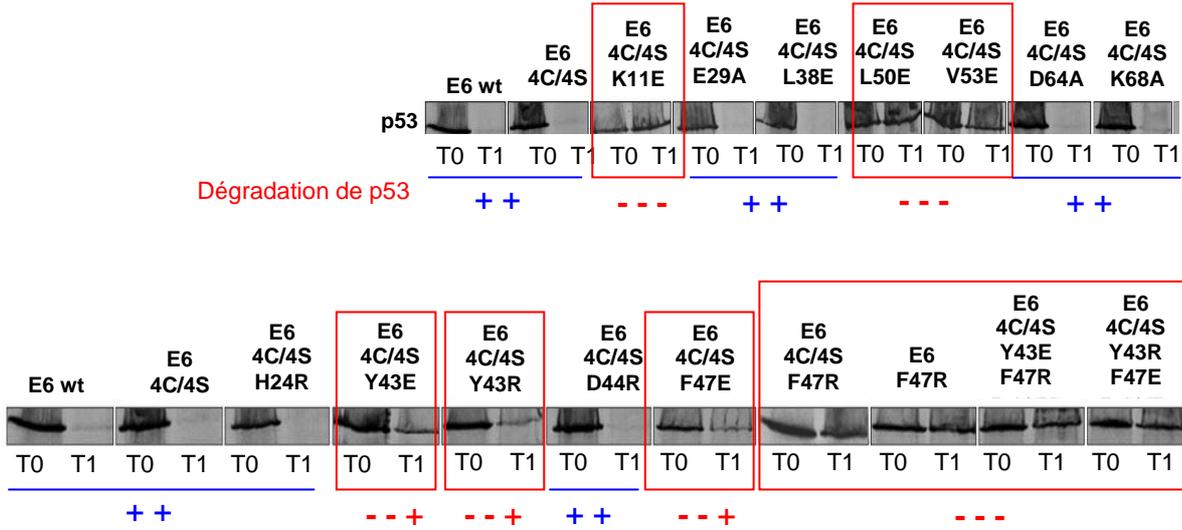
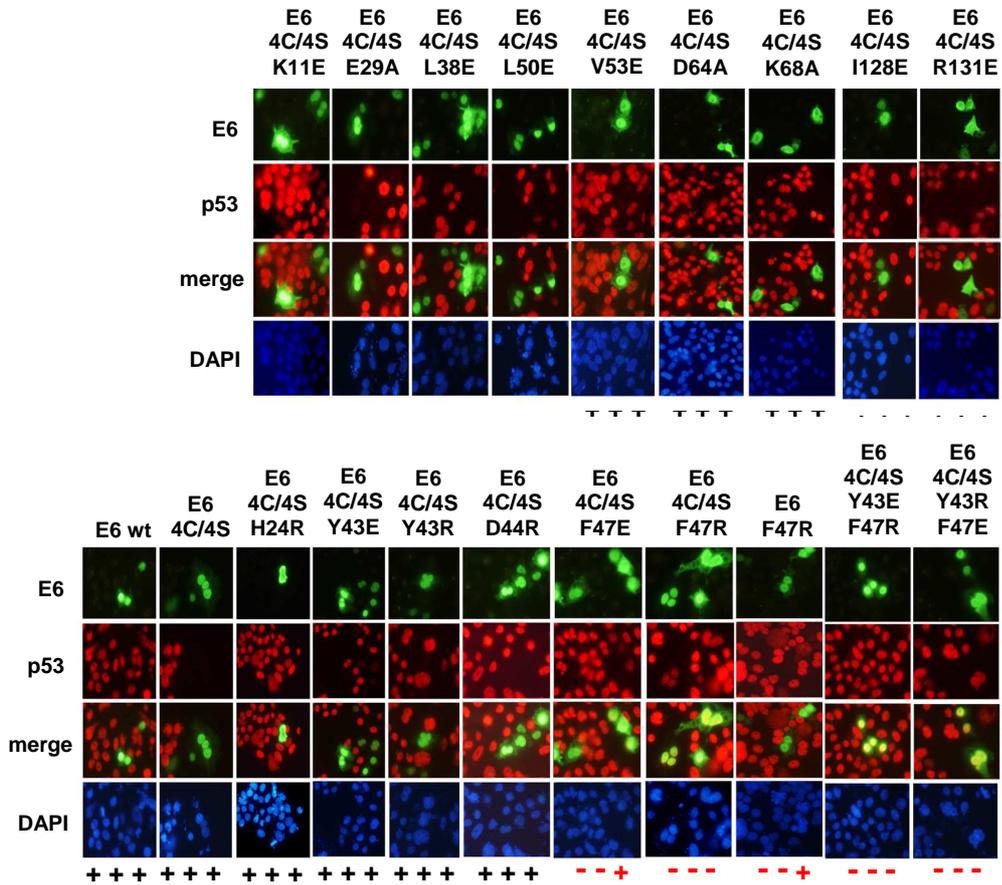
A**B**

Figure 30. Eude fonctionnelle de deux types de mutants de E6: (A) tests *in vitro* de la dégradation de p53 par la protéine E6 dans les lysats de réticulocytes de lapin. Les mutants encadrés en rouge affectent la dégradation de p53. (B) dégradation de p53 *in vivo* par la protéine E6 dans les cellules HPV négative (C33A). (+) dégradation de p53, (-) absence de dégradation de p53.

Deuxième partie : Caractérisation biophysique de la protéine E6 de BPV

Manuscrit 1 :

Biophysical characterisation of Bovine Papillomavirus E6 oncoprotein and its interaction with a LXXLL motif derived from paxillin.

Dans le présent article, sont présentés nos résultats sur la protéine E6 issue du papillomavirus bovin BPV1. Cette protéine contient deux domaines à zinc et possède des activités oncogéniques à l'instar des protéines E6 de HPV à haut risques. Pour explorer des conditions permettant l'étude structurale de cette protéine, nous nous sommes inspirés de stratégies déjà utilisées pour HPV E6 et nous avons également développé de nouvelles approches. Différentes formes (sauvage, mutée ou fusionnée à un peptide ligand) de BE6 ont été clonées en fusion avec l'extrémité N-terminale de la protéine MBP ou la GST. Depuis l'expression jusqu'à la dernière étape de purification, des précautions particulières se sont également avérées nécessaires pour remédier aux différents problèmes de solubilité et d'agrégation.

Tout au début, nous avons constaté que la protéine MBP-BE6 (forme sauvage) formait majoritairement des agrégats solubles qui précipitaient lors de l'étape d'ultracentrifugation de longue durée que nous appliquons systématiquement aux protéines de fusion purifiées pour l'étude structurale. Cependant, nous parvenions à isoler une fraction minoritaire de MBP-E6 monomérique. Toutefois, la récupération de protéine BE6 à l'état monomérique après le clivage par la protéase TEV, suivi d'un passage sur colonne de chromatographie d'exclusion de taille, est conditionnée par la concentration de l'échantillon. En effet, la protéine BE6 est soluble à des basses concentrations (5 – 10 μ M), mais une fois que cet intervalle soit dépassé, l'agrégation de la protéine est immédiate, empêchant l'obtention d'échantillons aptes à l'étude structurale.

Les deux domaines de BE6 ont été également étudiés. La frontière entre les deux domaines a été définie sur la base d'alignements de séquence avec les protéines HPV E6. Le domaine N-terminal BE6-N était très soluble contrairement au domaine BE6-C. Ceci peut être dû au grand nombre de cystéines présentes dans ce domaine, qui pourraient s'oxyder durant la purification. Alternativement, on peut imaginer que le repliement du domaine BE6C sous forme isolée est insuffisamment stable, et ne se stabilise qu'en présence du domaine BE6N. Nous avons obtenu des données structurales sur la partie BE6-N que nous produisons en échantillons suffisamment concentrés pour des mesures de RMN. Le spectre HSQC du domaine BE6-N (300

μM) démontre une bonne qualité du repliement, alors que seul, le domaine BE6-C riche en résidus cystéines non conservées ne supporte pas les concentrations nécessaires pour les mesures RMN. Nous avons effectué plusieurs mutations visant les cystéines non conservées du domaine BE6-C, mais ces mutations n'ont pas aidé à améliorer la qualité des échantillons contrairement à ce que nous avons auparavant obtenu avec le HPV E6-C.

En quête de nouvelles stratégies alternatives, des expériences de phage display effectuées par nos collaborateurs américains dans ce projet (équipe de Scott Vande Pol) ont montré que la protéine BE6 lie préférentiellement une protéine cellulaire appelée la paxilline. La liaison se fait via une séquence de motif consensus LXXLL se trouvant dans séquence de la paxilline et d'autres protéines cellulaires ciblées par E6. Nous avons par la suite montré que l'affinité de l'interaction BE6-LXXLL est relativement forte pour une interaction protéine-peptide ($K_d = 0,1\mu\text{M}$). Sur cette base, nous avons, en concertation avec nos collaborateurs, cherché à mettre à profit cette interaction pour la production et la purification de la protéine.

Nous avons effectué différentes tentatives de co-expression et/ou co-purification des deux partenaires. Tout d'abord la protéine BE6 sauvage, produite sans fusion, a été co-exprimée avec son ligand peptidique minimal issu de la paxilline ("peptide paxilline") fusionné au C-terminus de la GST (Gluthatione-S-Transférase). La co-existence des deux vecteurs dans une même cellule bactérienne augmente le niveau d'expression de BE6. Nous avons ensuite cherché à co-purifier le complexe GST-peptide / BE6. Cette stratégie a fonctionné mais les rendements sont restés insatisfaisants pour la production en vue d'études structurales. Une approche voisine a été tentée, consistant à produire la fusion GST-peptide et la protéine MBP-E6 dans des bactéries séparées, et de mélanger les deux expressions dès le début de la purification, lors de l'étape de sonication. Nous avons effectivement réussi à co-purifier des complexes GST-peptide / MBP-E6, mais en quantité insuffisante, probablement parce que les protéines MBP-BE6 sont majoritairement produites sous formes d'agrégats solubles inactifs pour la liaison au peptide.

Nous nous sommes alors opté pour une autre approche consistant à produire une chimère covalente où le peptide paxilline est fusionné à la protéine BE6 par l'intermédiaire d'un "linker" de trois résidus, le tout étant produit sous forme de protéine de fusion à la MBP. Cette approche a amené à une amélioration impressionnante sur le plan qualitatif et quantitatif. Les agrégats solubles, majoritaires pour la forme MBP-E6, sont pratiquement inexistantes pour la forme MBP-peptide-BE6. De plus, la protéine chimérique peptide-BE6, une fois séparée de MBP par protéolyse spécifique, peut être isolée sous la forme d'un monomère hautement soluble. Nous en avons obtenu des échantillons marqués ^{15}N suffisamment concentrés pour les mesures

RMN. Néanmoins, à haute concentration (600 μ M), les spectres HSQC ont montré des phénomènes d'élargissement de raies qui pourraient être imputables à un début d'agrégation, et des spectres effectués à des concentrations plus basses révèlent des phénomènes d'échange qui pourraient être induits par la fusion covalente de la protéine avec son peptide ligand.

Enfin, l'analyse comparée des deux approches (co-purification et fusion covalente) nous a permis d'imaginer une troisième voie cumulant les avantages de l'une et l'autre. Dans cette nouvelle approche, la protéine BE6 est fusionnée avec le peptide LXXLL lui-même placé au C-terminus de la protéine de fusion (MBP ou GST). La séquence sensible à la protéase TEV, qui se trouvait entre la MBP (ou GST) et le peptide dans la construction précédente est désormais placée dans la région séparant le peptide et BE6. L'expression et le repliement restent ainsi optimaux lors de la production bactérienne grâce à la présence stabilisatrice du peptide. Lors de la purification sur une colonne d'affinité BE6 est séparée du peptide qui reste cependant lié de façon non covalente à la protéine MBP-peptide (ou GST-peptide selon la nature de la protéine de fusion). Lors de cette étape cruciale, nous ajoutons un excès de la forme synthétique du peptide LXXLL, qui vient s'échanger avec le peptide de fusion tout en préservant la stabilité et la solubilité de la protéine BE6. Nulle perte n'est remarquée après une longue étape d'ultracentrifugation; la solubilité et la concentrabilité du complexe non covalent purifié restent identiques sinon meilleurs que ce qui était observé pour le complexe covalent. A l'issue de cette étude, nous disposons désormais d'un protocole et de données biophysiques préliminaires permettant d'aborder l'étude structurale de la protéine BE6 complexée avec son principal ligand. De plus les approches innovantes que nous avons utilisées pourraient être adaptables pour étudier la biochimie et la biophysique d'autres protéines difficiles.

Biophysical characterisation of Bovine Papillomavirus E6 oncoprotein and its interaction with a LXXLL motif derived from paxillin.

Abdellahi Ould M'hamed Ould Sidi¹, Nicole Brimer², Katia Zanier¹, Brian Vande Pol², Yves Nominé¹, Christophe Romier³, Scott Vande Pol² and Gilles Travé¹

¹UMR7100, Ecole Supérieure de Biotechnologie de Strasbourg, Boulevard Sébastien Brant, BP 10413, 67412 Illkirch, France

²Department of Pathology, University of Virginia. P.O. Box 800904, Charlottesville, VA 22908, USA.

³Institut de Génétique et Biologie Moléculaire et Cellulaire, 1 Rue Laurent Fries, BP 163, Illkirch CEDEX, France.

Running title: Biophysical Characterisation of BPV E6 protein

SUMMARY

E6 is a small oncoprotein involved in tumorigenesis induced by papillomaviruses (PVs). E6 binds and often targets to proteasomal degradation numerous target cell proteins acting in key cellular pathways such as apoptosis, cell cycle control, transcriptional control and cellular adhesion. E6 often recognises its targets by binding to short motifs presenting the consensus LXXLL. E6 proteins from Human Papillomaviruses (HPVs) have long resisted structural analysis. Here, we explored the suitability of the E6 oncoprotein of Bovine Papillomavirus (BE6) and its isolated N-terminal (BE6N) and C-terminal (BE6C) zinc-binding domains for biophysical and structural analysis. Whereas the BE6N domain was highly soluble, the BE6C domain as well as full-length BE6 protein, were prone to aggregation. Phage display experiments and BIAcore analysis demonstrated that full-length BE6 bound to the LXXLL motif of paxillin and homologous peptides with high affinity, suggesting that the motif might stabilize the protein. Indeed, the presence of the LXXLL motif, either fused to the N-terminus of BE6 or non-covalently bound to it, permitted the obtention of highly soluble constructs of BE6 suitable for NMR analysis. Samples of BE6N, BE6 non-covalently bound to LXXLL motif, and BE6 covalently fused to the LXXLL motif, were (15N, 13C)-labelled and multidimensional NMR spectra were recorded. These data open the way to the first novel structural and functional analysis studies of full-length E6 oncoprotein, in complex with its preferential target motif.

INTRODUCTION

Papillomaviruses (PV) are small DNA viruses that infect epithelial tissues within mammals and birds. Almost 200 PV types have been identified (de Villier *et al.*, 2004). PVs display distinct tropisms for different body sites (skin, mouth and genitalia) as well as distinct levels of pathogenicity. Infections by “high-risk” PVs can progress into cancer. In human, cervical cancers are caused by infection of high-risk genital Human Papillomaviruses (HPVs) (zur Hausen, 1999). HPV 16 is the most common high-risk genital type, followed by HPV 18, 45, 31, 33, 52 and 58 (Bosch, 2003). The high-risk cutaneous HPV types 5 and 8 have been instead associated to the occurrence of squamous cutaneous carcinoma (SCC) linked to the genetic disease *epidermodysplasia verruciformis* (EV) (Pfister, 2003). Among papillomaviruses infecting other animals, the high-risk bovine papillomavirus (BPV) has also been extensively studied because of its strong transforming properties.

PV tumorigenesis is mainly supported by the activities of two oncoproteins, E6 and E7. E6 proteins from high-risk genital HPVs form a trimeric complex with the tumor suppressor p53 and the ubiquitin ligase E6-associated protein (E6AP) (Scheffner *et al.*, 1993), which leads to ubiquitin-mediated degradation of p53. High-risk genital HPV E6 also target many other cellular proteins to degradation (reviewed in (Chakrabarti & Krishana, 2003), including the MAGI family of PDZ-containing proteins (Glaunsinger *et al.*, 2000; Thomas *et al.*, 2002) which recognise E6 *via* their short motif at the extreme C-terminus of E6 () multiple PDZ domains. In contrast, little is known about the targets of the High-risk cutaneous HPV E6 proteins, which might carry out their functions *via* alternative mechanisms. For instance, E6 from HPV 5 and 8 do not target p53 to degradation, but they promote the degradation of the Bak pro-apoptotic protein, which is normally induced after UV irradiation of the skin we.

HPV16 E6 binds several of its targets proteins, including E6AP (Huibregtse *et al.*, 1993), E6BP (Chen *et al.*, 1998), IRF-3 (Ronco *et al.*, 1998), and possibly p300/CBP (Patel *et al.*, 1999; Zimmermann *et al.*, 1999) *via* “charged leucine” motifs, also called “LXXLL” motifs. These motifs contain the consensus LxxLsh sequence, where L indicates conserved leucine residues, 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 (Chen *et al.*, 1998; Vande Pol *et al.*, 1998; Elston *et al.*, 1998). The bovine papillomavirus E6 (BPV E6) also interacts with the LXXLL motif of E6AP, as well as LXXLL motifs found on the focal adhesion protein paxillin and possibly the AP1 adapter complex (Das *et al.*, 2000; Tong *et al.*, 1997, Vande Pol *et al.*, 1998). This interaction is required for cellular transformation by BPV E6 (Bohl *et al.*, 2000; Vande Pol *et al.*, 1998). Finally, recent data also suggest

that the low risk genital HPV-11 also interacts with the LXXLL motif of E6AP (Brimer *et al.*, 2007) suggesting that binding to LXXLL motives may be a common feature of most, if not all, E6 proteins.

Most mammalian PV E6 proteins are cysteine-rich small proteins of about 150 amino acids, which consist of two zinc-binding domains, E6-N and E6-C (Lipari *et al.*, 2001) Nominé *et al.*, 2003). We recently solved the a solution structure of the a mutant of HPV16 E6-C domain with four cystein residues changed to serines to limitate oxidative aggregation, and proposed a model for the structure of the full-length protein (Nominé *et al.*, 2006). Yet the high-resolution tridimensional structure of the entire E6 protein has not been solved yet, due to difficulties in its recombinant production. Attempts to produce the native full-length protein in *E. coli* have mainly focussed on E6 from HPV 16 (E6₁₆) (Daniels *et al.*, 1997; Lipari *et al.*, 2001; Nominé *et al.*, 2001; Degenkolbe *et al.*, 2003). Recently, we also explored the behaviour of several E6 proteins of other HPV types, including major high-risk genital HPV 18, 33, 45, 52, 58, low-risk genital HPV 11 and high-risk cutaneous HPV 5 (Zanier *et al.*, 2007). None of the HPV E6 proteins investigated raised samples amenable to structural analysis.

In the present work, we explored the suitability of the BPV E6 protein for biophysical and structural analysis, as an alternative to E6 proteins from Human papillomaviruses. On the one hand, we show by phage-display and BIAcore that the BPV E6 protein binds LXXLL motifs with high affinity and selectivity. On the other hand, we have obtained stable (¹⁵N, ¹³C)-labelled samples of both an isolated BPV-E6N domain and full-length construct of BPV-E6. Bearing a LXXLL motif from paxillin fused at its N-terminus, and the first 2D- and 3D-NMR spectra have been recorded. Remarkably, the presence of the fused LXXLL motif, covalently or non-covalently bound to the protein, was required for the obtention of a stable construct of full-length BPV E6 suitable for NMR analysis. Moreover, we have also produced double-labelled samples of a construct corresponding to the N-terminal zinc-binding domain of BPV E6. This domain also gave raise to NMR spectra of excellent quality. These data open the way to the first structural analysis of full-length E6 oncoprotein, in complex with its preferential target motif within cellular proteins.

MATERIAL AND METHODS

Expression constructs

The DNAs encoding full-length BPV1 E6 (BE6) (137 residues), and its isolated domains BE6-N and BE6-C (66 and 71 residues respectively) were amplified by PCR and cloned into the pETM41 vector (generous gift from Gunter Stier, EMBL-Heidelberg) allowing production of the protein of interest fused to the C-terminus of 6His-tagged maltose binding protein (MBP) *via* a linker sensitive to TEV protease. These three plasmids allow the production of 6His-MBP-TevSite-BE6N, 6His-MBP-TevSite-BE6C, and 6His-MBP-TevSite-BE6 constructs, respectively. Another construct consisting in the 12 amino acids E6-binding LXXLL motif of paxillin fused to the N-terminus of full-length BE6 *via* a short linker of 3 glycine residues was also cloned into the pETM41 vector, allowing the expression of a 6His-MBP-TevSite-LXXLL-BE6 construct. Finally, a last construct was designed, in which we fused directly the LXXLL motif to the C-terminus of 6his-MBP without a TEV-sensitive linker, and introduced the TEV sensitive linker at the junction between the LXXLL motif and the full-length BE6 sequence. This plasmid allowed the expression of a 6His-MBP-LXXLL-TevSite-BE6 construct.

The full-length sequence of BE6 was also cloned into the expression vector pACYC-11b carrying chloramphenicol resistance and a p15A origin of replication which permits co-expression with a second protein expressed in a classical PET vector (Romier *et al.*, 2006). Co-expression of protein complexes in prokaryotic and eukaryotic hosts: experimental procedures, database tracking and case studies. Acta Cryst. D62, 1232-1242).

DNA oligomers coding for the negative control peptide (GSNSGNGNS), the paxillin-derived peptide (DDL DALLADKE) and the E6AP peptide (ESSELT LQELLGEER) were cloned into the NcoI/KpnI sites of the pETM-30 expression vector containing a N-terminal His6-GST tag and a TEV protease cleavage site. pETM-vectors (Nominé *et al.*, 2003) were kindly provided by G. Stier, EMBL, Heidelberg. All clones were verified by DNA sequencing.

Expression and purification procedures

The vectors were transformed into either BL21(DE3) and BL21(DE3) pLysS *E. coli* competent cells according to standard procedures. For a preculture of 10 ml LB medium containing kanamycin (15 µg/mL), kanamycin and chloramphenicol (34 µg/mL) for the strains carrying a chloramphenicol-resistant plasmid (pLysS). Expression cultures of 500 ml were grown at 37 °C until an OD₆₀₀ of 0.6 was reached, induced with a final concentration of 0.5 mM IPTG and further grown overnight at 15 °C in either LB medium or labelled M9 medium cultures. Expression cultures were harvested by centrifugation at 4000 rpm and 4 °C for 20 min.

For purification, the pellet was resuspended in 50 ml buffer containing 2.5 µg/ml DNase I and 2.5 µg/ml

RNase I. Anti-protease cocktail (EDTA-free; Roche Applied Science) was added at the concentration recommended by manufacturer. Initial buffer was buffer A (Tris 50 mM pH 6.8, NaCl 400 mM, and DTT 2 mM). All purifications buffers were extensively degassed then bubbled with Argon. Cells were broken by sonication on ice and then centrifuged at 38,000 rpm for 1h at 4 °C. The supernatant was filtered (Millipore 0.22 µm) and loaded on the affinity amylose resin column (New England Biolabs) equilibrated with the same buffer. The column was washed with three volumes of buffer A. The protein fusion was then eluted from resin by using 2 volume of buffer A containing 20 mM maltose. The eluate was ultracentrifuged overnight at 160,000g and 4 °C in a SW41 rotor (Beckman). Subsequently, the MBP moiety was removed by adding 50 µl of TEV protease and performing an over night digestion at 8 °C. The digested product was concentrated to 2 ml using Vivaspin 15 filter devices with a MWCO of 5 kDa (Vivascience) and loaded on a HiLoad 16/60 Superdex 75 pg ($3 \times 10^3 \times 7 \times 10^4$ Da) gel filtration column (Amersham Biosciences) previously equilibrated with 240 ml buffer of (Sodium Phosphate 20 mM pH 6.8, NaCl 200 mM, and DTT 2 mM). Finally, the fractions corresponding to the recombinant protein were collected, and the concentration was estimated by absorption spectroscopy at 280 nm.

Synthetic peptides

The exact sequence of the synthetic peptides used for the co-expression and co-purification is: EFTMDDL DALLADKE

Phage display experiments.

We performed phage display selection experiments with an M13 phage display library displaying a random 12 residue peptide fused to the gene III coat protein (Smith, 1991). BPV E6 protein as well as a panel of other E6 proteins from HPV 1, 5, 11, 16 and rabbit CRPV-1 were expressed in mammalian cells by vaccinia virus expression as a fusion to either GST or a chitin binding domain (CBD) followed by a TEV protease cleavage site and finally E6. The CBD-TEV-E6 and GST-E6 fusion proteins were expressed in CV1 cells, purified on chitin beads, and 1 g was then incubated with 10^{11} pfu of the phage library (2×10^9 complexity). Unbound phage were removed by extensive washing, bound phage released by TEV protease cleavage, followed by amplification in *E. coli*. For each E6 protein, CBD fusion selection was followed by GST fusion selection. After 5 rounds of selection and amplification of bound phage, a final round was performed using an input of 3×10^{10} input phage selected and amplified. The amounts bound and released phages were titered at each amplification cycle.

BIAcore experiments.

Surface Plasmon Resonance (SPR) experiments were performed on a BIAcore 2000 machine and by using an assay based on antibody-mediated capture of recombinant peptides fused to the C-terminus of GST, as described in our previous studies (Zanier *et al.*, 2005). The recombinant GST-fused peptides were LXXL peptides derived from paxillin or E6AP, as well as a control irrelevant peptide (GST-none). We used as an analyte monomeric MBP-E6 fusions isolated by gel-filtration. We previously demonstrated (Zanier *et al.*, 2005) 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 Surface Plasmon Resonance (SPR) specific signal. The HBS buffer usually used for standard BIAcore experiments contains the zinc chelator EDTA, which is incompatible with zinc-binding proteins such as E6, and it does not contain reducing agents, which are particularly required for stability of E6 proteins. Therefore, the HBS buffer was replaced by a running buffer without EDTA, including supplementary reducing agent dithiothreitol (DTT, 2 mM), and the concentration of salt was adjusted to 200 mM (50 mM Tris (pH 6.8), 200mM NaCl, 2 mM DTT, 0.005% P20). The recombinant gel filtrated MBP-BE6 was diluted in this buffer just before the start of the measurements to avoid long-time storage at this relatively low salt concentration. The use of this optimised running buffer considerably improved data quality. Injecting the MBP-BE6 analyte within one to five hours after dilution in running buffer yielded reproducible results, verifying that the solubility and monodispersity of analyte did not vary during the time required to perform a set of measurements.

RESULTS

Sequence alignments suggest that BE6 possesses two zinc-binding domains comparable to those of HPV E6.

A sequence alignment of BPV E6 versus E6 proteins from a representative set of HPV species (Figure 1) reveals that BPV E6 probably contains two zinc-binding domains adopting a similar fold. Zinc-binding cysteine residues as well as most of key buried hydrophobic positions appear to be conserved in the Bovine E6 protein. The borders of E6N and E6C structural domains seem also conserved across human and bovine papillomavirus species.

Recombinant BE6N domain is highly soluble and amenable to structural analysis in contrast to full-length BE6 and the BE6C, which are prone to aggregation.

Preliminary experiments using partially purified BPV E6 (Vande Pol, unpublished data) suggested that this protein might be more soluble than HPV E6. The wild-type sequence of BPV1a E6 was fused to the C-terminus of Maltose Binding Protein (MBP) via a linker sensitive to TEV protease. The resulting MBP-fusion (called MBP-tevs-BE6) was precipitated during the steps of purification.

According to our results from previous work on HPV-16 E6, the isolation of two production of isolated zinc-binding domains of the BPV E6 might be help to better understanding may be a useful approach for the biophysical characterisations of the E6 proteins. Therefore, residues 1- 66 (M₁DLK---RRLW₆₆) were fused to the C-terminus of MBP via a linker sensitive to TEV protease (Figure 2 A). The purified BE6-N sample (called MBP-tevs-BE6) was completely soluble and could be concentrated up to 600 μ M without undergoing any detectable aggregation.

By contrast, when the C-terminal domain called BE6-C (residues 67 – 137) (Q₁GVP---SKYP₁₃₇) was cloned and expressed according the same conditions used in a similar manner as for BE6-N, it cannot raise samples amenable be obtained with a concentration enough for to biophysical characterisation. Once separated from the MBP tag, the domain aggregates during the purifications steps and finally the separated sample form diluted amounts of monomeric domain (5 - 10 μ M) could be isolated by gel-filtration column (5 - 10 μ M) is no able to concentrate (Figure 2 B) but they could not be concentrated further, due to precipitation and retention on the concentration membrane.

To try to circumvent this problem, we have produced different mutants of the domain BE6-C. Sequence explorations of the BE6 showed that the sequence contains a several non-conserved cisterns residues at the moiety C-terminus of the protein. Firstly, we have generated a longer construct that include and three additional residues of BE6-Nat the N-terminus (63 – 137) of BE6-C, therefore spanning residues 63 – 137, but that not give a better results for the domain. Also we have also mutated two non-conserved cystein residues at the positions C91S and C128S ("2C/2S" mutants) in the context of both the BE6-C long (63 – 137) and the BE6-C (66 – 137) short and longer version of BE6C. The three resulting constructs (BE6-C (63–137), BE6-C 2C/2S (63–137), and BE6-C 2C/2S (66–137)) the same phenomena of aggregation have observed and were prone to aggregation and precipitated upon concentration in a manner comparable to the BE6-C (66–137) construct initially examined.

Initial NMR analysis of (¹⁵N)-labeled samples of the N-terminal zinc-binding domain of BPV E6.

The purified samples of ¹⁵N-labelled BE6-N behaved as a monomer could be readily concentrated up to 500 μ M to the requirements of allowing NMR analysis. Preliminary ¹H,¹⁵N-HSQC spectra indicate a folded domain in which all amides peaks can be counted (Figure 3). The high quality of these initial NMR spectra together with the small size of the BE6-N domain make it a very promising target for fast

solution structure determination using standard methods. Obviously, we could not obtain spectra for the samples of BE6 and the domain BE6-C, because the concentrations required for NMR are not reached. Analytical ultracentrifugation measurements performed on a fraction of the NMR sample indicated that BE6N was monomeric in these conditions (data not shown).

Recombinant full-length BE6 displays high affinity for the LXXLL motif of paxillin and homologous peptides.

□ *Phage display experiments*

In an alternative search, the Charlottesville team investigated the peptide-binding properties of a panel of recombinant E6 proteins from various papillomaviruses. We hypothesized that the E6 protein type that upon overexpression best bound its optimal peptide sequence *in vitro* would represent the type that best folded *in vitro*. Since we do not know the optimal LXXLL peptide binding sequence for E6 from most papillomavirus types. To investigate peptidic binders of BPV E6, we performed a phage display selection experiment with an M13 phage display library displaying a random 12 residue peptide fused to the gene III coat protein (Smith, 1991). Each BPV E6 protein as well as a control panel of E6 proteins from other papillomaviruses was expressed and produced in mammalian CV1 cells by vaccinia virus expression as a fusion to either GST or a chitin binding domain (CBD) followed by a linker sensitive to TEV protease cleavage site and finally E6. The Purified CBD-TEVtevs-E6 or corresponding GST fusion protein-tevs-E6 fusion proteins are, still attached to their cognate affinity resins, expressed in CV1 cells, purified on chitin beads, and 1 g was then incubated with 10^{11} pfu of the phage library (2×10^9 complexity). Unbound phage was removed by extensive washing, bound phage released by TEV protease cleavage, followed by amplification in *E. coli*. For each E6 protein, CBD fusion selection was followed by GST fusion selection and were alternated to avoid the amplification of CBD-binding or GST-binding sequences. After 5-6 rounds of selection and amplification of bound phage, a final round was performed using an input of 3×10^{10} input phage selected and amplified by each E6 type. Note that each E6 type was used to select its own optimal peptide sequence. The bound and released phage were titered with the results shown in (Table 1). Of all the E6 types (HPV types 1, 5, 11, 16, BE6 and cotton-tailed rabbit papillomavirus (CRPV E6)) only 16E6, 11E6 and BE6 selected a significant amount of phage-peptides that bound greater than 10 fold more than beads alone without E6. Since each E6 type had the opportunity to select its optimal peptide sequence, it seems likely that the control E6 proteins when overexpressed as fusions and purified in this fashion are were not folded in a conformation that can bind to a phage-displayed peptide. In striking contrast, BE6 was able to bind almost the entire input of the fifth round selected phage, and bound almost 1000 times more of its selected input phage than any other E6 protein (Table 1). We picked 6 individual plaques from the BE6 bound phage and sequenced

them to determine if a consensus sequence was present, with the results shown in (Figure 4). Each plaque contained a unique consensus sequence bearing close resemblance to the LXXLL sequences of E6AP and paxillin. This demonstrates that overexpressed fusions of BE6 fold properly in order to bind *in vitro* peptides similar to peptides found on associated proteins isolated *in vivo*. In contrast, peptide sequences bound by 11E6 and 16E6 *in vitro* from their respective phage display experiments were divergent from the LXXLL sequence of E6AP (data not shown). This indicates that while HPV E6 proteins prepared in this fashion did not select peptides well *in vitro* (confirming the solubility problems of HPV E6 proteins independently observed by the Strasbourg group), BE6 fusions folded appropriately and may be suited for initial structural E6 studies that examine the interaction of E6 with LXXLL.

□ **Preliminary Analysis of the Interactions of BE6 with the paxillin using surface plasmon resonance**

The Surface Plasmon Resonance experiments are performed as described in our previous studies (Zanier *et al.*, 2005). We use an analyte BE6 in fusion with MBP to occur the benefic role of MBP in high- yielded protein solubility in order to prevent the misfoldig of BE6. Especially the HBS buffer usually used for standard BIAcore experiments contains the zinc chelator EDTA and does not contain any reducing agent was replaced by a running buffer without EDTA, including supplementary reducing agent dithiothreitol (DTT, 2 mM), and the concentration of salt was adjusted to 200 mM. The recombinanate gel filtrated MBP-BE6 was diluted in this buffer just before the start of the measurements to avoid long-time storage at this relatively low salt concentration. The use of this optimised running buffer considerably improved data quality. We inject the MBP-BE6 analyte within one to five hours after dilution in running buffer yielded reproducible results, verifying that the solubility and monodispersity of analyte are unvarying during the time required to perform a set of measurements.

The immobilisation of antibody at the chip surface was performed in the same modified buffer, and the initial signal detected is about 1000 resonance units (RU). After immobilisation of the solute corresponding to diluted purified GST-LXXLL (E6-binding LXXLL motif of paxillin fused to the N-terminus of glutathione S-transferase (GST)) or the GST-none (as control) the real response signal reached to 100 RU.

Finally, the injection of a panel of low concentrations of the MBP-BE6 has permitted to collect the corresponding signals according the high or low affinity between the solute (GST- peptide) and the analyte (MBP-BE6). These measurements are reproducibly observed to avoid a possibly artefacts of Biacore instruments and aggregation problem of purified MBP-BE6.

We previously analysed by Surface Plasmon Resonance (BIAcore instrumentation) the kinetic and equilibrium constants of binding of a soluble mutant of HPV16 E6 to a peptide containing the consensus

sequence LXXLL and derived from the E6AP ubiquitin ligase (Zanier *et al.*, 2005). These data were obtained using an experimental setup based on recombinant peptides fused to the C-terminus of the Glutathione-S-transferase (GST). Here we decided to use the same experimental setup to analyse the interaction of BPV E6 with a minimal E6-binding sequence derived from paxilline (REF), which also contains the LXXLL consensus motif. The GST-fused paxillin peptide (the ligand) was immunocaptured via anti-GST antibody on the sensor surface, and put in contact with variable concentrations of purified monomeric MBP-tevs-BE6 fusion (the analyte). A dose-dependent specific response was observed (Figure 5). Although the preliminary sensorgrams shown here do not allow accurate analysis of the kinetics of interaction, the analysis of the dose dependent response indicated an approximate dissociation constant (K_d) of 0.1 M. Noteworthy enough, the dissociation phase was remarkably slow. The half-life of the complexes, estimated from the dissociation profiles, approximated 600 s. In comparison, the half-life of the complex formed between HPV16 6C6S mutant and its cognate E6AP-derived LXXLL peptide was approximately 6 s (Zanier *et al.*, 2005), for a dissociation constant of 2.6 M. Remarkably, SPR analysis of MBP-BE6 analyte versus GST-E6AP peptide as well as MBP-16E6 6C6S analyte versus GST-paxillin peptide showed no dose-dependent signal (data not shown), indicating that the interactions of BE6 and 16E6 with their respective LXXLL peptides are highly selective. We have remarked the high affinity between the MBP-BE6 and the GST-LXXLL (Figure 5), this affinity is predicted by the slowly dissociation time (about 600 minutes) and the estimated dissociation constant ($K_d = 0.1$ M). This constant of dissociation is relatively important if we compare it with the $K_d = 4.22$ M) that we have obtained for the interaction between HPV16 E6 and the GST fused to the LXXLL peptide from E6AP (Zanier *et al.*, 2005).

Co-expression with the GST-fused LXXLL motif of paxillin increases solubility and permits small-scale purification of full-length BE6.

Based on the high affinity and specificity displayed by BE6 towards LXXLL peptides, we explored strategies aimed at stabilizing the BE6 protein by means of its target LXXLL peptide.

First, the paxillin peptide fused to the C-terminus of GST (GST-LXXLL) and the BE6 protein fused to the C-terminus of MBP (MBP-BE6) were independently expressed in separate bacterial cultures. The cell pellets were mixed and sonicated together, and affinity purification of the complexes was attempted on either amylose (MBP-binding) or glutathione (GST-binding) resins. This approach yielded very little amounts of complexes, most probably because MBP-BE6 aggregates pre-formed upon bacterial expression were not proficient for binding to the peptide (data not shown).

A related approach, consisting in adding an excess of synthetic paxillin-derived LXXLL peptide

to the solution used for breaking MBP-E6 expression pellets, was also unsuccessful (data not shown).

Finally, we attempted to co-express, within the same bacterial cells, the genuine untagged BE6 protein and the GST-Paxillin peptide fusion.

We have expressed the BE6 cloned into the co-expression vector of pACYC-11b in aim to exploit the part of the interaction proved above between the peptide contains a motif *LXXLL* from paxillin and the protein by the coexpression experiments. The level of expression of the protein BE6 wild-type as non-tagged protein is increased when we upon coexpress it with the GST-paxillin peptide (Figure 6.A, and 6.B). Also, the Gluthatione affinity co-purification assays demonstrated co-purification of the GST-paxillin peptide with the BE6 protein with GST-paxillin peptide, suggesting that the two co-expressed proteins had formed a stable complex given rise to a purified BE6 and increases its solubility (Figure 6.C and 6.D). However, the levels of soluble BE6 protein purified by this approach remained modest, prompting us to investigate alternative approaches for producing high amounts of pure BE6 required for biophysical and structural studies.

Fusing the LXXLL motif of paxillin to the N-terminus of BE6 prevents its self-association and dramatically increases its solubility.

According our results from the phage display, biacore and the co-purification, Next we tried an approach with a longer construct tethered peptide BE6 fusion, called LXXLL-BE6, which consists of the E6-binding LXXLL motif of paxillin fused to the N-terminus of full-length wild-type BPV1a E6 via a short three-residue glycine linker (called in this paper MBP-LXXLL-BE6). This construct was fused to the C-terminus of Maltose Binding Protein (MBP) via a TEV-sensitive site, yielding the following fusion protein: MBP- TEV site - paxillin peptide - BE6 (MBP-tevs-LXXLL-BE6). The fusion protein was and expressed and purified as described above for BPV E6 (Figure 7.A). The affinity purification on amylose resin raised amounts of soluble MBP-tevs-LXXLL-BE6 fusion equivalent to the amounts of soluble MBP-tevs-BE6 previously obtained. However, the long ultracentrifugation step that we routinely use for probing the quality of MBP-E6 fusions (Zanier et al., Prep 2006) indicated that the MBP-tevs-LXXLL-BE6 preparations contained very little amounts of soluble aggregates as compared to MBP-tevs-BE6 preparations. TEV digestion of the MBP-tevs-LXXLL-BE6 fusion was unusually slow, suggesting that the conformation of the LXXLL-BE6 construct hindered accessibility of the proteolytic site. By contrast to what was previously observed for the MBP-tevs-BE6 construct, the LXXLL-BE6 moiety released by TEV proteolysis did not undergo precipitation neither during the proteolysis reaction nor during the subsequent concentration step. Gel filtration also raised large amounts of monomeric LXXLL-BE6 (Figure 7.B). By contrast to unfused BPV E6, this monomeric sample could be

concentrated up to 300 M without undergoing any detectable aggregation.

Purified BPV E6 undergoes time-dependent organised self-association into fibrous structures whereas LXXLL-fused BPV E6 does not.

We have also recently found that purified, initially monomeric HPV E6 proteins slowly self-associate into organised fibrous structures (Zanier K. et al., manuscript in preparation). We investigated whether BPV E6 would form such structures. Gel-filtrated samples of either unfused BPV E6 or LXXLL-BE6 fusion protein were stored at 4°C for one week and were subsequently adjusted to a concentration of 30 ng/μl, deposited on a carbon-coated grid and negatively stained using a uranyl acetate solution. The BPV E6 preparations present numerous fiber-like structures while the LXXLL-BE6 fusion preparations do not present such structures. These data confirm that the presence of the tethered LXXLL peptide efficiently prevents the self-association of BE6 protein (data not shown).

An alternative strategy for obtention of concentrated samples of BE6 non-covalently bound to the LXXLL peptide.

We searched for an alternative approach, which would allow us to obtain large amounts of folded soluble BE6 protein non-covalently bound to its cognate peptide. However, we also wanted to keep taking advantage of the solubilising effects of both the Maltose Binding protein and the LXXLL paxillin peptide. To this aim alternatively, we have made built an expression vector a new construction, which consists allowing to express the following chimeric protein: of the MBP (or GST) complex with paxillin peptide - TEV site - the E6-binding LXXLL motif of paxillin BE6 (MBP-LXXLL-tevs-BE6). Fused to the N-terminus of full-length wild-type BPV1a E6 via a linker TEV protease sensitive (called MBP-LXXLL-TEV-BE6). This fusion protein was highly expressed as a soluble form. By contrast to MBP-tevs-BE6 construct, and in agreement with the MBP-tevs-LXXLL-BE6 construct, the amylose preparations of MBP-LXXLL-tevs-BE6 contained only minor amounts of aggregated material when submitted to overnight ultracentrifugation. This design permits the obtention of BE6 non-covalently bound after TEV digestion, but at high concentration because we used a chemical synthetic peptide of E6-binding LXXLL motif of paxillin to optimise the purification of the protein.

The small-scale mini-copurification of BE6 on the glutathione resin permits to obtain the protein BE6 as a pure form after elution, but we have observed a better output if we use the amylose resin for the co-purification (Figure 6 (C and D)).

For the purification, the samples are purified by affinity on MBP- amylose resin, then the MBP-

LXXLL-tevs-BE6 preparation was eluted complex subjected to TEV digestion simultaneously with addition of a four-fold stoichiometric excess of the chemical synthetic LXXLL paxillin peptide in excess (for times of concentration). Allowing the released BE6 moieties to be displaced from the MBP-fused LXXLL peptides to the small synthetic peptides. The BPV E6 moiety, supposedly bound to the target peptide, was then separated from MBP-paxillin peptide and TEV by gel-filtration. Non-covalently complex of the gel-filtrated sample BPV E6 with peptide from paxillin was eluted and can be could be concentrated up to 300 μM as well as just like the covalent the covalently LXXLL-BE6 fusion complex of the same partners described above, but in striking contrast with the unfused and non-complexed BE6 protein, which was always recovered in low amounts from gel-filtration and did not stand being concentrated above 50 μM .

We also constructed a GST-LXXLL-tevs-BE6 construct that we attempted to produce in a similar manner (data not shown). However, the solubility of this construct was poor as compared to that of the MBP-LXXLL-tevs-BE6. This suggests that the Maltose Binding Protein acts in synergy with the LXXLL peptide for increasing the soluble amounts of BE6 protein during the bacterial production step.

Production of (^{15}N , ^{13}C)-labelled samples and NMR analysis of BE6 covalently or non-covalently bound to the LXXLL peptide.

The expression of both the MBP-tevs-LXXLL-BE6 and MBP-LXXLL-tevs-BE6 fusion constructs was then reproduced in appropriate minimal media, giving rise to ^{15}N - and (^{15}N , ^{13}C)-labeled samples that were subsequently purified, concentrated up to 300 μM and submitted to standard multidimensional NMR measurements. The ^1H , ^{15}N -HSQC spectrum of the covalent LXXLL-BE6 fusion (Figure 8) displays a very good quality and is typical for a folded protein. Approximately 147 amide resonances can be counted in the HSQC spectrum. This number is close to that predicted from the sequence of the LXXLL-BE6 construct (consisting of 155 amino acids of which 6 are proline residues). The ^1H , ^{15}N -HSQC spectrum of the non-covalent complex of BE6 with non-labelled synthetic LXXLL peptide (Figure 9) also displays a good quality. We have also reproduced the construct MBP-LXXLL-TEV-BE6 in minimal media ^{15}N , and we obtained a good ^1H , ^{15}N -HSQC spectrum. The superimposition of the two spectra shows the absence of about 10-15 peaks in the non-covalent complex, which must account for the fact that the which represent only the amides peaks of the protein BE6 in complex with the non-labelled synthetic LXXLL peptide is non-labelled (Figure 10).

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 Glutathione-S-Transferase (GST) generate no or little solubilization of E6 (Kapust et al.). 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 in which the E6 moieties are self-associated under a biologically inactive form. We have previously developed various assays to detect these soluble aggregates and used ultra-centrifugation approaches to eliminate them (Nominé *et al.*, 2001a; Nominé *et al.*, 2001b; Zanier *et al.*, 2007). However, the folded E6 proteins obtained by these procedures generally remain prone to self-association at high concentration (Zanier et al., manuscript submitted to publication), hindering their biophysical and structural analysis.

The BPV1a E6 protein (BE6) studied in this work generally behaves in a manner very comparable to that of the E6 proteins studied to date. GST-BE6 fusions are poorly soluble (data not shown) whereas MBP-BE6 fusions are soluble but largely under the form of soluble aggregates. These aggregates can be cleared by ultracentrifugation, allowing one to isolate diluted amounts of monomeric BE6 after proteolytic removal of the MBP tag. However, the BE6 samples obtained in this manner do not stand the concentration steps required for biophysical analysis.

To turn around this problem, we have searched for means of improving the solubility of BPV-1a E6. The phage display studies revealed a strong preference of BE6 for peptides containing the LXXLL consensus, whereas BIAcore indicated that BE6 formed a strong, slowly dissociating complex with a LXXLL peptide derived from paxillin. In addition, previous studies had shown that a construct of BE6 fused to the C-terminus of the same paxillin-derived LXXLL peptide was inactivated, *in vivo*, for biological functions requiring interactions with partners possessing a LXXLL E6-binding motif, suggesting that the fused peptide was efficiently blocking the LXXLL-binding site of BE6. The conserved leucine residues of the LXXLL motif are likely to interact with a hydrophobic patch exposed on the surface of BE6, which may promote aggregation of the unbound protein. We therefore reasoned that BE6 solubility might be improved in the presence of its cognate peptide. We tested several alternative strategies. On the one hand, approaches consisting in stabilising BE6 via an intermolecular interaction with the cognate peptide (based either on co-expression or mixing at the early steps of purification) were unsuccessful or yielded insufficient amounts of complexes. On the other hand, approaches consisting in stabilising BE6 via an intramolecular interaction with its cognate peptide,

provided immediately at the protein synthesis level, turned out to be remarkably successful, as it allowed the obtention of high yields of soluble and non-aggregated material.

To purify highly concentrated materials out of MBP-peptide-BE6 fusions, we used two alternative strategies, which were equally successful. One strategy, based on the MBP-tevs-LXXLL-BE6 construct, led to the obtention of a covalent peptide-BE6 fusion construct, which proved highly soluble at high concentrations and amenable to biophysical analysis by NMR. The second strategy, based on the MBP-LXXLL-tevs-BE6 construct, led to the obtention of a non-covalent BE6-peptide complex, which was equally soluble at high concentrations and is also amenable to NMR studies. The failure of all our alternative attempts to isolate BE6 in the absence of its cognate peptide indicates the absolute requirement of the peptidic ligand for solubilization of BE6.

In a parallel work, we have tested the same approaches to try to stabilise the HPV 16 E6 (16E6) protein via its cognate peptide derived from the E6AP ubiquitin ligase (data not shown). The use of MBP-tevs-LXXLL-16E6 and MBP-LXXLL-tevs-16E6 fusions constructs, as well as GST-peptide / 16E6 co-expression or co-purification strategies, were all unsuccessful in stabilizing efficiently the 16 E6 protein. This may be related to differences in affinity and more specifically to the differences in kinetic off-rates of the complexes. Preliminary SPR analysis of the BE6-paxillin peptide complex (Fig. 5) showed that this complex displays a much slower off-rate as compared to the 16E6-E6AP peptide complex (Zanier *et al.*, 2005). Consequently, the BE6-paxillin peptide complex is strong enough to survive purification steps and to prevent self-association of the protein at high concentration. In the covalent tethered peptide fusions, the low affinity and rapid off-rate of the 16E6-peptide complex may result in a rapid-exchanging close-open form which may undergo aggregation via either intermolecular E6-protein interactions or self-association of hydrophobic surfaces of 16E6 not fully protected by the peptide. By the contrast, the strong affinity and low off-rate of the BE6-peptide complex may favour intramolecular versus intermolecular E6-peptide interactions and ensure a better protection of hydrophobic surfaces of BE6, thereby preventing aggregation at high concentration.

In our previous work, we have shown that mutagenesis of non-conserved cysteins to the isosteric non-oxidable residue serine notably increased the purificability and concentrability of recombinant HPV16 E6 protein (Nominé *et al.*, 2001). We have tried to reproduce this approach by producing various mutants of BPV E6 in which all or some of the non-conserved cysteins were mutated to serins. None of these mutants displayed improved solubility as compared to the wild-type BPV E6 protein (data not shown). 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), and each of these approaches succeeded for one protein but failed for the other

one. This shows that, even for closely related proteins, there are no definite standard recipes for enhancing sample solubility.

SPR analysis (this paper and previous data by Zanier et al., 2005) has shown that BE6 and 16 E6 both bind to peptides possessing the consensus LXXLL. However, we observed (data not shown) that 16E6 does not recognize the paxillin-derived peptide, whereas BE6 does not interact with the E6AP-derived peptide. This confirms previous functional data suggesting that 16E6 does not target paxillin and that BE6 does not target E6AP. In addition, this also suggests that the structural determinants of the E6-LXXLL interaction must possess distinctive traits, which result in the observed difference in selectivity.

We also observed that BE6 interacts stronger with its cognate motif within paxillin, than 16 E6 does with the equivalent motif within E6AP. This may have functional implications. The BE6-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.

Here 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, which are only stable and/or soluble in the presence of a peptidic partner.

ACKNOWLEDGMENTS

This work was supported by CNRS, University of Strasbourg, Association de Recherche contre le Cancer (ARC), Ligue Nationale Contre le Cancer, Agence Nationale de la Recherche (ANR), and the N.I.H. A.O.M.O.S. was supported by a grant of the ARC.

REFERENCES

- 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) May 25;271(1):163-70.
- Bosch FX.** Epidemiology of human papillomavirus infections: new options for cervical cancer prevention. *Salud Publica Mex.* (2003) ;45 Suppl 3:S326-39.
- Brimer N, Lyons C, Vande Pol SB.** Association of E6AP (UBE3A) with human papillomavirus type 11 E6 protein. *Virology*. (2007) Feb 20;358(2):303-10.
- Chakrabarti O, Krishna S.** Molecular interactions of 'high risk' human papillomaviruses E6 and E7 oncoproteins: implications for tumour progression. *J Biosci.* (2003) Apr;28(3):337-48. Review.
- 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) May 29;273(22):13537-44.
- Daniels PR, Sanders CM, Coulson P, Maitland NJ.** Molecular analysis of the interaction between HPV type 16 E6 and human E6-associated protein. *FEBS Lett.* (1997) Oct 13;416(1):6-10.
- Das K, Bohl J, Vande Pol SB.** Identification of a second transforming function in bovine papillomavirus type 1 E6 and the role of E6 interactions with paxillin, E6BP, and E6AP. *J Virol.* (2000) Jan;74(2):812-6.
- de Villiers EM, Fauquet C, Broker TR, Bernard HU, zur Hausen H.** Classification of papillomaviruses. *Virology*. (2004) Jun 20;324(1):17-27. Review.
- Degenkolbe R, Gilligan P, Gupta S, Bernard HU.** Chelating agents stabilize the monomeric state of the zinc binding human papillomavirus 16 E6 oncoprotein. *Biochemistry.* (2003) Apr 8;42(13):3868-73.
- Elston RC, Naphine 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) Feb;79 (Pt 2):371-4.
- Glaunsinger BA, Lee SS, Thomas M, Banks L, Javier R.** Interactions of the PDZ-protein MAGI-1 with adenovirus E4-ORF1 and high-risk papillomavirus E6 oncoproteins. *Oncogene.* (2000) Nov 2;19(46):5270-80.
- 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) Aug;13(8):4918-27.
- Lipari F, McGibbon GA, Wardrop E, Cordingley MG.** Purification and biophysical characterization of a minimal functional domain and of an N-terminal Zn²⁺-binding fragment from the human papillomavirus type 16 E6 protein. *Biochemistry.* (2001) Feb 6;40(5):1196-204.
- 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 oncoprotein: biophysical characterization of the E6 C-terminal DNA-binding domain. *Biochemistry.* (2003) May 6;42(17):4909-17.
- Nominé Y, Masson M, Charbonnier S, Zanier K, Ristriani T, Deryckère F, Sibler AP, Desplancq D, Atkinson RA, Weiss E, Orfanoudakis G, Kieffer B, Travé G.** Structural and functional analysis of E6

oncoprotein: insights in the molecular pathways of human papillomavirus-mediated pathogenesis. *Mol Cell*. (2006) Mar 3;21(5):665-78.

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)b Apr;14(4):297-305.

Nominé Y, Ristriani T, Laurent C, Lefèvre JF, Weiss E, Travé G. Formation of soluble inclusion bodies by hvp e6 oncoprotein fused to maltose-binding protein. *Protein Expr Purif*. (2001)a Oct;23(1):22-32.

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) Sep 15;18(18):5061-72.

Pfister H. Chapter 8: Human papillomavirus and skin cancer. *J Natl Cancer Inst Monogr*. (2003) ;(31):52-6.

Romier C, Ben Jelloul M, 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) Oct;62(Pt 10):1232-42.

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) Jul 1;12(13):2061-72.

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) Nov 5;75(3):495-505.

Thomas M, Laura R, Hepner K, Guccione E, Sawyers C, Lasky L, Banks L. Oncogenic human papillomavirus E6 proteins target the MAGI-2 and MAGI-3 proteins for degradation. *Oncogene*. (2002) Aug 1;21(33):5088-96.

Tong X, Salgia R, Li JL, Griffin JD, Howley PM. The bovine papillomavirus E6 protein binds to the LD motif repeats of paxillin and blocks its interaction with vinculin and the focal adhesion kinase. *J Biol Chem*. (1997) Dec 26;272(52):33373-6.

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) Jan 8;16(1):43-52.

Zanier K, Charbonnier S, Baltzinger M, Nominé Y, Altschuh D, Travé G. Kinetic analysis of the interactions of human papillomavirus E6 oncoproteins with the ubiquitin ligase E6AP using surface plasmon resonance. *J Mol Biol*. (2005) Jun 3;349(2):401-12.

Zanier K, Nominé Y, Charbonnier S, Ruhlmann C, Schultz P, Schweizer J, Travé 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) Jan;51(1):59-70.

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) Aug;73(8):6209-19.

zur Hausen H. Papillomaviruses in human cancers. *Proc Assoc Am Physicians.* (1999) Nov-Dec;111(6):581-7.

FIGURES LEGENDS

Figure 1

Alignment of E6 sequences (upper half E6N and lower half E6C) from the BPV and 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).

Figure 2

(A) The resulting MBP-fusion (called MBP-tevs-BE6) was overexpressed (lane 1), purified by affinity on MBP-specific amylose resin (lanes 2-3-4), ultracentrifugated for eliminating large soluble aggregates (lane 4), then and subjected to TEV digestion yielding a 40 kDa protein corresponding to MBP and an approximatively 20 kDa protein corresponding to BPV E6 (lane 5).

(B) Gel filtration profiles of BE6N and BE6C preparations after separation from the MBP tag. Affinity-purified MBP-E6 domain preparations were cleared from soluble aggregates by ultracentrifugation and digested using TEV protease. Subsequently, 2 ml were injected on a Superdex75 16/60 column. Peaks corresponding to the two domains are indicated.

Figure 3

^1H - ^{15}N HSQC spectrum of the BPV E6-N construct recorded at pH 6.8 and 298 K.

Figure 4

Translated DNA sequences of Phage Peptides Isolated by Binding to BE6. Six plaques selected after five rounds of phage display selection were sequenced across the M13 gene III-random peptide junction. Each plaque contained a unique sequence conforming to the LXXLL motifs of E6AP and paxillin.

Figure 5

Analysis of the interaction of MBP-BE6 with GST-LXXLL. The Figure illustrates concentration-dependent binding of BE6 to GST-LXXLL ligand. The dissociation phase is remarkably slow, the half-life of the complexes, estimated from the dissociation profiles, approximated 600 s.

Figure 6

(A) The BE6 was produced and micropurified on affinity resin. The sample (lane 1) was centrifuged and both pellet and supernatant resuspended in equal volumes and subsequently analyzed by SDS-PAGE. Lane 2: supernatant, line 3: pellet. The BE6 was co-expressed with the GST-LXXLL from Paxillin, Lane 4: mini co-purified BE6/GST-LXXLL, Lane 5: supernatant, lane 6: the pellet. **(B)** The co-purification of the produced unfused BE6 with the produced expression of GST-LXXLL. Lane 1: mixture of pellets, lane 2, 3: incubation of the samples with glutathione resin, lane 4 co-elution of the BE6 with the GST-LXXLL. **C and D:** mini co-expression and mini co-purification of the unfused sequence of BE6 with the MBP-LXLL or GST-LXXLL respectively. Pellets of BL21 CodonPlus bacteria expressing the constructs (lane 1) were sonicated and then ultracentrifuged. The cleared supernatant (lane 2) was incubated with the amylose resin (**C**), or glutathione resin (**D**). Lane 3: wash, lane 4: the elution, line 5: the samples was subjected to proteolytic cleavage on resin giving rise to elute the unfused BE6.

Figure 7

(A) The purification of the construct of MBP- TEV site - paxillin peptide - BE6 (MBP-tevs-LXXLL-BE6). The construct was overexpressed (lane 1) and purified by affinity on MBP- amylose resin (lane 2), then The MBP-tevs-LXXLL-BE6 preparation was ultracentrifuged (lane 3) and then subjected to TEV digestion (lane 4) followed by a gel-filtration separation of BE6 from MBP (lane 5). **(B)** The gel filtration profiles of LXXLL-BE6 and the unfused BE6 preparations after separation from the MBP. Peaks corresponding to the two constructs are indicated.

Figure 8

The ^1H - ^{15}N HSQC spectrum of the LXXLL-tevs-BE6 fusion construct.

Figure 9

The ^1H - ^{15}N HSQC spectrum of ^{15}N labelled BPV E6 bound to unlabelled LXXLL peptide.

Figure 10

Superposition of ^1H - ^{15}N HSQC spectra of ^{15}N labelled LXXLL-BPV E6 fusion (black) and of ^{15}N labelled BPV E6 bound to unlabelled LXXLL peptide (red).

Table 1

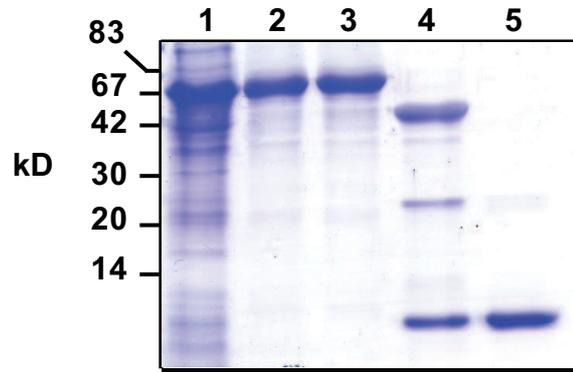
Titers of bound and released phage after 5 rounds of selection of phage displayed random peptides. 3×10^{10} input phage were bound to 1 μg immobilized CBD-TEV-E6 immobilized on chitin beads. After washing away unbound phage, bound phage were released by TEV protease and titered.

BPV1-nt MDLKPFA^RTNPFSGLDCLWCREPLTEVDA^FFR^CMMV^KDFHVVIR. EGCRYGACTICLENCLATERRLWQ
 HPV16Rnt dPQERPRKLPQLCTELQTTIHDIILECVYCKQQLRRRE^VYDFAF^RDL^CIVYR. DGNPYAVCDKCLKFYSKI^EYRHY
 HPV18Rnt dPTRRPYKLPDLCTELN^TSLQDI^EITCVYCKTVLELT^EVFEF^AFKDLFV^VYR. DSIPHAACHKCIDFYSRIR^ELRHY
 HPV34nt nPEERPYKLPALCEEVNI^SIHEIE^LDCVY^CERQLYRCE^VYDFIFRDL^CVVYR. KGKPLGVCQPCLLFYSKVR^QYRRY
 HPV26nt dPRERPRTLHELCE^SLN^TTLQNLQV^QCVYCKETLQWAD^VYNFAICDLR^VVYR. DRSPYAACKRCVIFYSKITEY^RRY
 HPV53nt nTEERPRTLHQ^LCEV^VNKPLLELQ^LGCV^FCKKALTAS^EVYNFAYTDLR^VVYR. DGYPYGVCKFCLLFYSKVR^KLRYY
 HPV40nt rCGSQARTLYELCDQC^NITLPTLQIDCV^FCKTVL^KTAEVLAFA^FRELYV^VWR. DDFPHAACPRCLDLHGKVN^QYRNF
 HPV27nt eENPCPRNIFLLCKQY^GLELEDLRL^LLCVY^CRRALS^DAVLAF^AIKEL^SVVWR. KGFPFGACGKCLIAAGKLR^QYRHW
 HPV11Rnt dASTATSIDQLCKTFNLSL^HTLQIQ^CVFCRNAL^TTAELIYAY^AYNLKV^VWR. DNFPFAACACCLELQ^GKIN^QYRHF
 HPV72nt MGLHNP^TNIWLLCKEIEVDLE^DLRITC^IFC^KNELTTE^ELLAIAIKEL^QIVWR. DNWPFGVCAPCLARAT^KVRELRYW
 HPV10nt MGAQEP^RNILL^LCRNCGIPLE^DLRLLC^IFC^TKQLTAA^ELLA^FFALRELYL^VWR. AGVPY^GACARCLL^LQGI^VRRLKYW
 HPV32nt sASSQPSTLYQLCKD^FGLTL^RNLQI^CCIW^CKNHLS^AEAYAY^HFKDLH^VVWK. KGFPYAACAFCL^EFYSK^VCALRHY
 HPV12nt sTPELPTTIKELADL^LDIPLV^DCLVPC^NFCGKFLDF^LEV^CDFDK^QLTLI^WK. GHFVTACCRSCCA^ATAI^EYEFNEFY
 HPV20nt lEPPLPATICGLAKLLEI^PLDCLIP^CNFCGN^FLTHLE^VCEFE^DEK^LTLI^WK. DHLV^FACCRVCC^SATAT^EYEFN^QFY
 HPV17nt MDRPKPQTVRELAD^TL^CIPLVD^ILLPCR^FCN^RFLAYI^ELVAF^DLKGL^QLIWTE^EDFV^FACCS^SCAYATA^QYEF^SKFY
 HPV23nt MDSTRPLTVQQLSDK^LTVPV^VDLLP^CRF^CS^RFLTYL^ELE^RFDY^KH^LQLI^WTE^EDFV^FACCS^GCAYASA^QFEI^QQFY
 HPV1Ant .MATPIRTVRQLSESLCIPY^IDVLL^PCN^FCNY^FLSNA^EKLL^FDFH^DFLH^LVWR. DNLV^FGCCQ^GCARTV^SLLE^FVLYY
 HPV48nt MEPQ^FPTDLD^SYCKY^FNI^SFFDL^VLK^CIF^CK^FSV^SIV^DLA^SFHN^KRL^SVI^WR. DNT^PFAC^TKCL^RL^TALY^EKDN^FF

BPV1-ct GVPVTGEEAELLHGK^TLDRLCIR^CCCYCGGK^LTKNEK^HRRH^VLFNE^PFCK^TRANI^IRGR^CYDC^RHGSR^SKYP
 HPV16Rct CYSLYGTTLEQQY^NKPLCD^LLIR^CINC^QKPL^CPEEK^QRHL^DKK^QR^FHNI. RGRWTGRC^MSC^RSR^TRRET^QL
 HPV18Rct SDSVYGD^TLEK^LTNTGL^NLLIR^CLCR^QKPL^NPAEK^LRHL^NEK^RR^FHNI. AGHYR^GQC^HSC^CN^RAR^QERL^QRR^RET^QV
 HPV34ct NQSVYGR^TLENL^TNK^QLCN^LLIR^CCGK^QKPL^CPLEK^QRHV^DENK^RF^HQI. ADQWTGRC^TQC^WRSAT^VV
 HPV26ct TCSVYGATLEAL^TKKSLC^NLLIR^CHRC^MPLG^PEK^QRIV^DEK^RR^FHEI. AGQW^KGL^CTN^CWR^PRR^QTET^QV
 HPV53ct NCSVYGASLEAL^TKKKLS^DLIR^CYRC^HPL^TPEEK^QLH^CDY^KK^RF^HKI. SHMWTG^SCL^TC^WRHT^TATES^AV
 HPV40ct RYAAAYAPTVEEET^GL^TIL^QV^RIR^CCK^CH^KPL^SPEK^TNH^IV^KKT^QFF^KL. KDSWTG^YCL^HC^WKK^CMEK^QR^SET^LC
 HPV27ct HYS^CYGD^TVE^TET^GI^PI^QL^FM^RCY^ICH^KPL^SWE^EK^EALL^VGN^KR^FHNI. SGRWTG^HCM^QCG^STCT^APD^PAS^RTL^H
 HPV11Rct NYAAAYAPTVEEET^NED^IL^KV^LIR^CYL^CH^KPL^CEI^EK^LK^HIL^GK^AR^FIK^L. NNQW^KGR^CL^HC^WTT^CMED^LLP
 HPV72ct TYSGYGPTVEQE^TG^KSLA^ELY^IR^CHAC^CK^PLS^QE^KEY^QV^QTG^IH^FH^KI. SGLWTG^RCC^QCG^ACTAR^WQ^P
 HPV10ct DYSYYVEGVEEET^KQ^SIY^TQ^LIR^CYM^CH^KPL^VRE^EK^DR^HR^NERR^LH^KI. SGYWR^GS^CEY^CW^SRCT^VRI^PQ
 HPV32ct DRSAF^WHTVEQE^TGL^LLE^QI^IRC^AIC^QK^PLS^PSE^KD^HHI^YNG^RH^FRF^I. LNRWTG^RCT^QCRE
 HPV12ct QQTVLGRDIE^LAT^GK^SI^FDL^KIR^CQT^CLS^FLD^TIE^KL^DSC^GR^GLP^FH^KV. RDRW^KG^ICR^QCK^HLY^LN^DR
 HPV20ct ESTV^LGRDIE^QV^TG^KSV^FDL^DV^RCY^TCM^KFL^DSI^EK^LD^ICG^RK^RF^FY^LV. RGSW^KG^ICR^LCK^HF^Q
 HPV17ct EQSVSGRELEIE^HK^PI^GE^IIR^CK^FCL^KLD^LLE^KL^DT^CY^RH^QQ^FH^KV. RRN^WK^GL^CR^HCG^SIG
 HPV23ct QLT^VY^GREIE^QE^QR^PI^GQ^IC^IR^CQ^YCL^KSL^DLI^EK^LD^IC^SFN^QP^FH^KV. RNHW^KG^RCH^CKE^IE
 HPV1Act QESYEVPEIE^EIL^DR^PL^LQ^IE^LRC^VT^CIK^KLS^VA^EK^LEV^SNG^ER^VH^RV. RNR^LK^AK^CSL^CRL^YAI
 HPV48ct VCTAK^SHLL^TGL^VKK^ELS^DI^NIR^CQ^HC^YS^FLD^YLE^KLY^HLY^ND^VD^FLLI. RGT^WR^GV^CRN^CIS^HEGR

Figure: 1

A



B

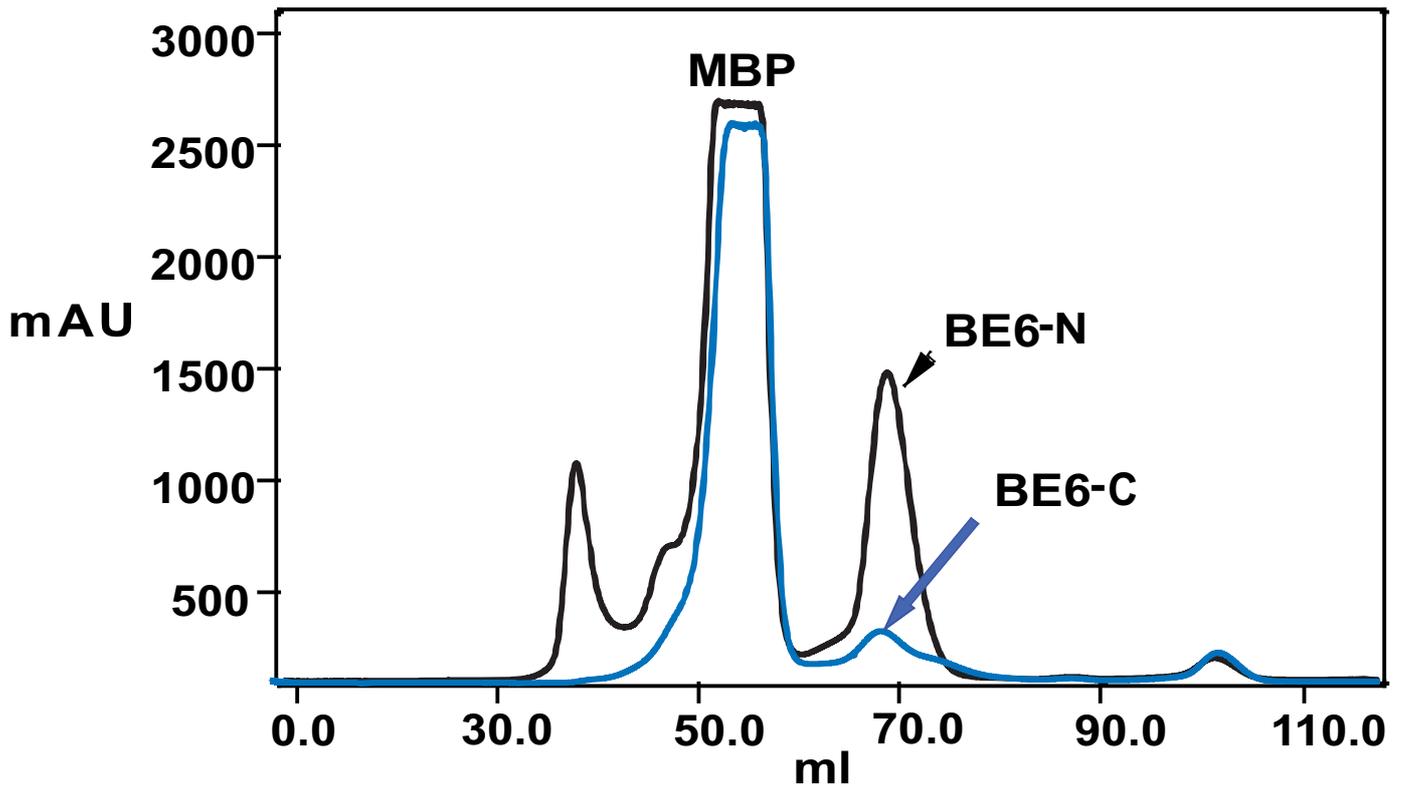


Figure: 2

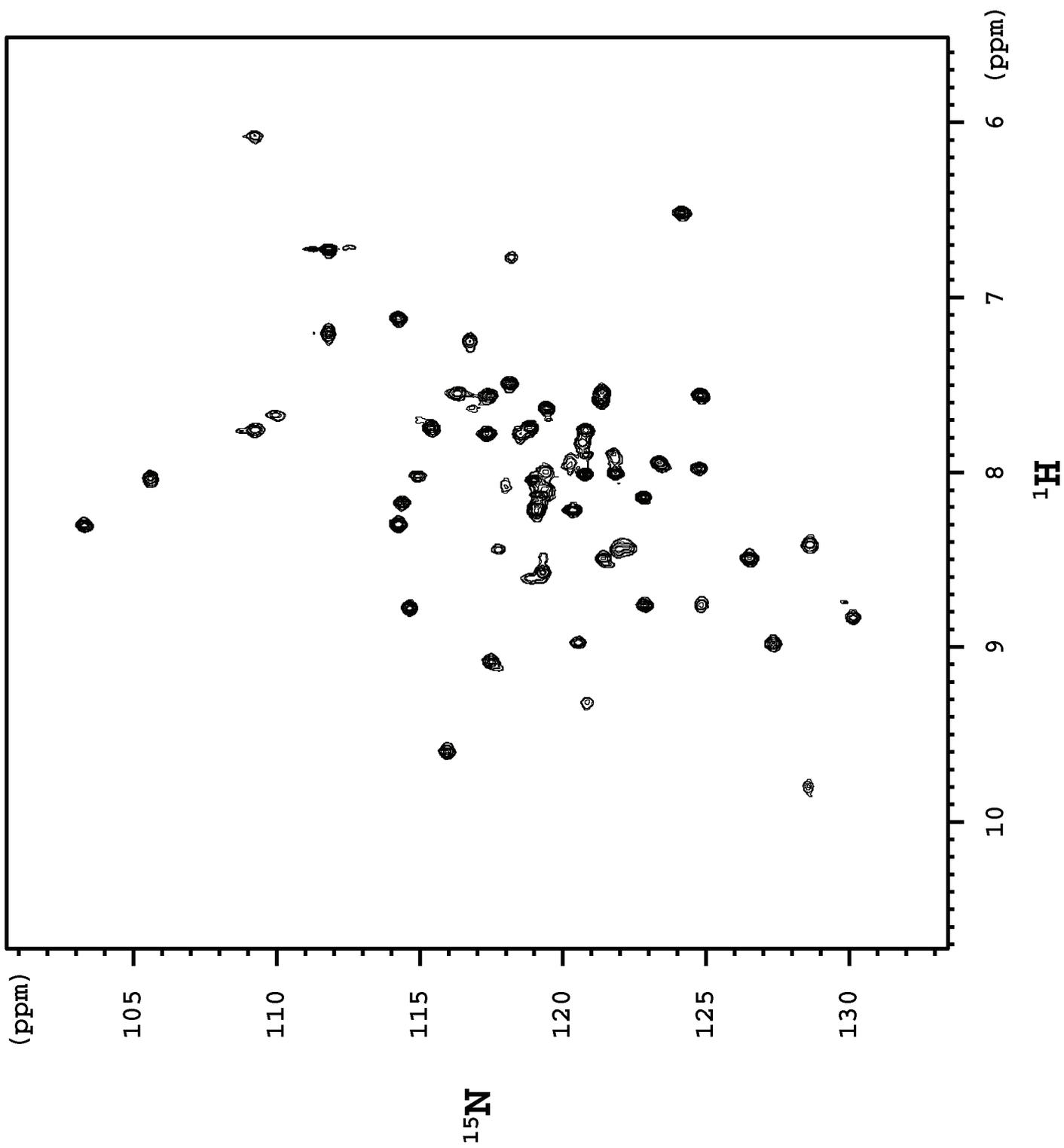


Figure: 3

E6 on Beads Released Phage Titer X 10⁵

BE6	270000
CRPV E6	1.0
1E6	36
5E6	61
11E6	800
16E6	150
Chitin Beads alone	2.9

Table: 1

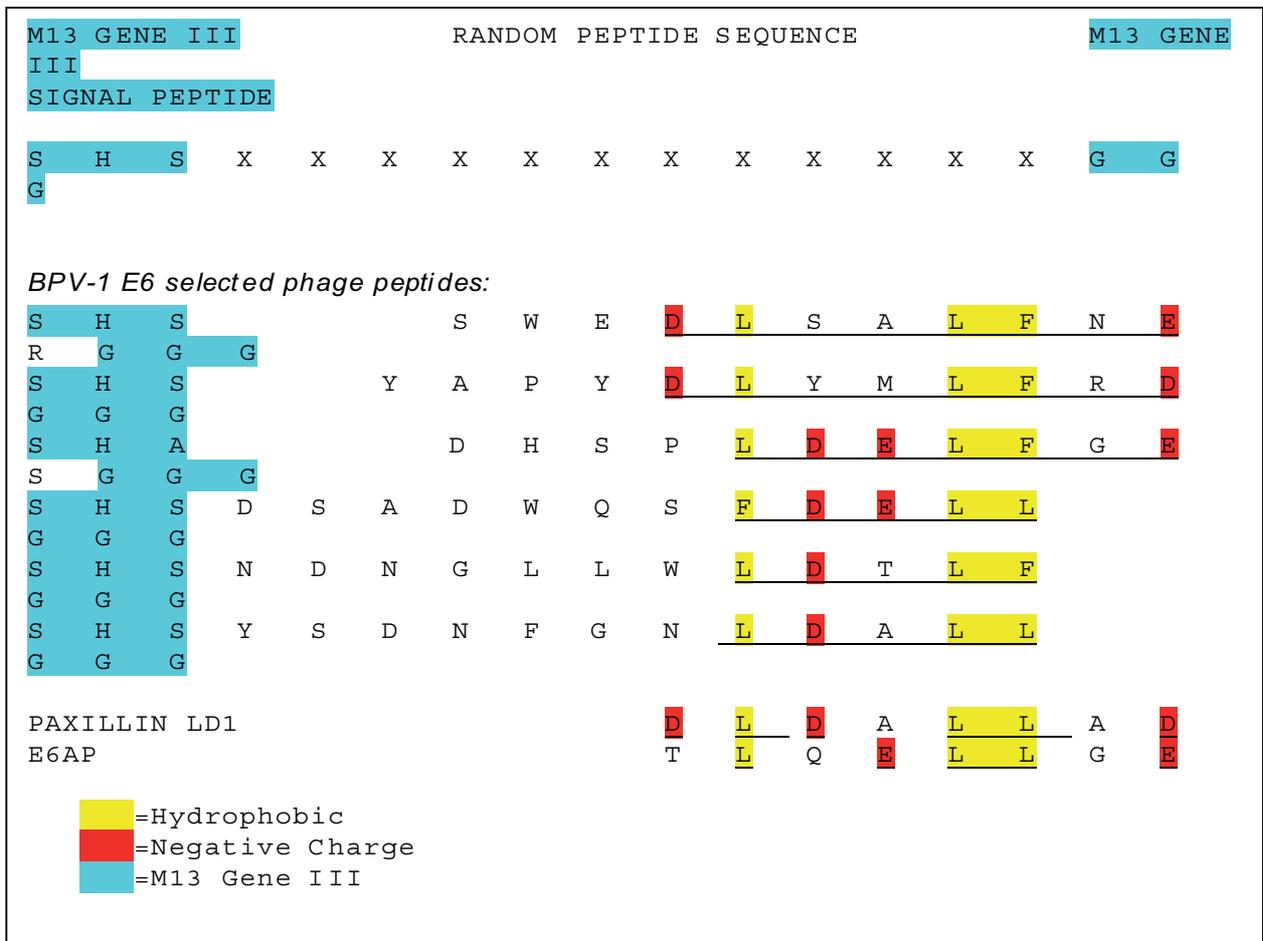


Figure: 4

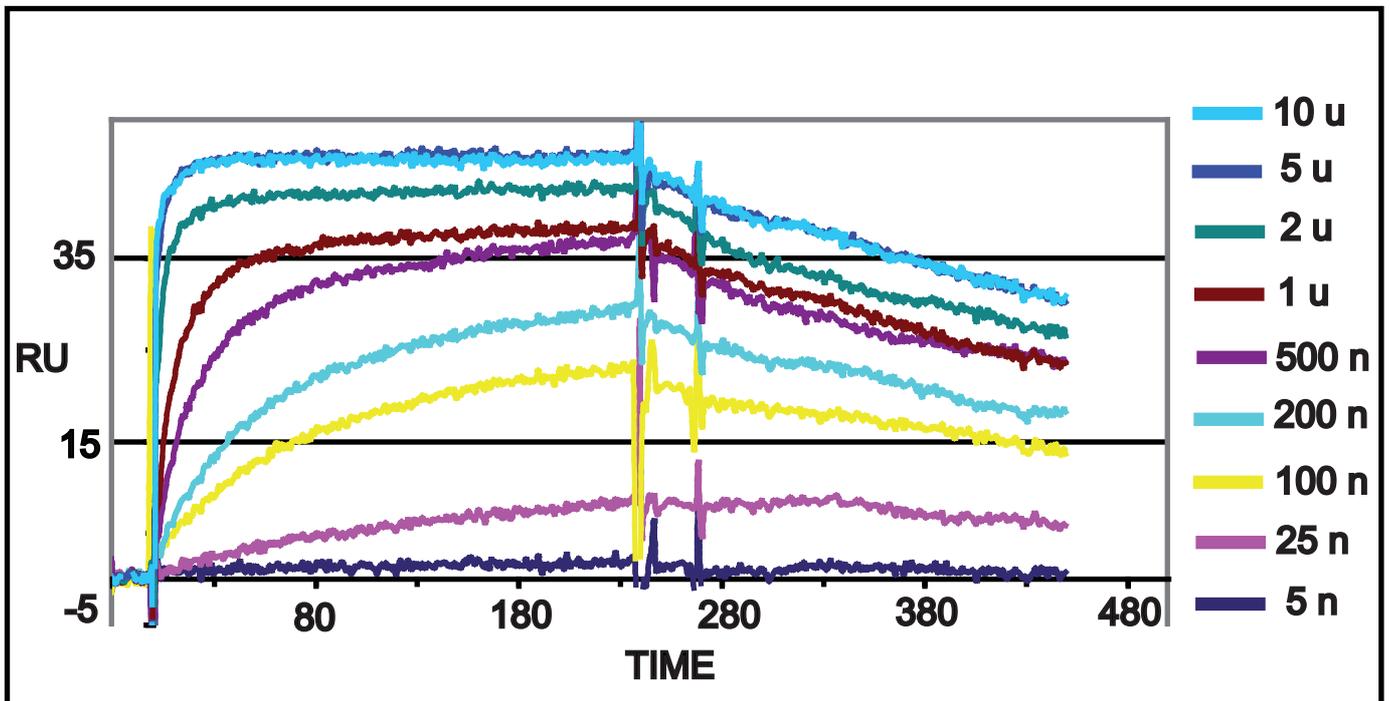


Figure: 5

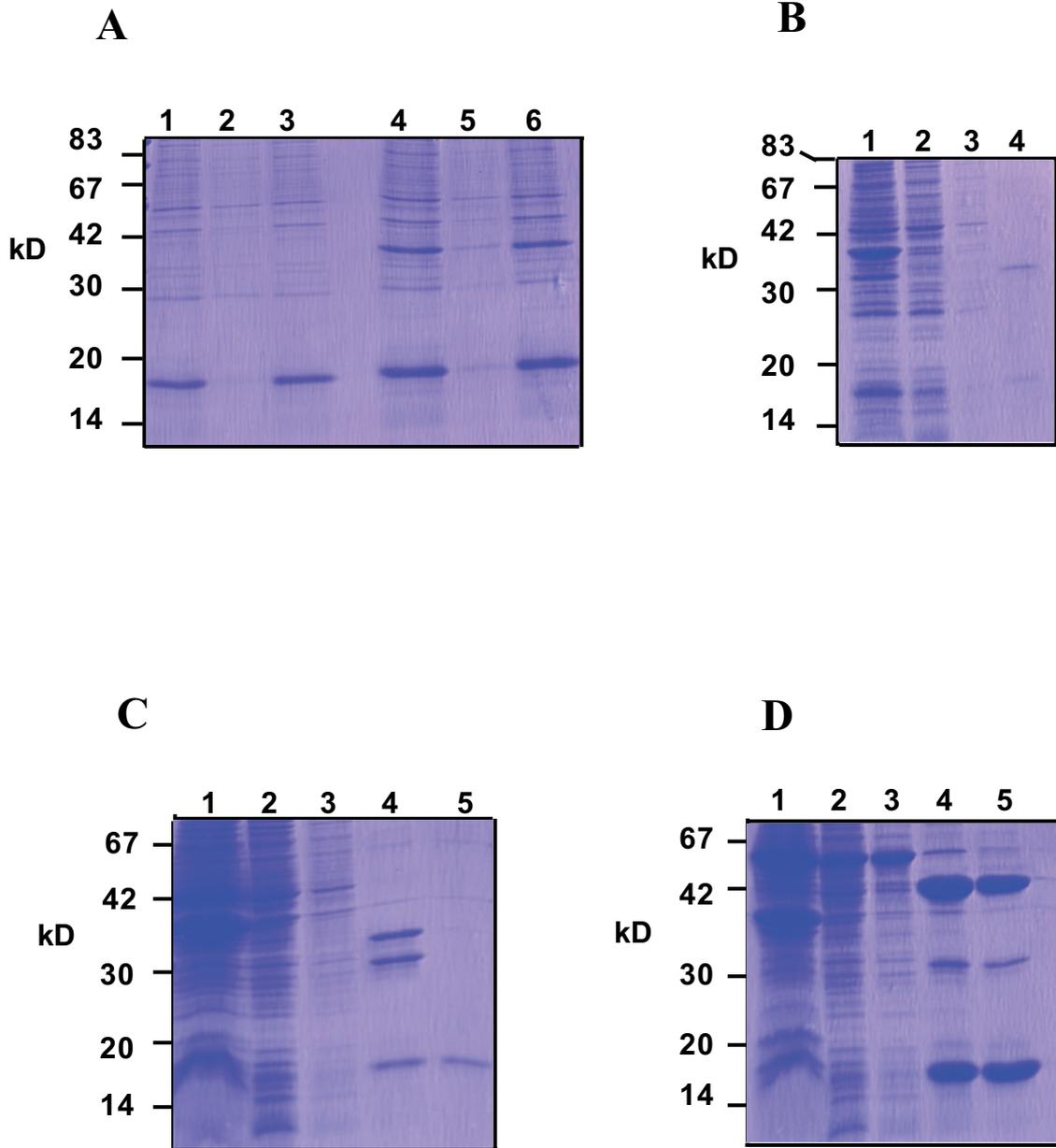
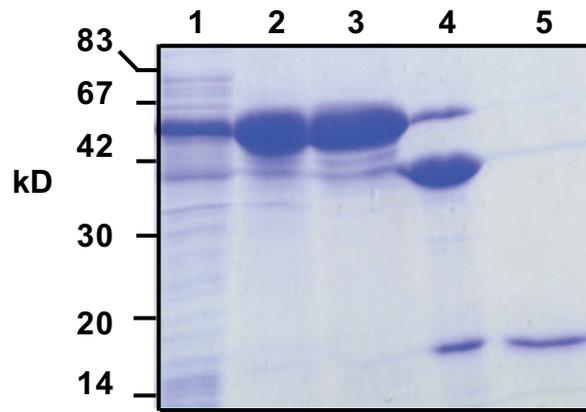


Figure: 6

A



B

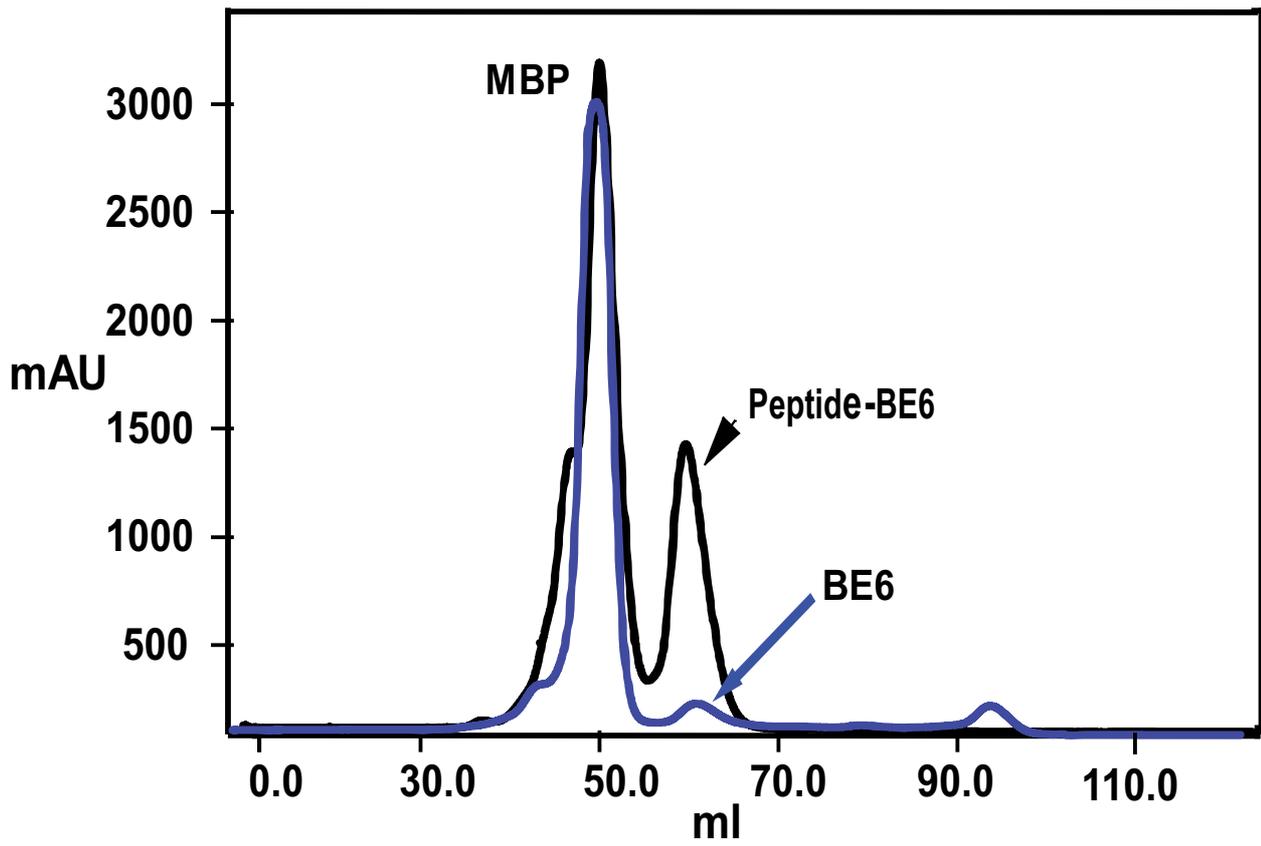


Figure: 7

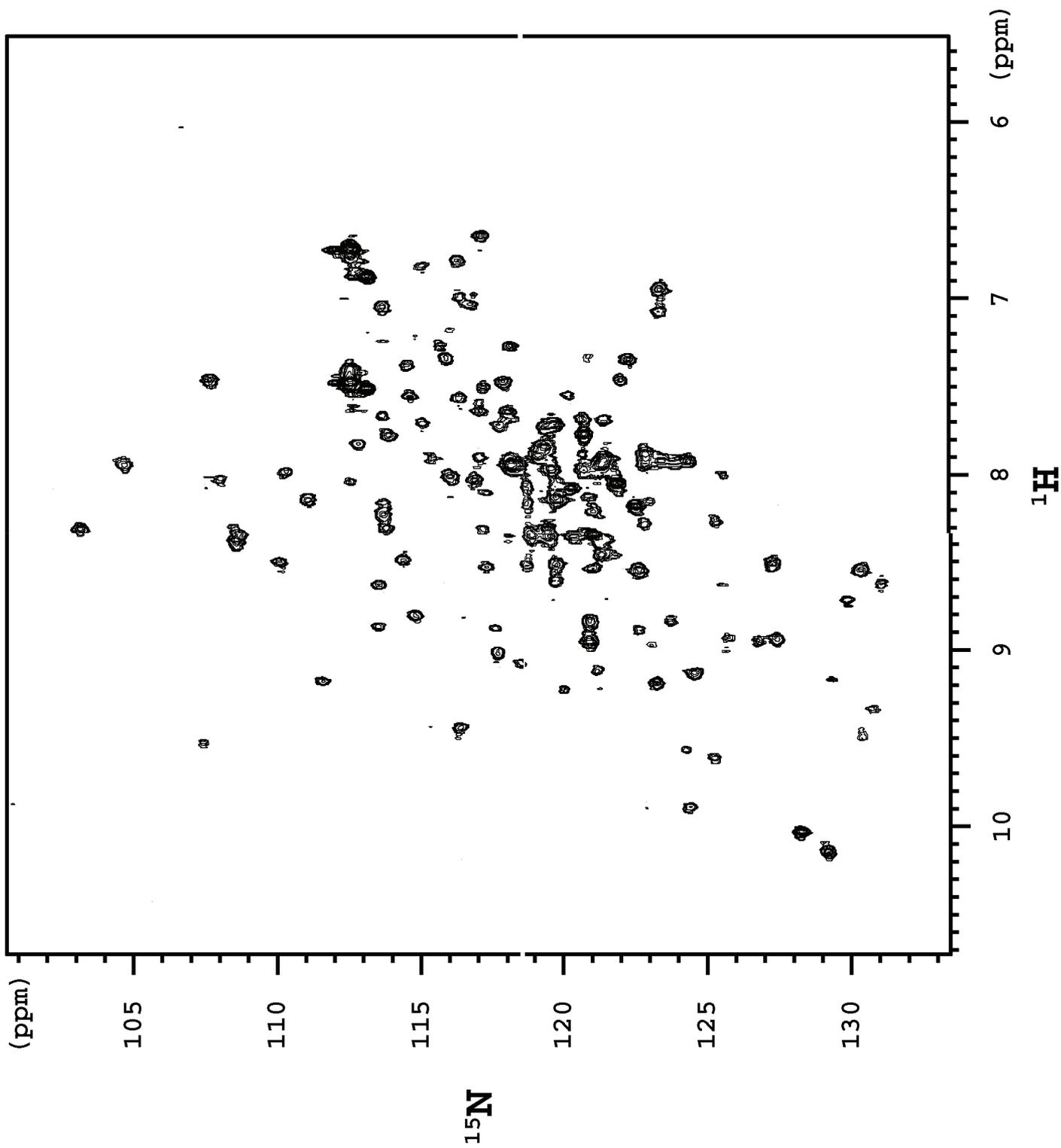


Figure: 8

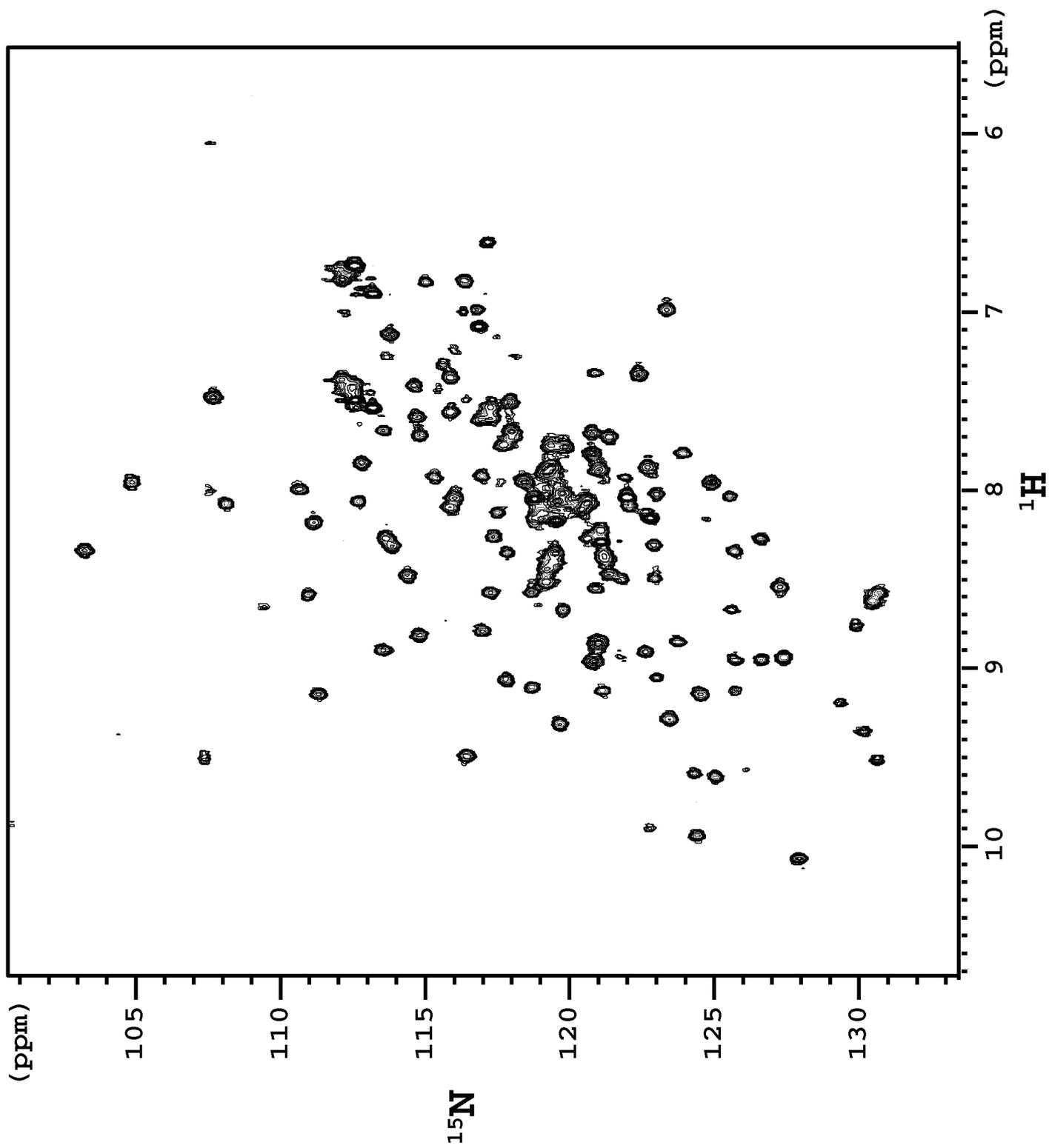


Figure: 9

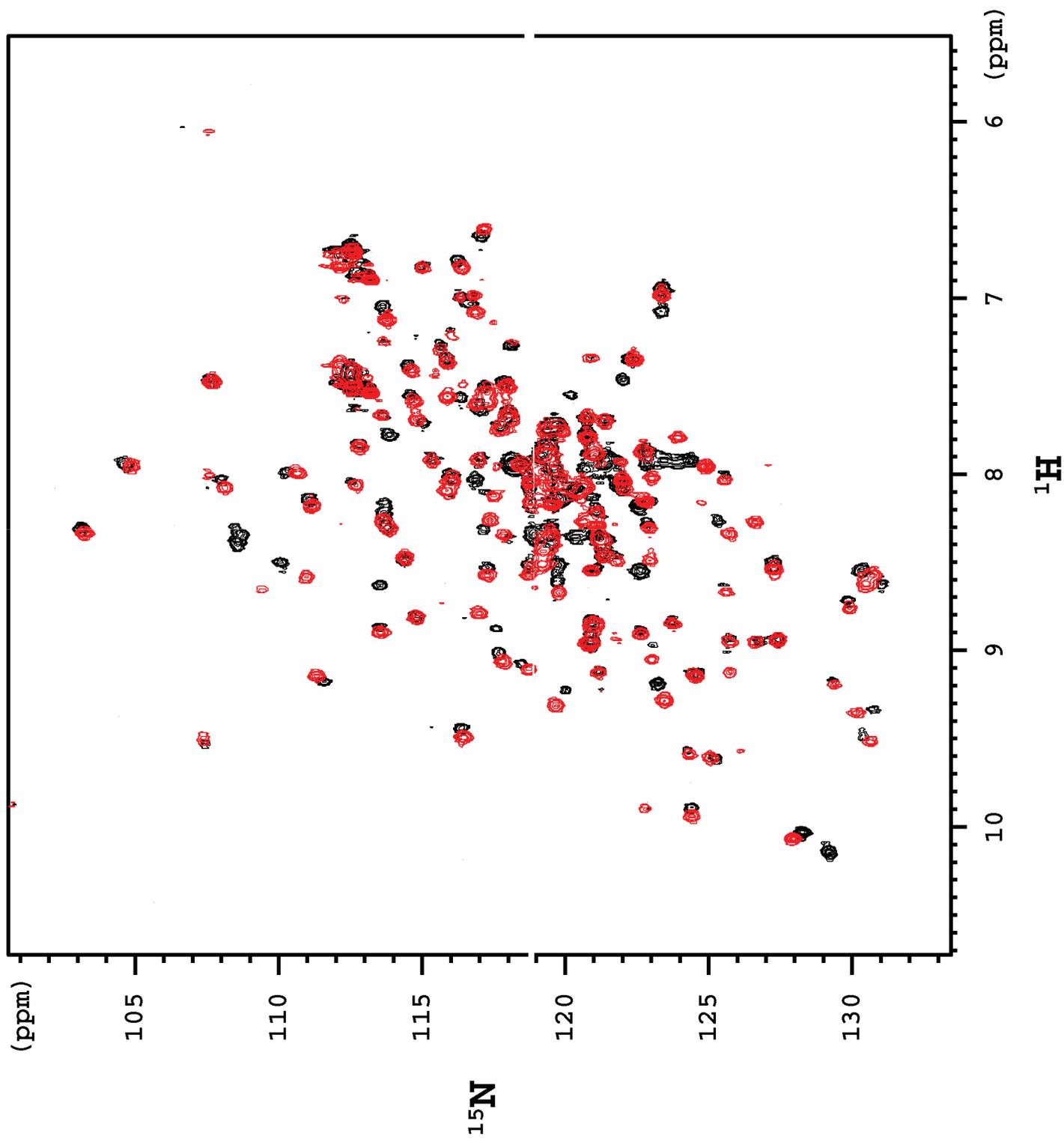


Figure: 9

Troisième partie : Identification et caractérisation des protéines précoces E6, E7 et E9 issues de papillomavirus aviaires.

Manuscrit 2:

Identification of unusual E6 and E7 proteins within avian papillomaviruses: cellular localization, biophysical characterization and phylogenetic analysis

Accepté dans Journal of Virology.

L'étude présentée dans cet article concerne un travail fondateur sur les papillomavirus aviaires, qui ont très peu été explorés jusqu'à présent. Un nouveau papillomavirus aviaire est identifié, et pour la première fois une étude de biochimie, de structure et localisation de protéines des papillomavirus aviaires est initiée. En outre, les données exposées dans cet article viennent proposer un nouveau regard sur la phylogénie des PVs.

Les gènes de E6 aviaires codent pour une petite protéine (83 et 85 résidus pour FIPV-1 et FPV respectivement). Cette protéine présente une particularité atypique parce qu'il ne contient qu'un seul domaine capable de lier le zinc. Les séquences de deux protéines E6 de FIPV-1 et FPV sont identiques à 39%. L'alignement des séquences de deux protéines avec les domaines N-terminal (E6-N) et C-terminal (E6-C) de HPV16 E6 révèle l'existence des résidus conservés qui jouent un rôle important dans la conformation structurale de la protéine. Les quatre cystéines impliquées dans la liaison de zinc sont strictement conservées ainsi que quelques résidus qui se trouvent à l'intérieur du repliement de E6-N et E6-C. Ceci laisse émerger la théorie selon laquelle une forme ancestrale de E6 existait il y a longtemps et s'est transformée avec l'évolution des PVs.

La protéine E7 aviaire est quand à elle plus typique par rapport à ses orthologues mammifères. Elle comprend environ 60 à 100 résidus selon le codon d'initiation utilisé comme départ et elle contient deux domaines: le domaine N-terminal, prédit comme non structuré, et comprenant le motif consensus LxCxE largement conservé dans les E7 de PVs et important dans la liaison avec pRb; et le domaine C-terminal, avec les deux motifs CxxC liant le zinc retrouvés dans tous les autres domaines C-terminaux de E7 des PVs. Nous avons également révélé l'existence d'un nouveau cadre de lecture (ORF) (couramment désigné ORFx) qui code pour une protéine précoce nommée E9. Les résultats obtenus sur cette dernière ne sont pas inclus dans l'article mais trouveront place dans une publication ultérieure.

Après avoir reçu les génomes viraux de la part de nos collaborateurs, nous avons clonés les séquences qui correspondent aux ORFs des protéines E6, E7 et E9 dans des vecteurs d'expression de protéines recombinantes (pETM-41) permettant à la protéine d'intérêt d'être fusionnée à une autre protéine MBP par un "linker" sensible à la protéase TEV. Les mêmes séquences de E6, de E9 ainsi que de E7 ont été également clonées dans des vecteurs d'expression eucaryotes permettant d'ajouter une séquence GFP ou myc en amont du N-terminus ou bien un signal myc en aval du C-terminus la protéine.

Nous avons adopté des protocoles optimisés pour la purification de ces protéines recombinantes. Ceci a eu un effet très positif puisque nous avons dépassé aisément le seuil critique d'agrégation des échantillons E6, et avons obtenu des échantillons très concentrés (plus de 300 μ M pour E6 et E7 de FIPV). La protéine E6 du FPV (l'autre papillomavirus aviaire abordé dans cet article) est peu soluble et commence à s'agréger durant les étapes de la purification. L'explication plausible mise en exergue dans l'article est le grand nombre d'acides aminés hydrophobes dans la séquence de FPV E6, et en particulier sur des positions prévues pour être exposées à sa surface.

La gel-filtration analytique a montré que la protéine E6 FIPV (9,5 kDa) est éluée à un volume correspondant à un monomère de 9,2 kDa. Ceci est d'ailleurs confirmé par l'ultracentrifugation analytique qui exclut la présence d'oligomères. Par contre, la partie C-terminale de E7 FIPV se comporte comme un dimère stable de 13 kDa alors que le poids moléculaire attendu d'un protomère est 6,5 kDa. Enfin, la protéine E9 recombinante ne supporte pas l'étape de concentration une fois séparée de MBP et commence à s'agréger et précipite dans la colonne de gel filtration.

Les deux protéines E6 et E7 de FIPV ont ensuite été exprimées dans un milieu marqué (M9, 15 N). Ces échantillons marqués atteignent également sans difficulté des concentrations de 300 μ M. Les mesures RMN bidimensionnelles (1 H, 15 N) de E6 FIPV ont montré un spectre typique d'une protéine repliée non agrégée. L'ensemble des résidus sont représentés par des taches (pics) dans ce spectre homogène. Pour la région C-terminale de E7 FIPV, le spectre HSQC démontre la bonne qualité et le repliement de la protéine. Nous avons constaté la présence des pics homogènes (entre 35 et 40) voisinant avec d'autres pics non homogènes centrés au milieu du spectre. Un tel profil a été observé précédemment pour la protéine E7 du HPV 45 et il a pu être interprété comme résultant de la présence d'un domaine C-terminal structuré liant le zinc, flanqué d'une région N-terminale non structurée (Ohlenschlager *et al.*, 2006).

Un deuxième volet dans cet article montre des résultats préliminaires sur la biologie cellulaire de ces protéines. Les vecteurs d'expression eucaryotiques contenant les gènes codant pour E6, E7 et E9 ont été transfectés dans des lignées de cellules aviaires (cellules SP6 de cailles) et de cellules humaines (HeLa et HaCat) en utilisant le réactif jetPEI. Les

cellules infectées ont été ensuite soumises après 24 heures d'incubation à des méthodes de détection de localisation *via* les signaux (GFP et myc) fusionnés aux protéines.

Les constructions GFP-E6, E6-myc et GFP-E7 sont localisés à part égale dans le cytoplasme et le noyau, tandis que la construction E7-myc se localise peut être un peu plus préférentiellement dans le noyau, avec toutefois un signal résiduel dans le cytoplasme. Remarquablement, la localisation des constructions GFP-E9 et E9-myc rappelle celle d'une protéine mitochondriale. Certains programmes bioinformatiques en ligne prédisent un signal d'import mitochondrial dans la séquence de la protéine, avec un score largement informatif, mais une confirmation expérimentale serait requise.

Au terme de ces expériences, nous avons constaté que la localisation des protéines E6, E7 et E9 ne dépend pas de la nature (myc, GFP) ni la position de la fusion. La distribution homogène des signaux suggère l'absence de formes agrégées des protéines dans les cellules, et aussi d'une façon surprenante la même localisation dans les deux souches cellulaire humaine utilisées et aviaire; ce qui laisse présupposer un rôle biologique pour ces protéines dans l'infection et le cycle viral des papillomavirus aviaires.

Signalement bibliographique ajouté par le :

UNIVERSITÉ DE STRASBOURG
Service Commun de Documentation

Identification of unusual E6 and E7 proteins within avian Papillomaviruses : cellular localization, biophysical characterization and phylogenetic analysis

Koenraad VAN DOORSLAER, Abdellahi OULD M'HAMED OULD SIDI, Katia ZENIER, Vladimir RYBIN, François DERYCKERE, Annabel RECTOR, Robert D. BURK, E. Kurt LIENAU, Marc VAN RANST and Gilles TRAVE

Journal of Virology, published online ahead of print on 24 June 2009

Manuscrit 2 : p 98-... :

La publication présentée ici dans la thèse est soumise à des droits détenus par un éditeur commercial.

Les utilisateurs de l'UdS peuvent consulter cette publication sur le site de l'éditeur :

<http://dx.doi.org/10.1128/JVI.01777-08>

La version imprimée de cette thèse peut être consultée à la bibliothèque ou dans un autre établissement via une demande de prêt entre bibliothèques (PEB) auprès de nos services :

<http://www-sicd.u-strasbg.fr/services/peb/>

Quatrième partie: l'auto-association des protéines E6

Manuscrit 3:

Natively folded Papillomavirus E6 proteins form biologically active oligomers and spontaneously assemble into multimeric ribbon-like structures

Dans cet article nous étudions l'oligomérisation des protéines E6 provenant de différentes souches de HPV (HPV 5, 11, 16, 18, 33, 45, 52, et 58) et la protéine E6 de BPV 1. Nous avons analysé les différentes protéines par le biais de plusieurs méthodes biochimiques et biophysiques incluant la gel-filtration analytique, microscopie électronique, l'ultracentrifugation analytique, Biacore et FTIR (La Spectroscopie Infrarouge à Transformée de Fourier). Afin d'augmenter la solubilité des préparations, nous avons produit les protéines E6 en fusion avec la protéine MBP. Ceci permet davantage de ralentir la formation de formes multimériques lors de l'oligomérisation spontanée de E6. Nous avons remarqué que la proportion de solubilité de chaque oligomère varie selon la souche virale. Les protéines E6 à haut risque s'oligomérisent sous formes repliées capables de lier l'E6AP et conservent la capacité de dégradation de p53. Nous montrons que la protéine E6, bien que de petite taille, elle est capable de s'assembler et s'organiser sous forme de structures semblables aux fibres. Cette conformation est accompagnée par un changement au niveau de la structure secondaire de la protéine E6. Nous avons également démontré par des expériences de co-localisation de la protéine E6 dans la cellule que l'auto-association est un événement qui se produit *in vivo*. Les résultats que nous avons obtenus suggèrent que l'auto-association de la protéine E6 semble être une activité naturelle de E6 et pourrait ensuite avoir des répercussions fonctionnelles.

Natively folded Papillomavirus E6 proteins form biologically active oligomers and spontaneously assemble into multimeric ribbon-like structures

Katia Zanier^{(1)(&)}, *Abdellahi ould M'hamed ould Sidi*⁽¹⁾, *Frederic Melin*⁽²⁾, *Christine Ruhlmann*⁽³⁾, *Tutik Ristriani* (1), *Vladimir Rybin* (4), *Mireille Baltzinger*⁽¹⁾, *Scott Vande Pol*⁽⁵⁾, *Petra Hellwig*⁽²⁾, *Patrick Schultz*⁽³⁾ and *Gilles Travé*^{(1)(&)}

⁽¹⁾ Ecole Supérieure de Biotechnologie de Strasbourg (IREBS, FRE 3211), Boulevard Sébastien Brant, BP 10413, 67412 Illkirch, France.

⁽²⁾ Institut de Chimie (UMR 7177), University of Strasbourg, 1 rue Blaise Pascal, 67000 Strasbourg, France.

⁽³⁾ Institut de Génétique et de Biologie Moléculaire et Cellulaire, CNRS/INSERM/ULP, 1 rue Laurent Fries, BP 163, 67404 Illkirch, France

⁽⁴⁾ EMBL Heidelberg, Meyerhofstrasse 1, 69117 Heidelberg, Germany

⁽⁵⁾ Department of Pathology, University of Virginia, Charlottesville, VA 22908, USA.

Running title:

Self-Association of Papillomavirus E6 proteins

(&) Corresponding authors

Phone: 0033 3 90 24 47 20, fax: 0033 3 90 24 4683

e-mails: zanier@esbs.u-strasbg.fr, trave@esbs.u-strasbg.fr

ABSTRACT

Papillomavirus (PV) E6 oncoproteins bind and often provoke the degradation of many cellular proteins important for the control of cell proliferation and/or cell death. Structural studies on E6 proteins have long been hindered by the difficulties of obtaining highly concentrated samples of recombinant E6. Here, we show that E6 proteins from eight human PV strains (HPV 5, 11, 16, 18, 33, 45, 52, and 58) and one bovine papillomavirus (BPV 1) exist as oligomeric as well as multimeric species. In this work we have analyzed E6 preparations using a variety of biochemical and biophysical techniques including analytical gel filtration, activity assays, SPR, EM, FTIR. The characterization of E6 oligomers is facilitated by the fusion to the maltose binding protein (MBP), which slows down the formation of even higher-order multimeric species. Oligomers of high-risk mucosal HPV E6 proteins appear to consist of folded E6 units, which retain binding to the ubiquitin ligase E6AP and the capacity to degrade the pro-apoptotic protein p53. The proportion and stability of each oligomeric form vary depending on the viral strain considered. In addition to the small-size oligomers, E6 proteins spontaneously assemble into large organized multimeric structures, a process which is accompanied by significant changes at the level of secondary structure. Finally, co-localisation experiments using E6 equipped with different tags indicate that E6 self-association also occurs *in vivo*. The ensemble of these data suggest that self-association is a general activity of E6 proteins which occurs both *in vitro* and *in vivo* and might therefore be functionally relevant.

Key words: HPV, E6, cervical cancer, self-association, oligomerization, ordered aggregation.

Abbreviations: MBP, maltose binding protein; HPV, human papillomavirus; BPV, bovine papillomavirus; TEV, tobacco etch virus; SPR, surface plasmon resonance; RU, resonance units; EM, electron microscopy; FTIR, Fourier Transform Infrared Spectroscopy;

INTRODUCTION

Papillomaviruses (PV) are small DNA viruses that induce squamous epithelial neoplasia. They are classified into phylogenetic groups with distinct tropisms for different body sites (skin, mouth and genitalia) and are divided into “low-risk” and “high-risk” types according to their carcinogenic potential. There are nearly 100 different Human papillomaviruses (HPV) so far identified. Cervical carcinomas are induced by high-risk genital HPVs and represent one of the main causes of cancer-related death for women (1). HPV16 is the most common oncogenic strain associated with 50% of the cervical cancers, followed by HPV18 covering 14% of the cancers (2). HPV types 5 and 8 are associated with squamous cutaneous carcinomas (3). Among the PV types infecting other animals, bovine papillomaviruses (BPV) have also been extensively studied due to their strong transforming properties (4).

In anogenital cancers caused by high-risk HPVs, two oncoproteins, E6 and E7, are expressed. E6 proteins from “high-risk” genital HPVs are thought to promote tumorigenesis by binding and often provoking the degradation of a number of proteins involved in cell proliferation and cell death control (5). In particular, E6 forms a complex with the ubiquitin ligase E6-associated protein (E6AP) (6), which in turn is able to target the tumor suppressor p53 (7), leading to ubiquitin-mediated degradation of p53. Another group of E6 targets include the tumour suppressor human discs large hDlg (8) and the MAGI family of proteins (9, 10). However E6 proteins from other HPV species may carry out their functions *via* alternative mechanisms. For instance, high-risk cutaneous HPV5 and 8 E6 proteins do not appear to degrade p53, but promote the degradation of the Bak pro-apoptotic protein, which is induced after UV irradiation of the skin (11). BPV1 E6 interacts with the focal adhesion protein paxillin and the AP-1 adaptin complex and these interactions correlate with cellular transformation (12, 13).

Papillomavirus E6 proteins are rather small in size (about 150 amino acids). Most of them share a common architecture consisting of two zinc-binding domains (E6-N and E6-C) (14). While the structure of the E6-C domain has recently been solved in our laboratory (15), structures of full-length E6 proteins have not been reported yet, due to the difficulties encountered in the preparation of highly concentrated and monodisperse samples of these proteins. Most attempts to produce the recombinant full-length protein in *E. coli* have focussed on the main oncogenic HPV16 variant and mutants of this protein (16-18). Recently, we have also explored E6 proteins issued from other HPV types, such as high-risk genital HPV18, 33, 45, 52, 58, low-risk HPV11 and high-risk cutaneous HPV5. We found that all these E6 proteins form large soluble aggregates upon fusion to the carrier protein MBP *in vivo* during biosynthesis in bacteria (16, 19). Other authors have also observed multimeric structures in HPV16 and HPV18 E6 preparations obtained using refolding techniques.

In the present work we focus on E6 self-association and show that this property is shared among E6 proteins issued from different phylogenetic viral groups. We identify and characterize both oligomeric and multimeric E6 species that exist *in vitro*. In addition, we detect E6 self-association in eukaryotic cells using nucleocytoplasmic indirect transport techniques. The conservation of self-association within the E6 family suggests that such mechanisms may have a functional role.

RESULTS

E6 proteins self-associate into oligomeric species.

We examined the oligomeric status of a panel of E6 proteins issued from seven different HPV strains (i.e. types 5, 11, 16, 18, 33, 52 and 58) and from BPV1. E6 constructs from HPV5, 11, 16, 33 and 52 are mutated on non-conserved cysteines in order to reduce intermolecular disulfide bridge formation as previously described (19). For clarity reasons, throughout the text, the terms HPV5, 11, 16, 33 and 52 E6 designate these stabilised cysteine mutants. By contrast, E6 constructs from HPV18 and BPV1 correspond to the wild-type sequences (20). MBP-fused E6 proteins were purified by amylose affinity and cleared from the high molecular weight MBP-E6 aggregates by ultracentrifugation as described in (19). Sample concentrations of all MBP-E6 fusions were adjusted to 2.5 μ M and analysed by size exclusion chromatography. The elution profiles indicate multiple oligomeric species in most of the preparations (Figure 1). For MBP-fused HPV18 E6, two peaks of similar intensity corresponding to monomeric and dimeric forms are detected, whereas for HPV11 monomeric, dimeric and trimeric forms of MBP-E6 are observed. On the other hand, the monomer is the most abundant form for MBP-fused E6 proteins from HPV5, 16, 52 as well as BPV1. For reasons of practicality, from here onwards we focus our studies on a diverse subset of E6 proteins, including most of the times HPV11, 16 and 18 E6, which belong to different phylogenetic groups of the Papillomavirus family and are amenable to purification.

To further confirm the gel filtration data we performed analytical ultracentrifugation sedimentation velocity experiments on HPV11, 16 and 18 MBP-E6 preparations. Fitting of the sedimentation profiles points to the existence of species with molecular sizes consistent with those of monomeric, dimeric and trimeric MBP-E6 (Figure 2). Although the polydispersity of the system does not allow for a precise extrapolation of molecular weights, these results support the findings of analytical size exclusion chromatography described above.

We proceeded to separate E6 from the MBP tag by TEV proteolysis and analyzed the digestion products by gel filtration chromatography. In the case of HPV11 E6 (Figure 3A) the elution profile contains three separate peaks at 9.1, 10 and 12.0 ml. SDS-PAGE analysis shows

that elution peaks at 9.1 and 12 ml contain predominantly E6, while the peak at 10 ml contains both of E6 and the MBP tag (Figure 3B). Hence E6 elutes at three distinct volumes, which are expected for trimeric, dimeric and monomeric E6 according to the column's calibration, showing that self-association is mediated by residues in the E6 protein rather than by the MBP carrier protein.

Noteworthy, all E6 proteins analysed in this study are prone to precipitation once separated from their MBP carrier. However, the concentration threshold for precipitation varies for each E6 protein. HPV18 E6 and BPV1 E6 precipitate immediately upon TEV cleavage, so that they cannot even be analysed by gel-filtration. On the other hand, monomeric and trimeric HPV11 E6 can be separated from MBP by analytical gel-filtration, but cannot be further concentrated due to adsorption onto the membranes of the concentrating devices. Indeed, the previously described HPV16 E6 6C/6S construct (16) remains the most soluble with a precipitation threshold of about 50 μM .

Stability of E6 oligomers upon dilution and concentration.

Next we investigated the influence of dilution on the stability of the E6 oligomers. MBP-E6 proteins from HPV11 and 18 were affinity-purified, cleared from high-molecular weight aggregates by ultracentrifugation, applied onto a Superdex 200 10/30 gel filtration column and fractionated to the different oligomeric species. Protein concentrations in fractions corresponding to the isolated monomeric, dimeric and trimeric species were adjusted by dilution to either 6 μM or 0.8 μM for HPV11 and to either 3 μM or 0.4 μM for HPV18 E6. After about two hours of incubation at room temperature, aliquots of these diluted samples were re-loaded individually for a second time on the same column. The resulting elution profiles are reported in Figures 4 and 5 for HPV11 and 18, respectively. Reloading of monomer fractions of both HPV11 and 18 MBP-E6 generates elution profiles with essentially only monomer peaks (Figures 4B and 5B). When reloading the oligomeric fractions instead we observe different behaviours for the HPV18 and 11 strains. HPV11 E6 dimer fractions give rise to monomers and trimer forms (Figure 4C). Most probably this is related to the resolution of the first gel filtration column where dimer populations cannot be fully separated from other forms. Instead HPV 11 E6 trimeric fractions give rise to a prominent monomer peak (Figure 4D), which points to the weak stability of HPV11 trimers dissociating into monomers. On the other hand, the HPV18 E6 dimer form strongly persists in the second gel-filtration column (Figure 5C) with only a minor monomer peak for the lowest concentration (accounting for about 10% of the total protein according to peak area integration values). This latter result highlights the higher stability of the HPV18 dimeric species compared to HPV11 oligomers. Monomer and dimer fractions of HPV18 MBP-E6 were analyzed by SDS-PAGE in the presence and in the absence of the reducing agent β -mercaptoethanol. We could not detect any band migrating at molecular

weights higher than that corresponding to an MBP-E6 monomer (i.e. 60 kDa) ruling out the possibility that E6 18 dimer stability arises from disulfide cross-bridging (Figure 5D).

We also investigated E6 oligomerization upon concentration. A preparation of affinity-purified HPV18 MBP-E6 cleared from soluble aggregates was analyzed by gel filtration before and after 5-fold concentration (Figure 6). For the concentrated sample, we observe a large increase of higher-order oligomers eluting at the column's void volume in addition to a small increase of the MBP-E6 dimer peak. This shows that MBP-E6 oligomers further assemble into larger species upon concentration.

Oligomers possess native-like E6 in vitro activities.

Next, we characterized the zinc coordination properties of the different E6 forms. Zinc concentrations were measured in gel filtration fractions of monomeric and oligomeric MBP-E6 from strains 11 and 18 using the selective fluorophore TSQ (17). For all samples we observed a nearly perfect 2:1 Zn^{2+} :E6 stoichiometry which is expected from the common architecture of E6 proteins consisting of two zinc-binding domains (Table 1). These data indicate that the E6 self-association into small oligomers is not related to loss of metal ion coordination.

Monomeric and dimeric species of the high-risk HPV18 MBP-E6 were also tested for their p53 degradation activities immediately after fractionation on the gel filtration column. Radiolabelled ^{35}S p53 was incubated with varying concentrations of monomeric and dimeric MBP-E6, corresponding to 400, 200, 100 and 50 nM for 1 hour at 29 °C. Given the observations on HPV18 E6 dimer stability reported above, it is likely that the dimer persists at least for the highest concentrations (i.e. 400 and 200 nM). The results clearly show very similar degradation profiles for monomeric and dimeric HPV18 E6 at all concentrations tested (Figure 7A).

Subsequently, we investigated the binding activity of dimeric HPV18 MBP-E6 to a 15-mer peptide containing the Lxx□Lsh motif of E6AP using Biacore technology. For such purpose we followed the protocols described in a previous report (21), where we carried out an extensive analysis of the interactions of different monomeric E6 proteins to the same peptide. Dimer fractions of HPV18 MBP-E6 were injected at concentrations ranging from 1.6 to ~0.2 μM on a surface capturing GST-E6AP peptide fusions. The resulting sensorgrams show concentration dependent binding (Figure 7B). Unfortunately, like in the case of monomeric HPV18 E6, extrapolation of precise kinetic and/or affinity parameters for this interaction is difficult. Kinetic analysis of these sensorgrams is hampered by the fast dissociation rates and the slight biphasic profiles, while equilibrium analysis requires sampling of high analyte concentrations that, in the case of E6 proteins, promote aggregation processes (see (21) for further discussion). However the magnitudes of binding responses are comparable for dimeric HPV18 MBP-E6 and the monomeric HPV18 MBP-E6 form analyzed in (21). This suggests that

the monomeric and dimeric forms have similar affinities for the cognate peptide corresponding to a K_D of approximately 30 μM as estimated in (21).

Likewise, we also tested preparations containing trimer and dimer forms of “low-risk” HPV11 MBP-E6. As in the case of monomeric HPV11 MBP-E6, we could not observe any significant activity of these samples both with respect to p53 degradation and E6AP binding (*data not shown*). Overall these observations indicate that the HPV 18 E6 dimer is an active species with respect to p53 degradation and E6AP binding. Together with the finding of a 2:1 Zn^{2+} : protein stoichiometry our results suggest that the E6-N and E6-C domains possess a native-like fold within the oligomeric E6 species.

Monomeric E6 proteins spontaneously assemble into large organized multimeric structures.

In our laboratory we have repeatedly experienced loss of activity (both E6AP peptide binding and p53 degradation activities) in samples of recombinant E6 upon storage during several weeks (*data not shown*). We thus wondered what are the physical processes leading to inactive E6 and investigated them by different biophysical methods. We initially selected HPV16 E6, which, in our hands, displays the highest concentration threshold of aggregation once separated from the MBP tag (50 μM , *data not shown*) and thus is the E6 protein most amenable to purification. Monomeric folded HPV16 E6 was isolated by gel-filtration, adjusted to a concentration of 40 μM and incubated at 20 °C for one week. During this period, freshly prepared DTT was added to the sample every 24 hours to maintain reducing conditions and avoid disulfide cross-bridging. Aliquots were taken at time intervals of 0, 1/2, 3 and 7 days, deposited on a carbon coated grid and stained using a uranyl acetate solution for direct EM observation. Figure 8A shows partial field images at the different time points. Immediately after purification ($t = 0$) we can detect mostly low molecular weight particles, likely to correspond to the monomeric protein, and small spherical deposits with an approximate diameter of 4 nm. After 12 hours ($t = 12$ hours) larger structures resembling a collection of spherical particles are visible in some parts of the grid. At day 3 ($t = 3$ days) we can instead observe well-defined ribbon-like structures with widths varying between 20 and 30 nm and measuring several hundreds of nanometers in length. After one week ($t = 1$ week) bundles of fibrous filaments are visible with an overall less regular morphology than the structures observed at day 3. Formation of the structures similar to those reported in Figure 8A was also detected in preparations of monomeric E6 incubated at either low temperature (5 °C) or at lower salt concentration (150 mM NaCl instead of 400 mM) (*data not shown*). Similar results were also obtained with monomeric MBP-E6 fusions of HPV16 E6, indicating that the presence of the MBP tag does not hinder the fiber assembly process. This latter finding allowed us to investigate other E6 proteins, which require fusion to MBP for the isolation of their monomeric form. We purified

monomeric MBP-E6 proteins issued from HPV11, HPV18 and the phylogenetically more distant BPV1 virus and incubated them as indicated above. EM images taken immediately after purification show that all MBP-E6 proteins are homogeneous preparations of low molecular weight particles (Fig. S1A-C). As the incubation time increases we detect formation of similar structures as those formed by HPV16 E6. Images in Fig. 8B-8D show that all MBP-E6 proteins analyzed develop fiber bundles and ribbon-like structures after one week of incubation. Hence, the ensemble of these data show that spontaneous self-association into large multimeric structures is a property shared by distantly related E6 proteins.

At this stage we wondered what would be the evolution of the soluble MBP-E6 aggregates described in (19) upon incubation. HPV16 MBP-E6 samples were affinity purified as described in (19). Immediately after purification these preparations contain the well-defined spherical structures previously described. After few days of incubation they also develop fiber bundles although with an overall less regular morphology as the structures shown in Fig. 8 (Fig. S1D). Remarkably, when incubation is performed at low temperature (5 °C) to slow down the kinetics of self-assembly, we clearly observe small E6 filaments growing from the spherical aggregates (indicated by an arrow in Fig. 8E), suggesting that soluble MBP-E6 spherical structures can also act as nuclei for fiber formation.

Secondary structure changes implicated in E6 self-assembly.

In an attempt to further characterize the structure of E6 multimers, we monitored the changes in Thioflavin-T fluorescence in preparations of monomeric HPV16 E6. We detected no significant differences in the intensity of the Thioflavin-T emission signal at the different time points of E6 incubation (*data not shown*), suggesting that E6 multimers do not contain the cross-beta structures typical of amyloid aggregation (22). In addition, far UV Circular Dichroism (CD) experiments failed to provide conclusive results, mainly due to the difficulties imposed by sample precipitation upon incubation and light scattering during the measurements. We therefore opted for FTIR spectroscopy, which enables to overcome these problems. Experiments were performed on 40 μ M monomeric HPV16 E6 preparations incubated at 20 °C in buffer containing either 150 or 400 mM NaCl (see Materials and Methods). E6 samples did not show any evolution in the amide I region (spectral range 1700/1600 cm^{-1}) during the first week of measurement. However a significant change in the bandwidth of the amide I band is observed starting from the second week of incubation which terminates after one month (Fig. 9A and B), Similar changes were also observed for HPV16 E6 samples which have been dialyzed against buffer containing 98% D2O immediately after purification (Figure S2A). Overall it appears that salt concentration has no significant effect on the time-evolution of E6 samples.

Secondary structure analysis was carried out on the Amide I band of spectra of the E6 protein in 150 mM NaCl (Table 2 and Fig. S2). For monomeric E6 at t=0 we obtained the following estimation: 30% β -sheet, 28% α -helix, 24% β -turn and 13% random coil. These results are in agreement with what has been previously reported by Lipari and coworkers on the HPV16 E6(Δ 143-151) construct which lacks the last 8 unfolded C-terminal residues (27% β -sheet, 26% α -helix, 24% β -turn) (23).

On the other hand, for the E6 sample incubated for 30 days we obtain a significantly different profile with 38% β -sheet, 17% α -helix, 27% β -turn and 10% random coil. It therefore seems that there is a significant increase in β -sheet structure (8%), which is paralleled by a decrease in α -helix.

E6 self-association occurs in vivo in transiently expressed eukaryotic cells.

Nucleo-cytoplasmic delocalization techniques allow detection of protein-protein interactions in transiently co-transfected eukaryotic cells (24). This assay requires one of the proteins to be fused to either a nuclear localisation or a nuclear export sequence (NLS or NES). As a result of protein-protein binding the interacting partner becomes targeted to either the nucleus or the cytoplasm.

Here we used nucleo-cytoplasmic delocalization to investigate E6 self-association *in vivo*. We selected HPV18 E6 based on its ability to form stable dimers *in vitro*. We cloned fusion constructs containing the wild-type HPV 18 E6 sequence N-terminally fused to either a Flag or a cMyc tag to allow for immunodetection. The cMyc-E6 constructs were additionally appended with either the SV40 T-antigen based-NLS sequence or a leucine-rich NES sequence at the N-terminus (NLS-cMyc-E6 and NES-cMyc-E6).

When transfected in HeLa cells, Flag-E6 localises preferentially in the nucleus due to the fact that wild-type E6 is naturally equipped with intrinsic NLS sequences (Fig. 10A). When an excess of NLS-Myc-E6 is co-transfected with Flag-E6, the NLS-Myc-E6 exhibits a strong nuclear signal. Remarkably, co-transfected Flag-E6 is prevalently nuclear with cytoplasmic expression levels close to the background (Fig. 10B). On the other hand, when an excess of NES-cMyc-E6 is co-transfected with Flag-E6, NES-Myc-E6 displays an even distribution between nucleus and cytoplasm. Indeed, a similarly even distribution is observed also for the co-transfected Flag-E6, which contrasts with the preferential nuclear staining observed for Flag-E6 alone (Fig. 10C). Therefore the distribution of Flag-E6 is strongly modified by co-transfection with an excess of cMyc-E6 equipped with localisation signals. These data indicate that differentially tagged E6 monomers interact with each other inside the cells, suggesting that E6 proteins self-associate in their natural *in vivo* environment.

DISCUSSION

E6 proteins share a common domain architecture but have variable surface properties (15). We have previously described soluble aggregates possessing miscellar structure that are formed upon biosynthesis by PV E6 proteins upon fusion to MBP (16, 19). Besides us, other authors reported on the existence of oligomeric E6 forms of HPV16 and 18 E6 (17, 18, 25). In this work we find that distantly related E6 proteins such as HPV16 E6 and BPV1 E6 form both soluble oligomers and multimeric ordered ribbon-like structures *in vitro*, which clearly highlights a conservation of self-association in the E6 protein family.

We found that E6 oligomers have *in vitro* properties very similar to the monomeric counterparts. E6 monomers, dimers and trimers all possess a native 2:1 Zn²⁺:protein stoichiometry expected from the E6 domain architecture, indicating that E6 oligomerization is not caused by loss of metal ion coordination. In addition, we found no detectable difference in the p53 degradation and E6AP binding activities of monomeric and dimeric HPV18 E6. We also observed that the MBP-E6 oligomers can re-dissociate, though with different rates depending on the E6 type, into the monomeric form upon dilution. The ensemble of these observations suggest that the small-size oligomers are formed by natively folded monomers engaging in reversible non-covalent interactions which do not alter the structural features required for their activities.

In this study we also observe that monomeric E6 proteins from different PV strains purified in native conditions can spontaneously assemble into large organized ribbon structures upon incubation at 20 °C for several days. FTIR experiments have shown that E6 self-association into these large multimers is accompanied by secondary structure changes involving a significant increase in the β -sheet content. However, the absence of thioflavin-T binding, which is in agreement with the findings of (25) on spheroid multimers of HPV18 and HPV16 E6 refolded in acidic conditions, suggests that E6 multimers do not possess amyloid structure. At present we do not know whether the conformational change related to the increase in β -sheet takes place at the level of E6 oligomerization or upon formation of larger multimers. Unfortunately we cannot measure FTIR spectra on the different oligomers of HPV11 and 18 E6, since these species can only be purified as MBP-E6 fusions and cannot be separated from the MBP tag that would dominate the spectra.

Whatever the structure of the E6 oligomers, these species might still act as intermediates in the pathways leading to larger multimers, as suggested by the observation, reported in this work, that concentration of E6 oligomers leads to further assembly into larger species. Remarkably, we have found evidence that MBP-E6 soluble aggregates purified from bacterial extracts can also serve as nuclei for the formation of small protofilaments. In a previous publication (19) we have proposed that E6 moieties in MBP-E6 soluble aggregates possess a

native-like structure based on the observation of an expected 2:1 zinc:protein stoichiometry. It is therefore possible that MBP-E6 soluble aggregates and the small oligomers described here share the same zinc-binding domain structure and pattern of intermolecular interactions. However, at present we cannot exclude that the mechanisms of E6 assembly into larger multimeric structures are distinct from those leading to small size oligomers, involving, for example, covalent interactions such as disulfide bridging and/or inter-molecular cross-linking *via* zinc ions. Interestingly, we also observed fibrous structures in preparations of the isolated HPV16 E6-N and E6-C domains (*data not shown*), albeit at higher protein concentrations (300 μ M), suggesting that the sequence and/or structure requirements for self-association are located on both of the E6 domains.

In this work we show that E6 self-association takes place in transiently expressed HeLa cells in its natural *in vivo* environment. The assay used relies on the indirect nucleocytoplasmic delocalization of E6 as a result of association with another E6 molecule equipped with either a nuclear localization or export signal peptide. Work by Garcia-Alai et al., which uses antibodies raised against soluble E6 oligomers obtained by refolding techniques, suggests that E6 oligomers occur *in vivo* in HPV-positive HeLa and CaSki cell lines (25). Hence, what could be the functional role of E6 self-association? Here we show that p53 degradation and E6AP binding activity profiles are similar for monomeric and oligomeric E6 forms. On the other hand we have repeatedly observed a loss of both E6AP peptide binding and p53 degradation activity in recombinant E6 samples that were stored for several weeks (*data not shown*), which is likely to be related to the formation of larger E6 multimers. It would be interesting to understand what are the physical events leading to activity loss, whether it is simple steric hindrance in the larger complexes or real conformational changes which occur at more advanced stages of self-association. It is possible that binding to certain ligands may promote E6 self-association. Results by Medcaf and coworkers have shown that complexes of HPV16 E6 and p53 consist of p53 tetramers and E6 dimers (26). One intriguing hypothesis for the functional role of active E6 oligomers could be the simultaneous targeting of multiple cellular proteins, resulting in the coordination of various virus-host interaction pathways. For example, it is well known that E6 binds to several host proteins and targets them to degradation (27, 28). The best characterized E6-mediated degradation pathway is the one occurring *via* the 26S proteasome, which requires ubiquitination of the substrate by the E6AP ligase. A single oligomer could bind at the same time to different cellular proteins which target either the same or different E6 sites (such as cellular proteins containing the Lxx□Lsh motif (29) or PDZ domains (30)). Consequently a single ubiquitination event on any of the substrates would be sufficient to target all these proteins simultaneously to the 26S proteasome. Such mechanism would increase the efficiency of the E6 proteins and therefore enhance their oncogenic potential.

ACKNOWLEDGMENTS

We thank Prof. G. Orth (Institut Pasteur, Paris) for providing the plasmids containing the HPV genomes. K.Z. was supported by a fellowship from the Fondation pour la Recherche Medicale (FRM). This work was funded by the following organizations: Arbor Vita Corporation (Sunnyvale, CA), Association pour la Recherche sur le Cancer (ARC), Institut National de la Santé et de la Recherche Médicale, Centre National pour la Recherche Scientifique (CNRS) and Université Louis Pasteur (Strasbourg).

AUTHOR CONTRIBUTIONS

Conceived and designed the experiments: KZ, FM, GT. Performed the experiments: KZ, AMS, FM, CR, VR, MB. Contributed reagents/materials: SVP. Analyzed the data: KZ, GT, FM, PH, PS. Wrote the paper: KZ, GT.

MATERIALS AND METHODS

Expression and purification of HPV E6 proteins.

E6 constructs from HPV 5, 11, 16, 33 and 52 are mutated in non-conserved cysteines in order to limit intermolecular disulfide bridge formation as previously described (14, 19, 31). Throughout the text, for reasons of clarity, these cysteine mutated constructs will be referred to by their strain names, omitting their full nomenclature designating the mutations. By contrast, E6 constructs from HPV 18 and BPV1 (20) correspond to the wild-type sequences.

The expression vectors for the HPV E6 proteins have been previously described (19). Similarly to the other E6 constructs, the wild type BPV1 E6 protein were cloned in the NcoI and KpnI sites of the pETM-41 vector (kindly provided by G. Stier, EMBL, Heidelberg, Germany).

MBP-E6 and E6 proteins used in this work were over-expressed in *E. Coli* BL21 DE3 by low temperature induction at 15°C for 17 hours. MBP-E6 fusion proteins were purified using amylose affinity chromatography and separated from soluble aggregates by extensive ultracentrifugation as described in (19). Buffer A (50 mM Tris pH 6.8, 400 mM NaCl and 2 mM DTT) was used throughout the purification. Preparations of monomeric MPB-E6 and E6 were obtained by further fractionation on a Superdex 75 16/60 gel filtration column equilibrated in buffer A. The individual oligomeric forms of MBP-E6 were instead obtained from fractionation on a Superdex 200 16/60 column equilibrated in buffer A as described in the text.

Concentrations were determined by absorbance at 280 nm based on the calculated extinction coefficient of the monomeric form of the protein.

Analytical gel filtration chromatography. Sample volumes of 100 or 500 μl were injected on Superdex200 10/30 and/or Superdex75 10/30 columns equilibrated in buffer A. The experiments were performed using an AKTA Purifier System (GE Healthcare) maintained at 10 $^{\circ}\text{C}$.

Analytical ultracentrifugation. Sedimentation velocity experiments were done at 4 $^{\circ}\text{C}$ using 2-channel charcoal centerpieces and speed 46 000 rpm in a Beckman Optima XL-A centrifuge fitted with a four-hole AN-60 rotor. Sedimentation velocity profiles were collected by monitoring the absorbance signal at 280 nm in buffer A. Sedimentation coefficient and molecular weight distributions were analysed by the C(s) method implemented in the Sedfit software package (32). Buffer density and viscosity corrections were made according to published data (33).

Determination of Zn^{2+} content (TSQ assay). Zn^{2+} concentrations in E6 samples were determined monitoring changes in fluorescence of the zinc selective fluorophore TSQ ((*N*-6-methoxy-8-quinoly)-*p*-toluenesulfonamide) (Molecular Probes, Eugene, OR) according to published protocols (17) with minor modifications described in detail in (19).

p53 degradation assays. p53 degradation assays were carried out as described in (34). Briefly, p53 proteins were synthesized *in vitro* using the TNT T7/SP6 reticulocyte lysate transcription/translation system (Promega). The p53 translation product was incubated with MBP-E6 or E6 proteins at concentrations between 400 and 25 nM in 25 mM Tris-HCl pH 7.5, 100 mM NaCl and 2 mM DTT, at 29 $^{\circ}\text{C}$ for 1 hour. Control reactions were made by replacing the MBP-E6 proteins with equivalent volumes of Buffer A. Reactions were stopped by addition of SDS-loading buffer, subjected to SDS-PAGE and analysed by autoradiography.

BIAcore experiments. Experiments were performed at 25 $^{\circ}\text{C}$ on a Biacore 2000 instrument (Biacore). All experiments were carried out using the protocols described in (21). In brief, the minimal E6 binding site of E6AP (i.e. ESSELTQELLGEER) was expressed fused to GST and captured onto the Biacore's surface *via* a polyclonal anti-GST antibody. Subsequently, the analyte E6 protein was injected and the interaction with the E6AP peptide monitored. A negative control GST-fusion peptide (peptide sequence: GNSGNSGNS) was used to correct the binding sensorgrams for buffer effects and non-specific binding.

Electron microscopy. Aliquots from MBP-E6 and E6 preparations were diluted to a final concentration of 30 ng/ μ l in buffer A. 5 μ l were deposited on an air glow-discharged carbon-coated grid. After 2 min adsorption samples were negatively stained with a 2% (w/v) uranyl acetate solution. Images were acquired on a Philips CM120 transmission electron microscope operating at 100 kV with a LaB6 filament. Areas covered with individual molecules were recorded under low dose condition, at a magnification of x 45.000 on a Pelletier cooled CCD camera (Model 794, Gatan, Pleasanton, CA).

FT-IR. 40 μ M preparations of monomeric HPV16 E6 in 20 mM phosphate buffer (pH 6.8), 400 NaCl (unless otherwise stated), 2 mM DTT were incubated at 20 °C for a period of up to 30 days. Middle infrared measurements were carried out at room temperature on a Vertex 70 (Bruker) Fourier transform spectrometer equipped with a globar source, a LN-MCT Photovoltaic detector, a KBr beam splitter and an ATR cell (Harrick) fitted with a diamond crystal (6 mm²) as the reflecting element. The spectrometer compartments were purged with dry air to reduce interference with H₂O and CO₂. At the different time points, 2-3 μ l of sample solution were deposited on the ATR crystal and allowed to dry at room temperature for 10 minutes. For each spectrum, a resolution of 2 cm⁻¹ was chosen and 256 scans were accumulated. Residual water in the gas phase was subtracted and all spectra were smoothed with 9 points. The data are presented in terms of arbitrary units. We noticed that sample concentration decreases during incubation due to adherence to the eppendorf's walls. However the relative ratio of each structural element is unaffected and secondary structure analysis was performed on the Amide I band (spectral range 1700/1600 cm⁻¹) using PeakFit software for deconvolution. The fitting procedure performed has an error of some percent.

Cell transfections and immunofluorescence assays. HeLa cells were maintained in Dulbecco's minimal essential medium (pH 7.2), supplemented with 10% foetal calf serum. Exponentially growing cells were transfected by calcium phosphate co-precipitation (35) with the different plasmids. The following transfections were performed : i) 0.8 μ g of Flag-18 E6; ii) 3 μ g of Flag-18 E6; iii) 3 μ g of NLS-Myc-18 E6 and 0.8 μ g of Flag-18 E6; vi) 3 μ g of NES-Myc-18 E6 and 0.8 μ g of Flag-18 E6. Each transfection was completed to 6.3 μ g of final quantity of plasmids using plasmid BlueSkript KS-. Transfected cells were let grow for 24 hours. Hence cells were washed with PBS, fixed in cooled methanol for 10 minutes at -20°C and then in cooled acetone for 1 minute at -20°C. Upon fixation, the cells were incubated with the anti-FLAG rabbit polyclonal antibody F7425 (Sigma) diluted in PBS for 1 hour at room temperature, then with the anti-Myc mouse monoclonal antibody 9E10 (kindly provided by Dr. M. Oulad, IGBMC, Strasbourg) subsequently with Alexa Fluor 568-conjugated anti-rabbit goat

immunoglobulin and Alexa Fluor 488-conjugated anti-mouse goat immunoglobulin (Invitrogen) and finally with DAPI (Sigma). After extensive washing with PBS, cells were dried and mounted with Fluoromount (Southern Biotech) on a slide. The cells were then examined with a Zeiss Axioplan Fluorescence microscope equipped with an Olympus DP50 camera. Images were collected with Zeiss 63X plan-neofluar objectives and processed using Adobe-Photoshop 5.5.

<i>Sample</i>	<i>Zn²⁺: protein stoichiometry</i>
HPV 11 (MBP-E6) ₁	2.30 ± 0.33
HPV 11 (MBP-E6) ₂	1.90 ± 0.03
HPV 11 (MBP-E6) ₃	2.10 ± 0.11
HPV 16 (MBP-E6) ₁	1.97 ± 0.24
HPV 18 (MBP-E6) ₁	1.70 ± 0.09
HPV 18 (MBP-E6) ₂	2.00 ± 0.04

Table 1.

Experimentally derived Zn²⁺:protein stoichiometries for MBP-E6 samples from HPV strains 11, 16 and 18. Monomer, dimer and trimer species are indicated as (MBP-E6)₁, (MBP-E6)₂ and (MBP-E6)₃ respectively. Zinc concentrations were measured by the TSQ assay and calculated relative to ZnSO₄ standards.

	<i>HPV 16 E6 (t=0)</i>		<i>HPV 16 E6 (t=30 days)</i>	
	<i>Position (cm⁻¹)</i>	<i>% ss</i>	<i>Position (cm⁻¹)</i>	<i>% ss</i>
β-sheet	1623, 1631, 1639, 1690	30	1625, 1637, 1685, 1695	38
α-helix	1657	28	1653	17
β-turn	1669, 1679	24	1663, 1674	27
random coil	1647	13	1654	10

Table 2. Secondary structure content (expressed as % ss) of monomeric HPV16 E6 immediately after purification (t=0) and after 30 days of incubation (t = 30 days). Secondary structure analysis was carried out by deconvolution of spectra of E6 in 20 mM sodium phosphate (pH 6.8), 150 mM NaCl, 2 mM DTT (Figure 9A, 9C). The frequency assignments are based on (36).

REFERENCES

- (1) zur Hausen, H. (1999) Papillomaviruses in human cancers. *Proc Assoc Am Physicians 111*, 581-587.
- (2) Bosch, F. X., Manos, M. M., Munoz, N., Sherman, M., Jansen, A. M., Peto, J., Schiffman, M. H., Moreno, V., Kurman, R., and Shah, K. V. (1995) Prevalence of human papillomavirus in cervical cancer: a worldwide perspective. International biological study on cervical cancer (IBSCC) Study Group. *J Natl Cancer Inst 87*, 796-802.
- (3) Pfister, H. (2003) Chapter 8: Human papillomavirus and skin cancer. *J Natl Cancer Inst Monogr. 31*, 52-56.
- (4) Lambert, P. F., Baker, C. C., and Howley, P. M. (1988) The genetics of bovine papillomavirus type 1. *Annu Rev Genet 22*, 235-58.
- (5) Narisawa-Saito, M., and Kiyono, T. (2007) Basic mechanisms of high-risk human papillomavirus-induced carcinogenesis: roles of E6 and E7 proteins. *Cancer Sci 98*, 1505-11.
- (6) Scheffner, M., Huibregtse, J. M., Vierstra, R. D., and Howley, P. M. (1993) The HPV-16 E6 and E6-AP complex functions as a ubiquitin-protein ligase in the ubiquitination of p53. *Cell 75*, 495-505.
- (7) Werness, B. A., Levine, A. J., and Howley, P. M. (1990) Association of human papillomavirus types 16 and 18 E6 proteins with p53. *Science 248*, 76-79.
- (8) Kiyono, T., Hiraiwa, A., Fujita, M., Hayashi, Y., Akiyama, T., and Ishibashi, M. (1997) Binding of high-risk human papillomavirus E6 oncoproteins to the human homologue of the Drosophila discs large tumor suppressor protein. *Proc Natl Acad Sci U S A 94*, 11612-11616.
- (9) Glaunsinger, B., Lee, S., Thomas, M., Banks, L., and Javier, R. (2000) Interactions of the PDZ-protein MAGI-1 with adenovirus E4-ORF1 and high-risk papillomavirus E6 oncoproteins. *Oncogene 19*, 5270-5280.
- (10) Thomas, M., Laura, R., Hepner, K., Guccione, E., Sawyers, C., Lasky, L., and Banks, L. (2002) Oncogenic human papillomavirus E6 proteins target the MAGI-2 and MAGI-3 proteins for degradation. *Oncogene 21*, 5088-5096.
- (11) Jackson, S., Harwood, C., Thomas, M., Banks, L., and Storey, A. (2000) Role of Bak in UV-induced apoptosis in skin cancer and abrogation by HPV E6 proteins. *Genes Dev 14*, 3065-73.
- (12) Tong, X., and Howley, P. M. (1997) The bovine papillomavirus E6 oncoprotein interacts with paxillin and disrupts the actin cytoskeleton. *Proc Natl Acad Sci U S A. 94*, 4412-4417.
- (13) Tong, X., Boll, W., Kirchhausen, T., and Howley, P. M. (1998) Interaction of the bovine papillomavirus E6 protein with the clathrin adaptor complex AP-1. *J Virol 72*, 476-482.
- (14) Nominé, Y., Charbonnier, S., Ristriani, T., Stier, G., Masson, M., Cavusoglu, N., Van Dorsselaer, A., Weiss, E., Kieffer, B., and Travé, G. (2003) Domain substructure of HPV E6 protein: biophysical characterization of E6 C-terminal DNA-binding domain. *Biochemistry 42*, 4909-4917.
- (15) Nomine, Y., Masson, M., Charbonnier, S., Zanier, K., Ristriani, T., Deryckere, F., Sibler, A. P., Desplancq, D., Atkinson, R. A., Weiss, E., Orfanoudakis, G., Kieffer, B., and Trave, G. (2006) Structural and functional analysis of E6 oncoprotein: insights in the molecular pathways of human papillomavirus-mediated pathogenesis. *Mol Cell 21*, 665-78.
- (16) Nominé, Y., Ristriani, T., Laurent, C., Lefèvre, J. F., Weiss, E., and Travé, G. (2001) Formation of soluble inclusion bodies by HPV E6 oncoprotein fused to Maltose-binding protein. *Protein Expr Purif. 23*, 22-32.
- (17) Degenkolbe, R., Gilligan, P., Gupta, S., and Bernard, H. U. (2003) Chelating agents stabilize the monomeric state of the zinc binding human papillomavirus 16 E6

- oncoprotein. *Biochemistry* 42, 3868-3873.
- (18) Daniels, P. R., Sanders, C. M., Coulson, P., and Maitland, N. J. (1997) Molecular analysis of the interaction between HPV type 16 E6 and human E6-associated protein. *FEBS Letters* 416, 6-10.
 - (19) Zanier, K., Nomine, Y., Charbonnier, S., Ruhlmann, C., Schultz, P., Schweizer, J., and Trave, G. (2007) 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* 51, 59-70.
 - (20) Bohl, J., Das, K., Dasgupta, B., and Vande Pol, S. B. (2000) Competitive binding to a charged leucine motif represses transformation by a papillomavirus E6 oncoprotein. *Virology* 271, 163-170.
 - (21) Zanier, K., Charbonnier, S., Baltzinger, M., Nomine, Y., Altschuh, D., and Trave, G. (2005) Kinetic analysis of the interactions of Human Papillomavirus E6 oncoproteins with the ubiquitin ligase E6AP using Surface Plasmon Resonance. *J. Mol. Biol.* 349, 401-412.
 - (22) Nilsson, M. R. (2004) Techniques to study amyloid fibril formation in vitro. *Methods* 34, 151-60.
 - (23) Lipari, F., McGibbon, G. A., Wardrop, E., and Cordingley, M. G. (2001) Purification and biophysical characterization of a minimal functional domain and of an N-terminal Zn²⁺-binding fragment from the human papillomavirus type 16 E6 protein. *Biochemistry* 40, 1196-204.
 - (24) Sibler, A. P., Nordhammer, A., Masson, M., Martineau, P., Trave, G., and Weiss, E. (2003) Nucleocytoplasmic shuttling of antigen in mammalian cells conferred by a soluble versus insoluble single-chain antibody fragment equipped with import/export signals. *Exp Cell Res* 286, 276-87.
 - (25) Garcia-Alai, M. M., Dantur, K. I., Smal, C., Pietrasanta, L., and de Prat-Gay, G. (2007) High-risk HPV E6 oncoproteins assemble into large oligomers that allow localization of endogenous species in prototypic HPV-transformed cell lines. *Biochemistry* 46, 341-9.
 - (26) Medcalf, E. A., and Milner, J. (1993) Targeting and degradation of p53 by E6 of human papillomavirus type 16 is preferential for the 1620+ p53 conformation. *Oncogene* 8, 2847-51.
 - (27) Liu, Y., and Baleja, J. D. (2008) Structure and function of the papillomavirus E6 protein and its interacting proteins. *Front Biosci* 13, 121-34.
 - (28) Tungteakkhun, S. S., and Duerksen-Hughes, P. J. (2008) Cellular binding partners of the human papillomavirus E6 protein. *Arch Virol* 153, 397-408.
 - (29) Chen, J. J., Hong, Y., Rustamzadeh, E., Baleja, J. D., and Androphy, E. J. (1998) Identification of an alpha helical motif sufficient for association with papillomavirus E6. *J Biol Chem* 273, 13537-13544.
 - (30) Thomas, M., Dasgupta, J., Zhang, Y., Chen, X., and Banks, L. (2008) Analysis of specificity determinants in the interactions of different HPV E6 proteins with their PDZ domain-containing substrates. *Virology* 376, 371-8.
 - (31) Nominé, Y., Ristriani, T., Laurent, C., Lefèvre, J. F., Weiss, E., and Travé, G. (2001) A strategy for optimizing the monodispersity of fusion proteins: application to purification of recombinant HPV E6 oncoprotein. *Protein Eng.* 14, 297-305.
 - (32) Schuck, P. (2000) Size-distribution analysis of macromolecules by sedimentation velocity ultracentrifugation and lamm equation modeling. *Biophys J* 78, 1606-19.
 - (33) Laue, T., Shah, B., Ridgeway, T., and Pelletier, S. (1992) Analytical Ultracentrifugation in Biochemistry and Polymer Science. *The Royal Society of Chemistry, Cambridge*, 90-125.
 - (34) Ristriani, T., Nomine, Y., C, L., Weiss, E., and Trave, G. (2002) Protein mutagenesis with monodispersity-based quality probing: selective inactivation of p53 degradation and DNA-binding properties of HPV E6 oncoprotein. *Protein Expr Purif.* 26, 357-367.
 - (35) Chen, C., and Okayama, H. (1987) High-efficiency transformation of mammalian cells by plasmid DNA. *Mol Cell Biol* 7, 2745-52.
 - (36) Barth, A. (2007) Infrared spectroscopy of proteins. *Biochim Biophys Acta* 176, 1073-101.

FIGURE LEGENDS

Figure 1. Analytical gel filtration chromatography of MBP-E6 preparations. Affinity-purified E6 samples were cleared from soluble aggregates by ultracentrifugation. Concentrations were adjusted to approximately 2.5 μ M and 100 μ l volumes injected on a Superdex200 10/30 column. Elution peaks corresponding to monomer (MBP-E6₁, 60 kDa), dimer (MBP-E6₂, 120 kDa) and trimer (MBP-E6₃, 180 kDa) forms of MBP-E6 proteins are indicated. The shoulder at 15 ml corresponds to monomeric MBP arising from residual proteolytic activity in the protein preparations. Molecular size markers are reported on top of the figure. 1: V₀; 2: ferritin (440 KDa); 3: mouse IgG (150 KDa); 4: BSA (67 KDa); 5: ovalbumin (43 KDa); 6: RNase (13.7 KDa). The elution profiles correspond to E6 proteins from the following viral stains. **(A)** HPV11 (magenta line), HPV16 (black line), HPV18 (blue line) and BPV1 (green line). **(B)** HPV33 (magenta line), HPV52 (black line), HPV58 (green line) and HPV5 (blue line).

Figure 2. Molecular weight distribution of MBP-E6 oligomeric species as derived from sedimentation velocity experiments. The E6 proteins reported are from HPV strains 11 (magenta line), 16 (black line) and 18 (blue line). C(M) indicates arbitrary units.

Figure 3. Gel filtration profiles of HPV11 MBP-E6 preparations after separation from the MBP tag. Affinity-purified MBP-E6 preparations were cleared from soluble aggregates by ultracentrifugation and digested using TEV protease. Subsequently, 500 μ l were injected on a Superdex75 10/30 column. **(A)** Peaks corresponding to monomeric (E6₁, 18 kDa), dimeric (E6₂, 36 kDa) and trimeric (E6₃, 54 kDa) E6 forms and MBP (MBP, 44 kDa) are indicated. Molecular size markers are reported on top of the figure. 1: V₀; 2: BSA (67 KDa); 3: ovalbumin (43 KDa); 4: Myoglobin (17.6 KDa); 5: RNase (13.7 KDa). Numbers 1-11 reported on the x-axis indicate fractions collected for SDS-PAGE analysis. **(B)** Lanes 1-11 correspond to fractions 1-11 indicated on the elution profile. Fractions were precipitated with 20% trichloroacetic acid. Pellets were re-suspended in 50 μ l of loading buffer and applied to a 12% SDS-polyacrylamide gel. Bands corresponding to MBP, TEV and E6 are indicated.

Figure 4 Stability of HPV11 MBP-E6 oligomers upon dilution. **(A)** Affinity purified HPV11 MBP-E6 was cleared from soluble aggregates and applied onto a Superdex200 10/30 column. **(B-D)** Aliquots of fractions corresponding to the monomeric **(B)**, dimeric **(C)** and trimeric forms **(D)** of MBP-E6 (corresponding to the indicated grey areas in **(A)**) were re-loaded individually onto the same gel filtration column at concentrations of 6.0 (continuous lines) and 0.8 (broken lines) μ M. See also legend of **Figure 1**.

Figure 5. Stability of HPV18 MBP-E6 oligomers upon dilution. **(A)** HPV18 MBP-E6 preparations were loaded on a Superdex200 10/30 column as described in **Figure 4**. Monomer **(B)** and dimer **(C)** fractions (indicated by the grey areas) were re-loaded a second time on the same column at concentrations of 3.0 (continuous lines) and 0.4 μ M (broken lines). (MBP-E6)_n refer to larger oligomers eluting at the column's void volume. See also legend of **Figure 1**. **(D)** Monomer and dimer fractions of HPV18 MBP-E6 analyzed on a 12% SDS-PAGE in the presence (+ β me) and in the absence (- β me) of the reducing agent β -mercaptoethanol. Lane 1 (L1): monomer + β me; L2: dimer + β me; L3: monomer - β me; L4: dimer - β me. Samples were migrated on the same gel but were separated by two empty wells in order to avoid cross-contamination by β -mercaptoethanol. The migration of the molecular weight markers (in kDa) is reported on the left hand side.

Figure 6 Assembly of larger oligomers of MBP-E6 upon concentration. Affinity purified HPV18 MBP-E6 was ultracentrifugated to remove soluble aggregates. A fraction of the recovered supernatant was concentrated 5-fold. Both unconcentrated (continuous line) and 5-fold concentrated (broken line) MBP-E6 samples were loaded on a Superdex 200 10/30 gel-filtration column.

Figure 7 Characterization of the activities of E6 oligomeric forms. Oligomeric MBP-E6 forms were fractionated on a Superdex 200 10/60 column. Protein concentrations were determined based on the calculated extinction coefficients of the monomeric forms. **(A)** Degradation of *in vitro* translated p53 by HPV18 MBP-E6 monomer (lanes (L): 1-4) and dimer (L: 5-8) species. p53 degradation reactions contained the following concentrations of HPV18 MBP-E6: 400 nM (L1 and L5); 200 nM (L2 and L6); 100 nM (L3 and L7); 50 nM (L4 and L8). The histogram indicates the residual p53 signal expressed as percentage relative to the intensity of the input control band (L9). **(B)** Binding of HPV18 MBP-E6 dimer to a peptide containing the Lxx□Lsh motif of E6AP as monitored by SPR. The figure illustrates the concentration dependent binding of MBP-E6 to approximately 600 RU of affinity-captured GST-E6AP peptide ligand. The MBP-E6 analyte concentrations for each sensorgram are indicated on the figure.

Figure 8 Spontaneous assembly of monomeric E6 and MBP-E6 proteins into ordered multimeric structures visualized by EM. **(A)** Partial field images of monomeric HPV16 E6 taken at 0, 1/2, 3 and 7 days of incubation Partial field images of **(B)** HPV11, **(C)** HPV18 and **(D)** BPV-1 MBP-E6 proteins after one week incubation. **(E)** HPV16 MBP-E6 soluble aggregates

incubated for 2 days at 5 °C. Bar = 50 nm for images in (A) t=0, t=12 hours, t=3 days and in (E). Bar = 100 nm for images in (A) t=1 week, (B) and (D). Bar = 500 nm for image in (C).

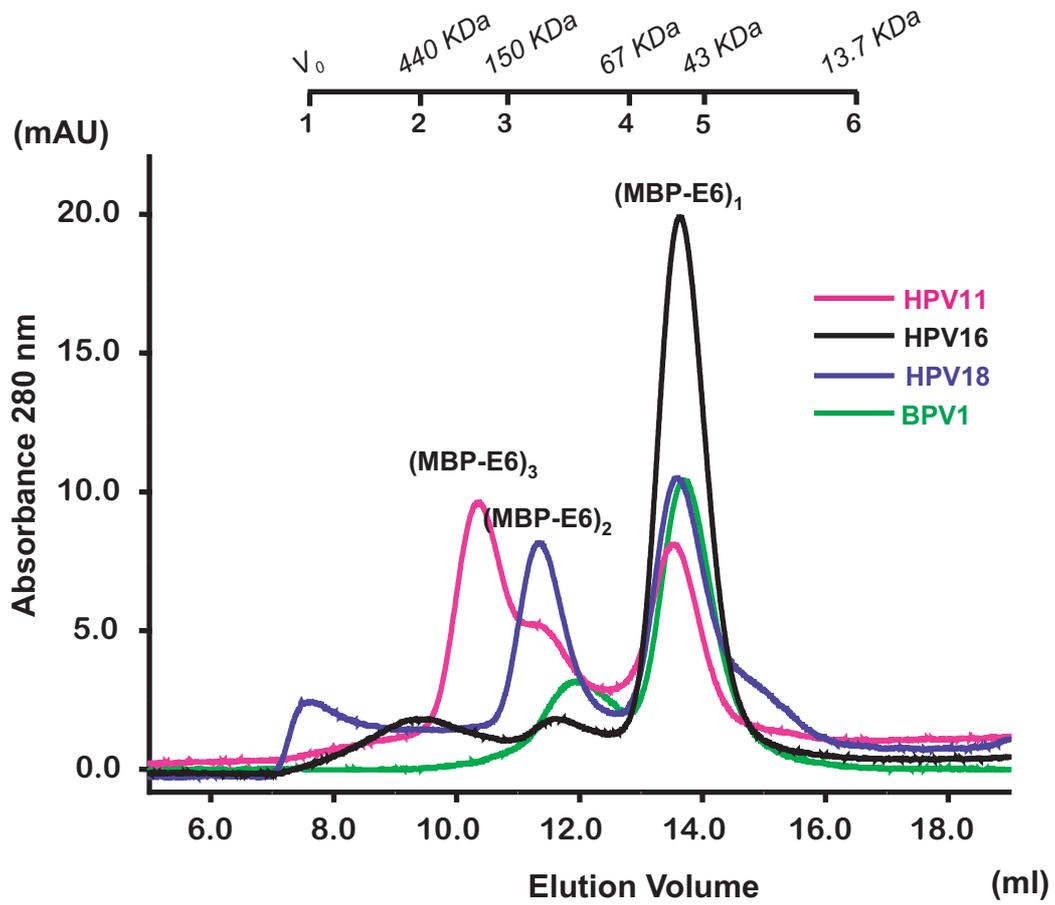
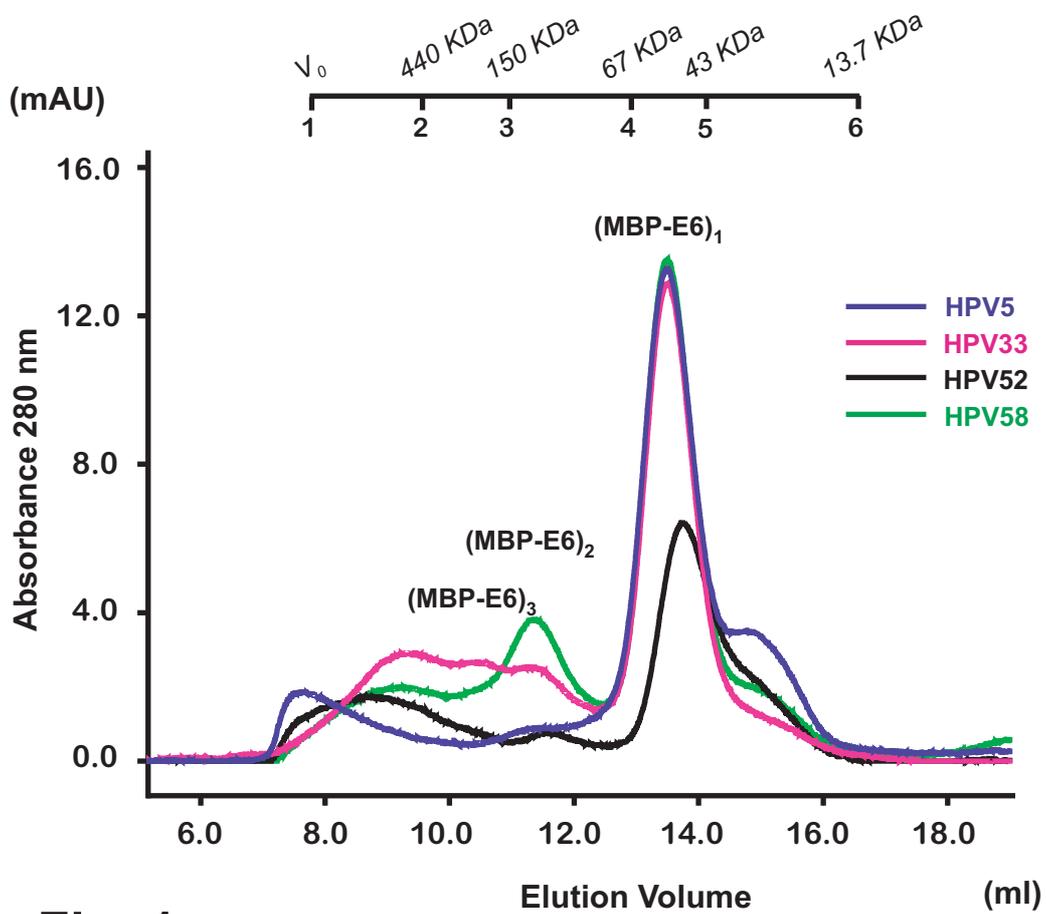
Figure 9 FTIR characterization of aggregates in samples of monomeric HPV16 E6. Superposition of IR spectra of HPV16 E6 in buffer containing either 150 mM (A) or 400 mM NaCl (B) recorded at different time points of the incubation.

Figure 10 Detection of HPV 18 E6 self-association *in vivo* using nucleo-cytoplasmic delocalization techniques. Localisation of Flag-E6, NLS-Myc-E6 and NES-Myc-E6 in HeLa cells transiently transfected: with: (A) 0.8 µg of Flag-E6 ; (B) 3 µg of NLS-Myc-E6 and 0.8 µg of Flag-E6 ; (C) 3 µg of NES-Myc-E6 and 0.8 µg of Flag-E6. Superpositions of cMyc and FLAG images are shown for (B) and (C).

SUPPLEMENTARY MATERIAL

Figure S1 Partial field images of monomeric MBP-E6 preparations from HPV11 (A), HPV18 (B) and BPV1 (C) viral strains immediately after purification (t=0 of incubation). (D) Affinity purified preparations containing HPV16 soluble MBP-E6 aggregates after 3 days of incubation at 20 °C.

Figure S2 Secondary structure analysis performed on the Amide I peak in of spectra of E6 after 1 day (A) and 30 days (B) of incubation (complete spectra shown in Fig. 9A).

A**B****Fig. 1**

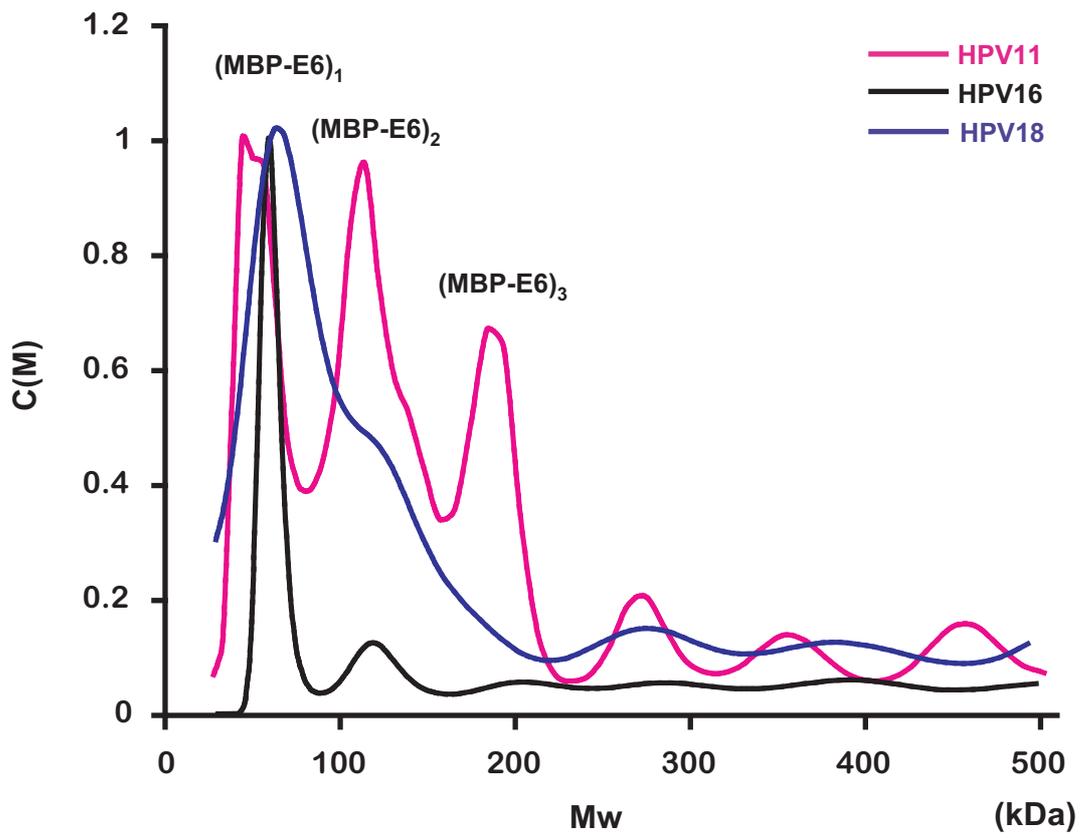


Fig. 2

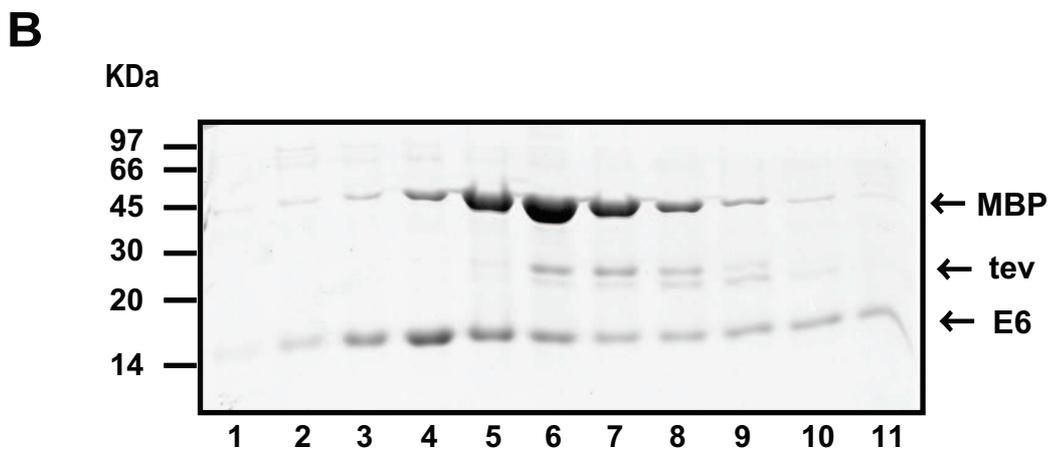
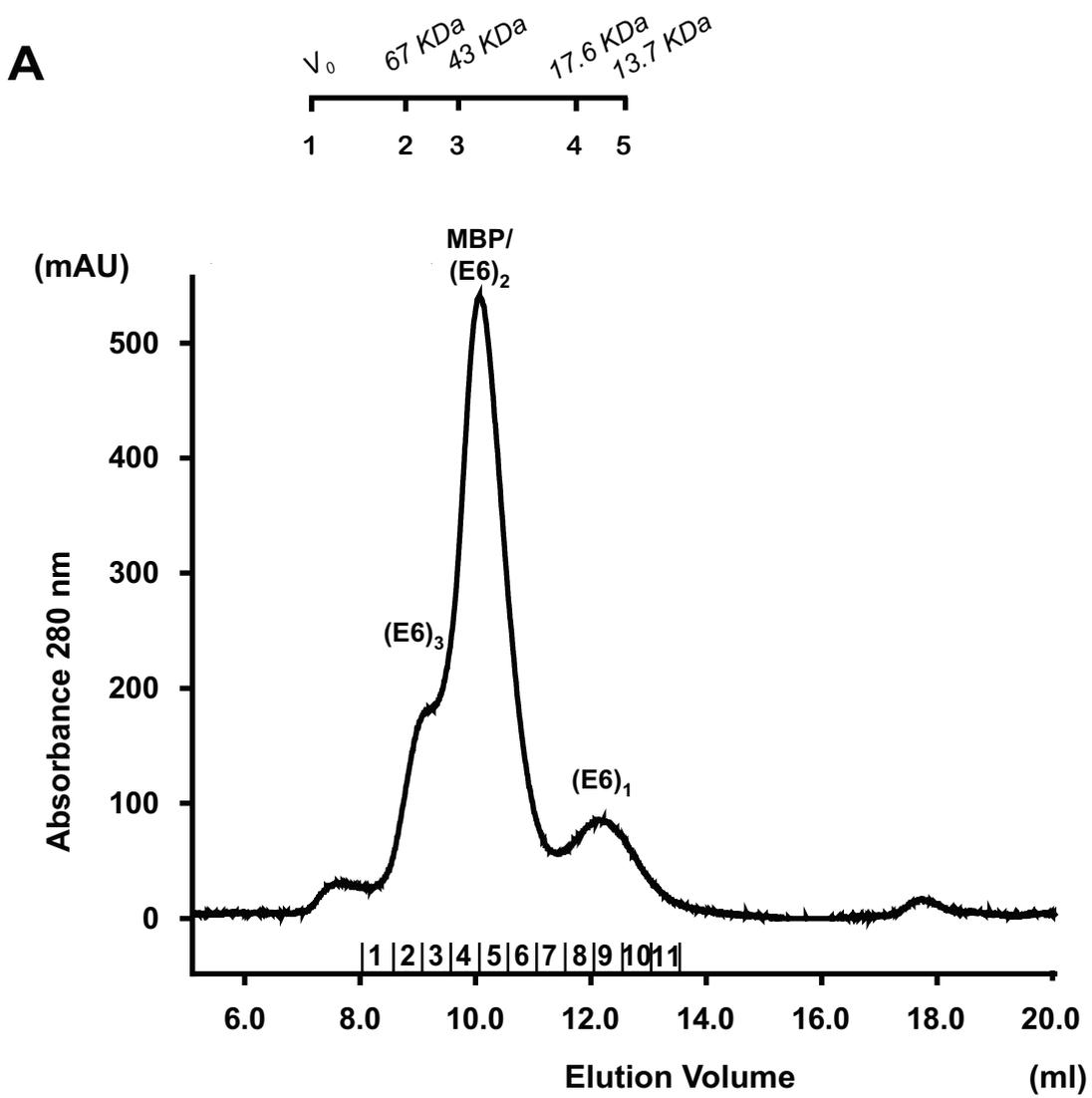


Fig. 3

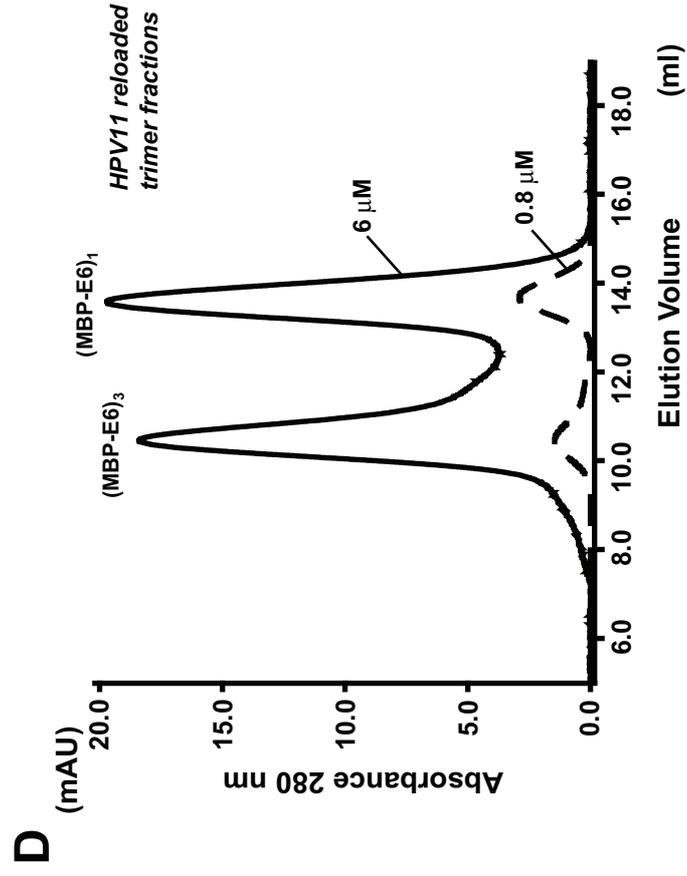
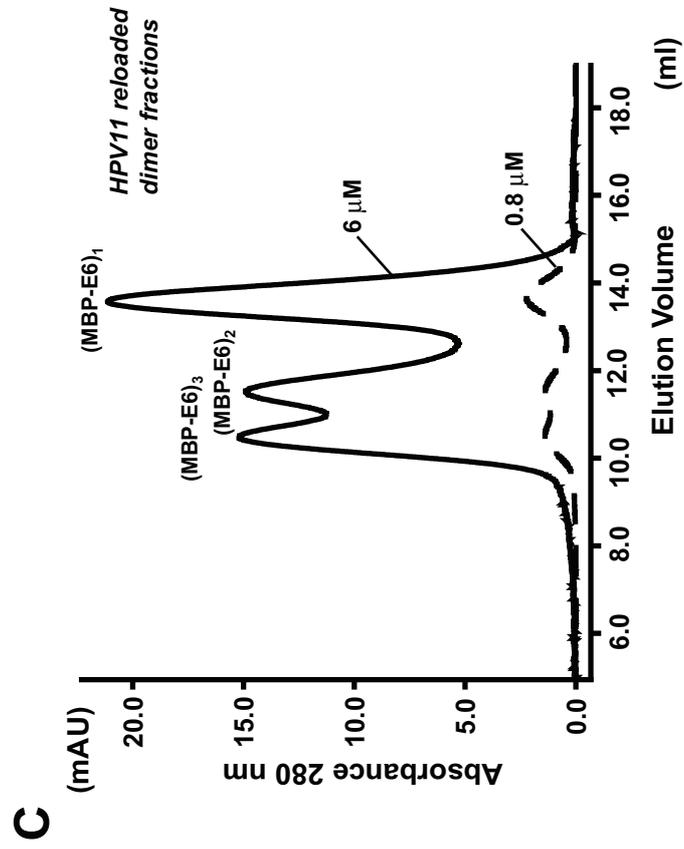
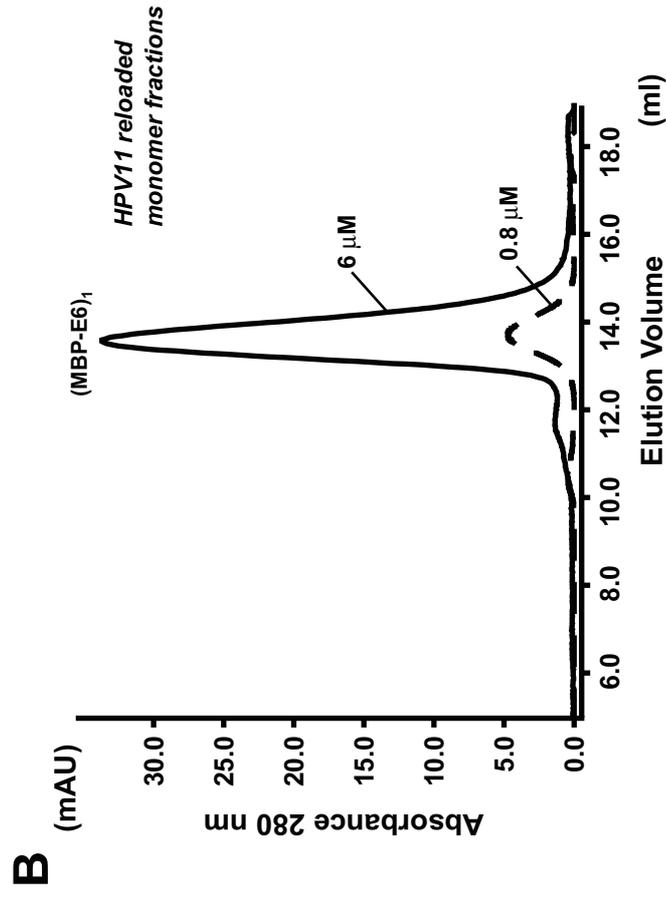
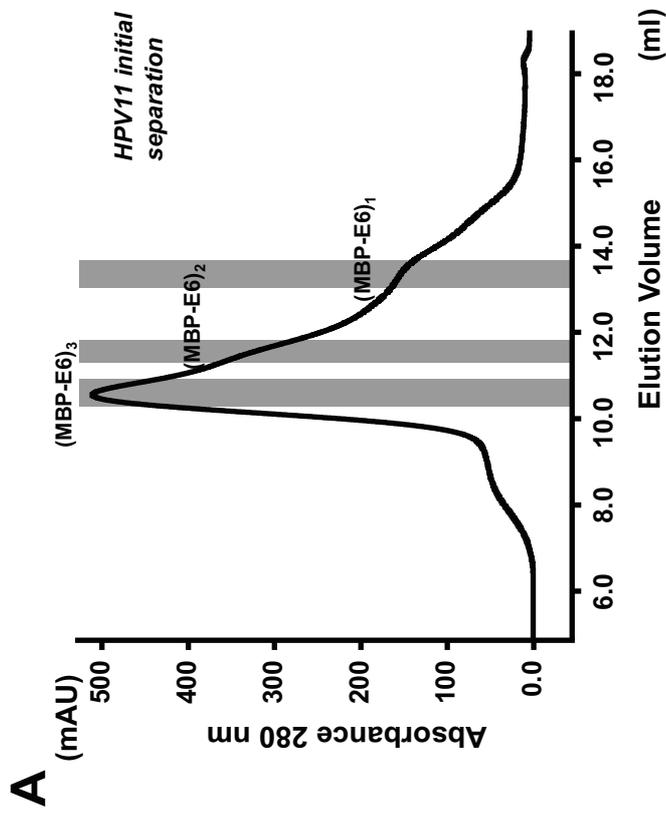


Fig. 4

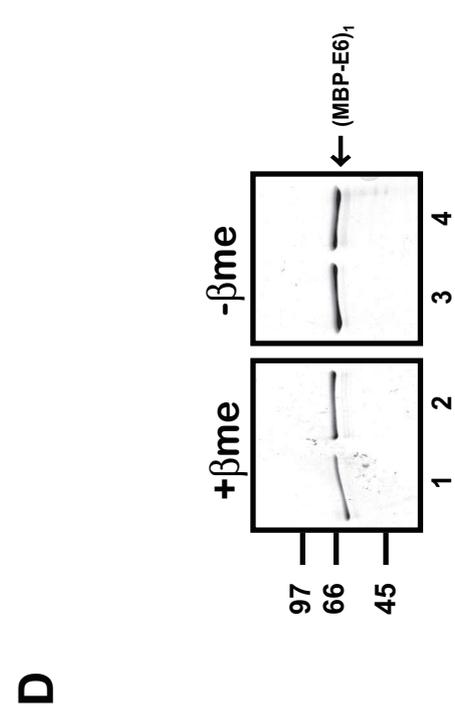
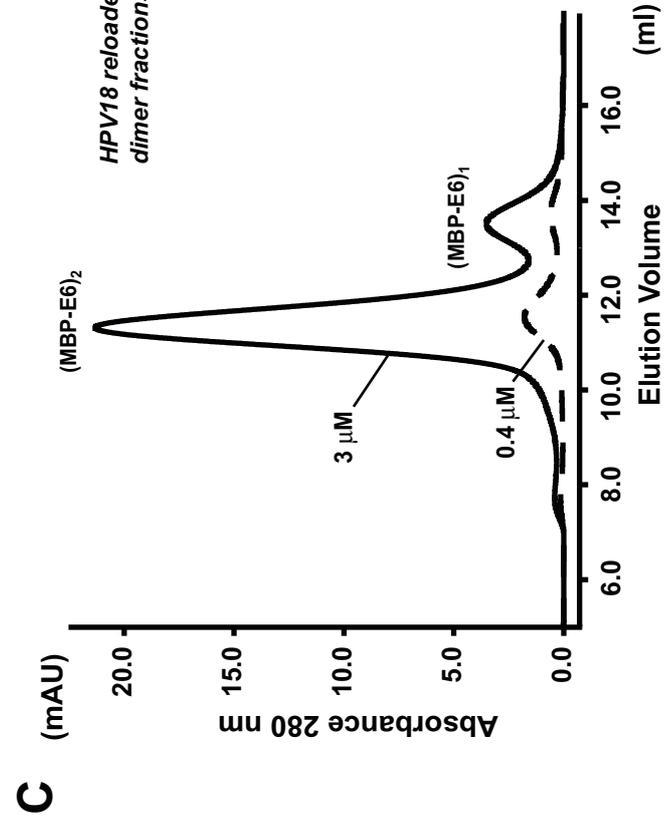
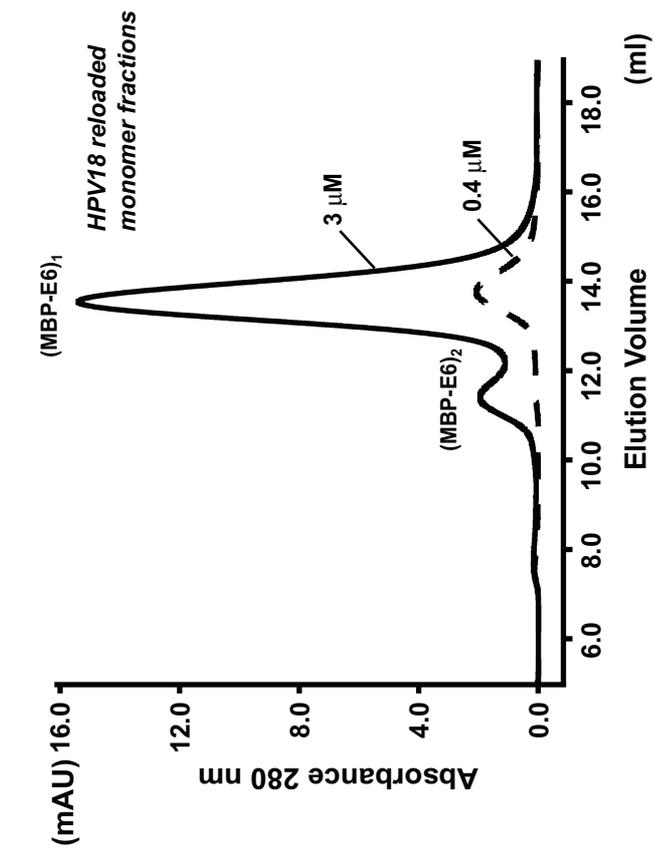
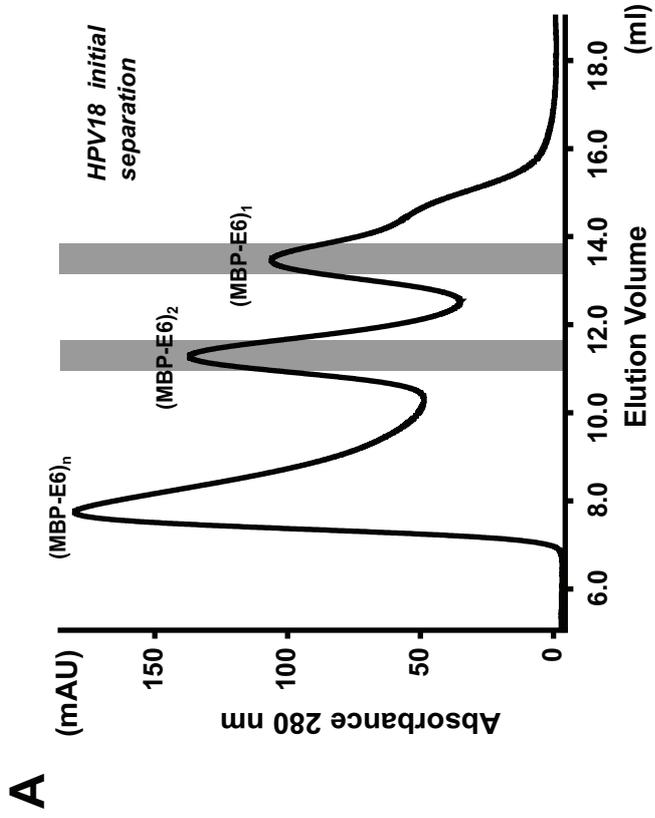


Fig. 5

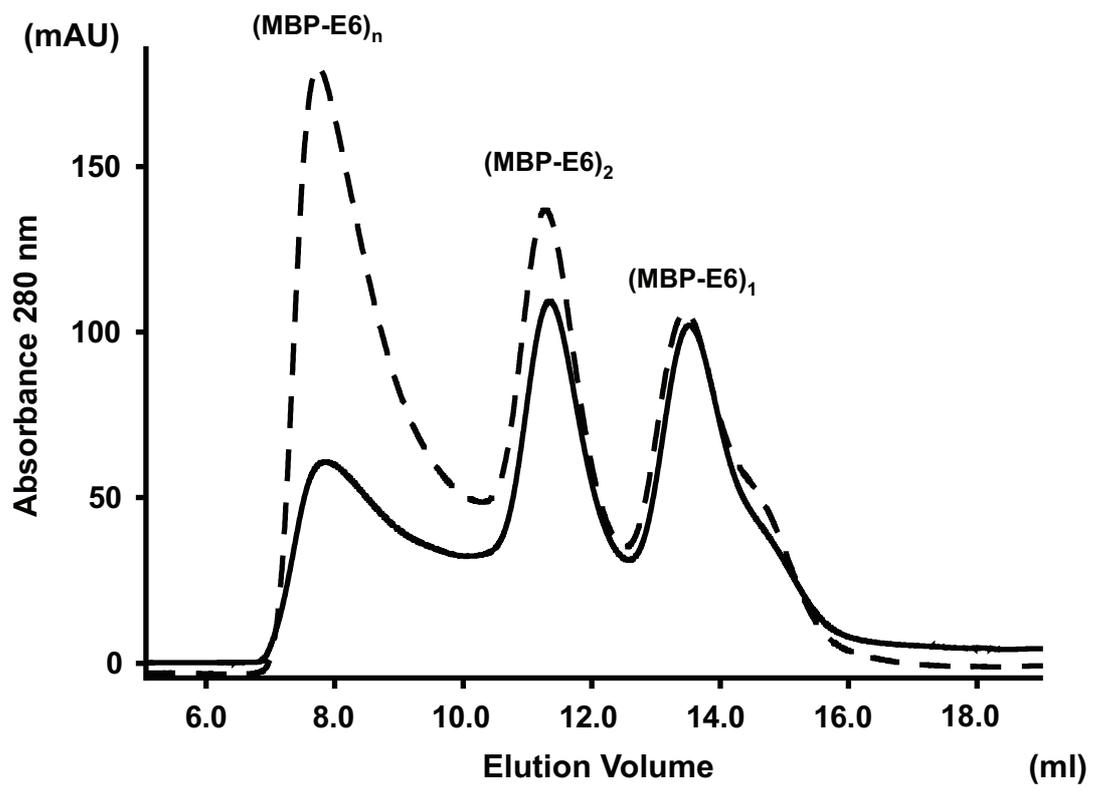
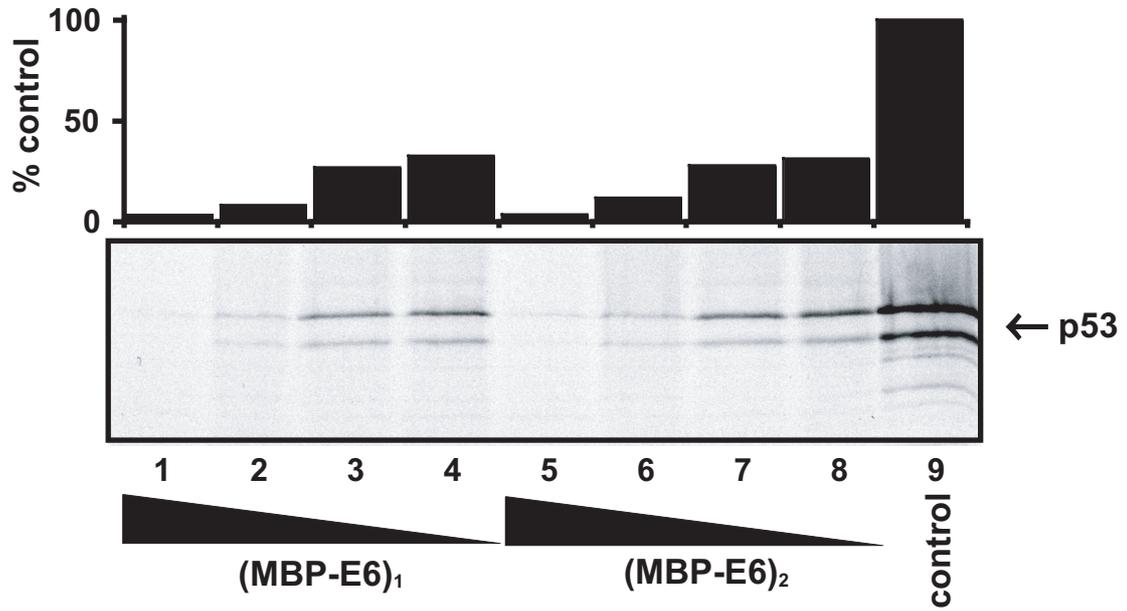
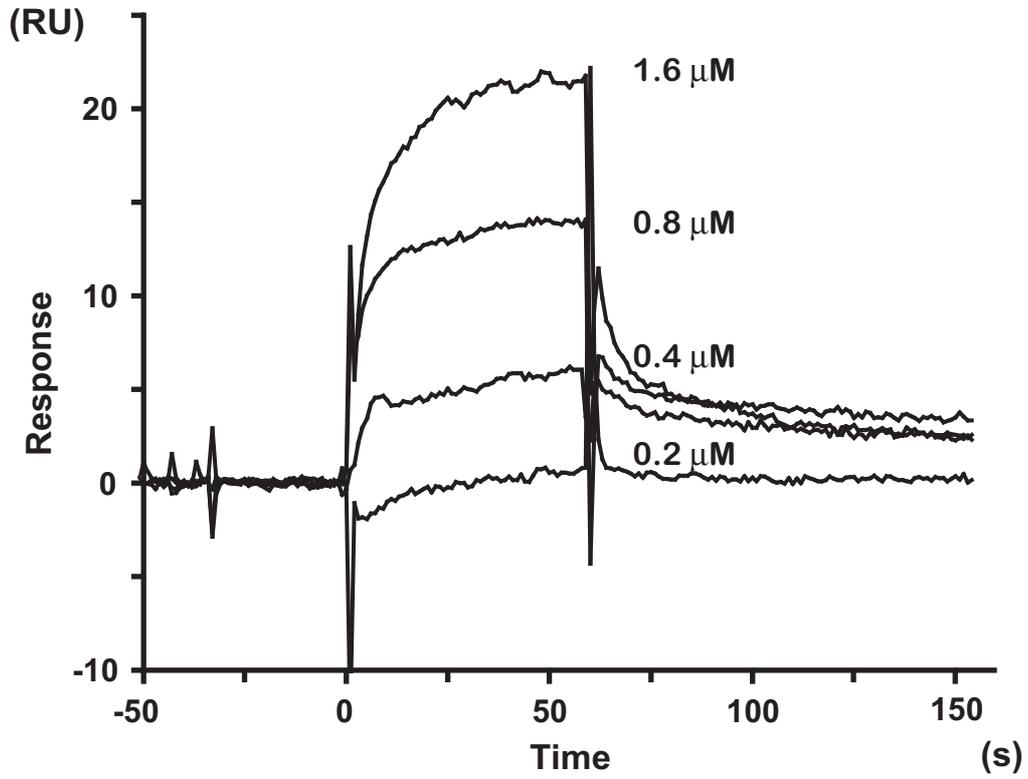


Fig. 6

A**B****Fig. 7**

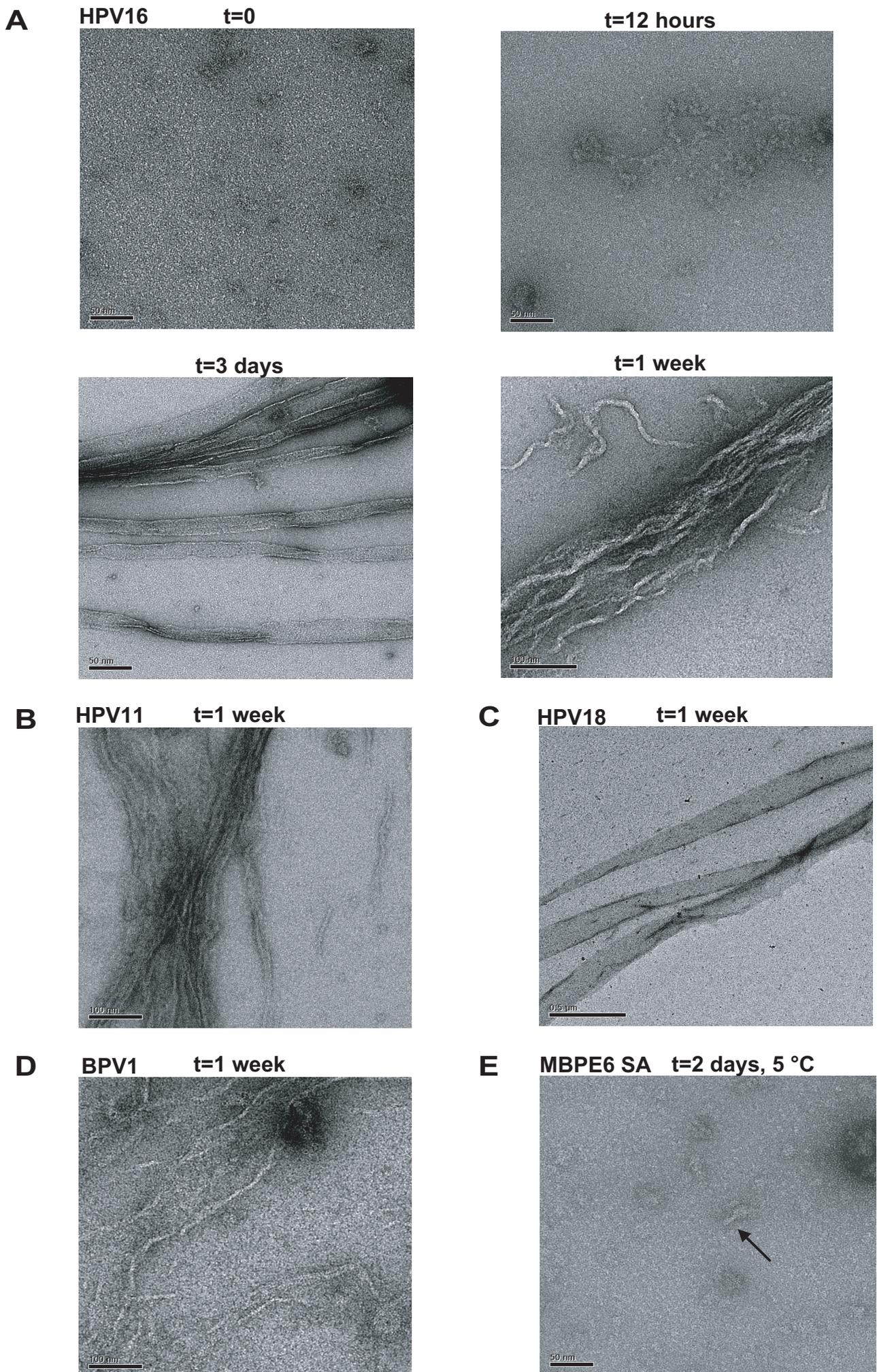


Fig. 8

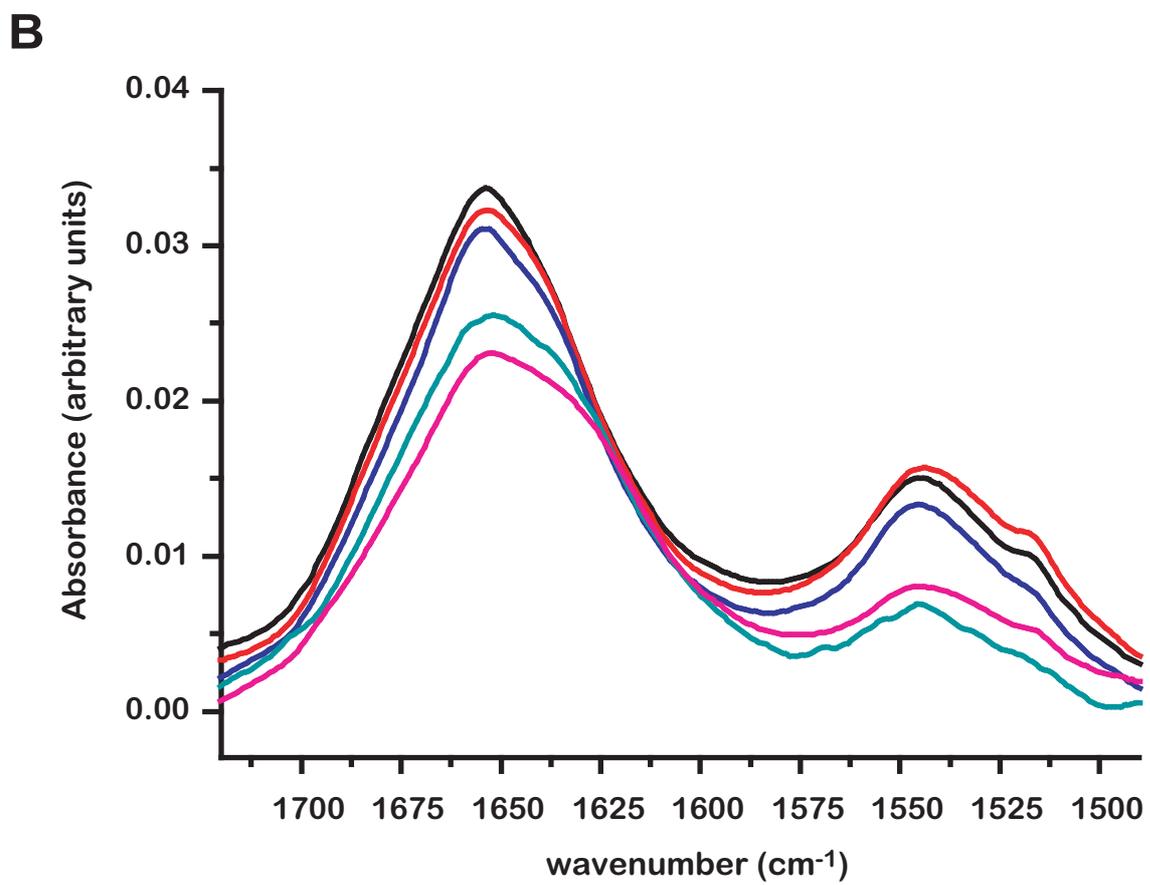
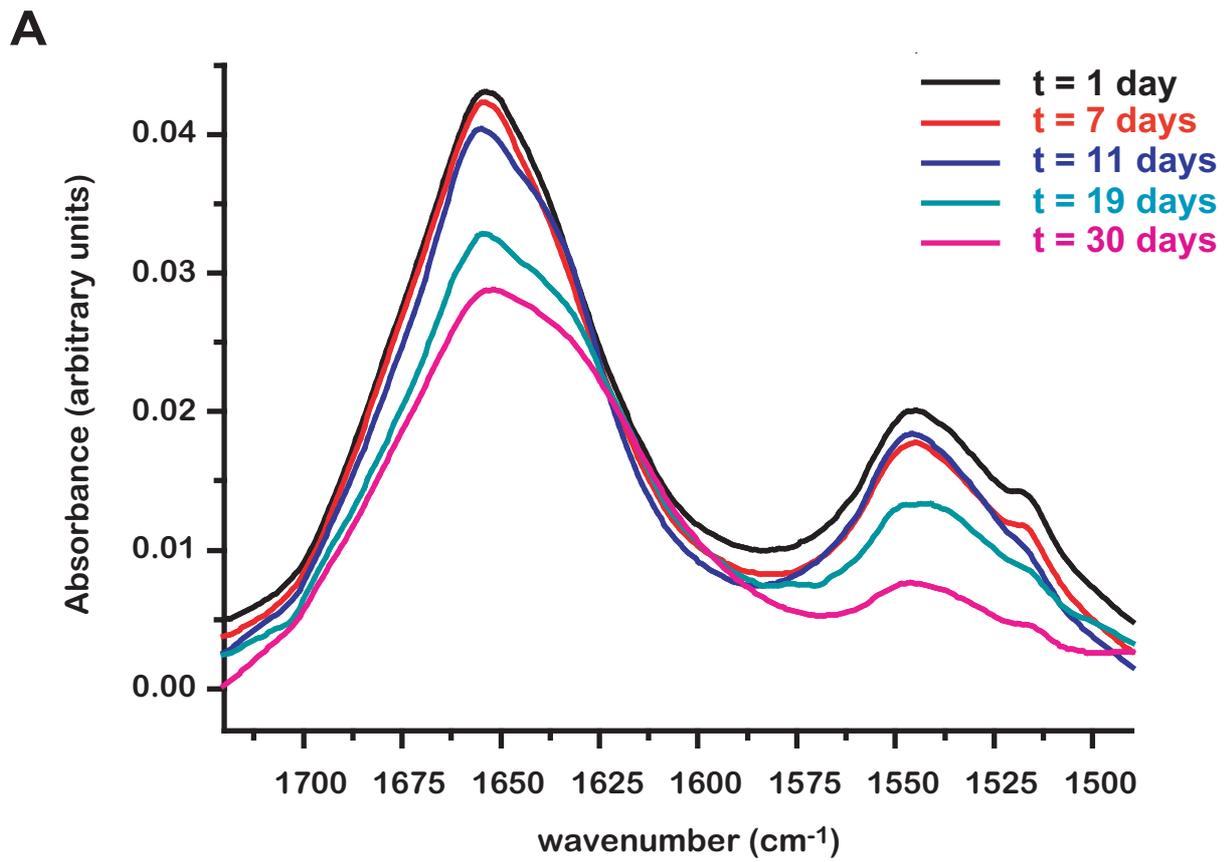


Fig. 9

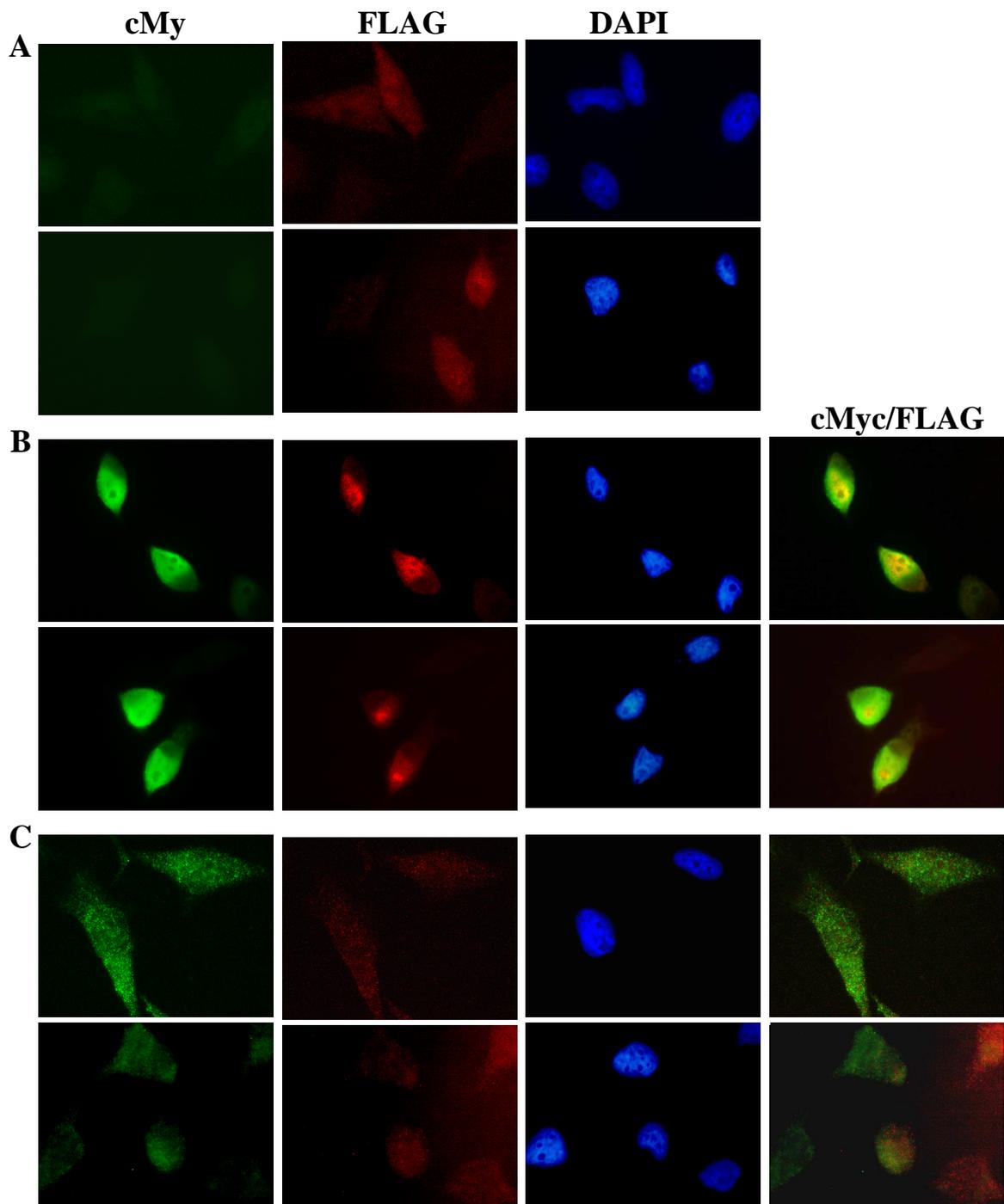


Fig.10

A

HPV11

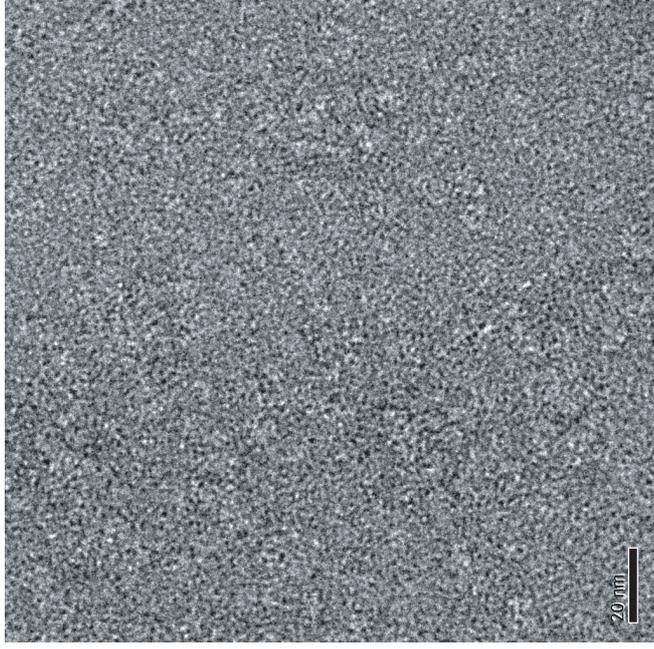
$t = 0$



B

HPV18

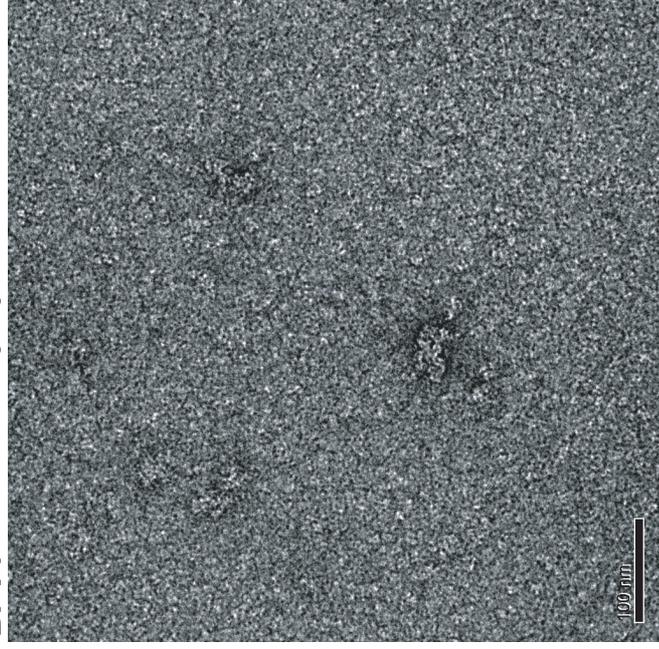
$t = 0$



C

BPV1

$t = 0$



D

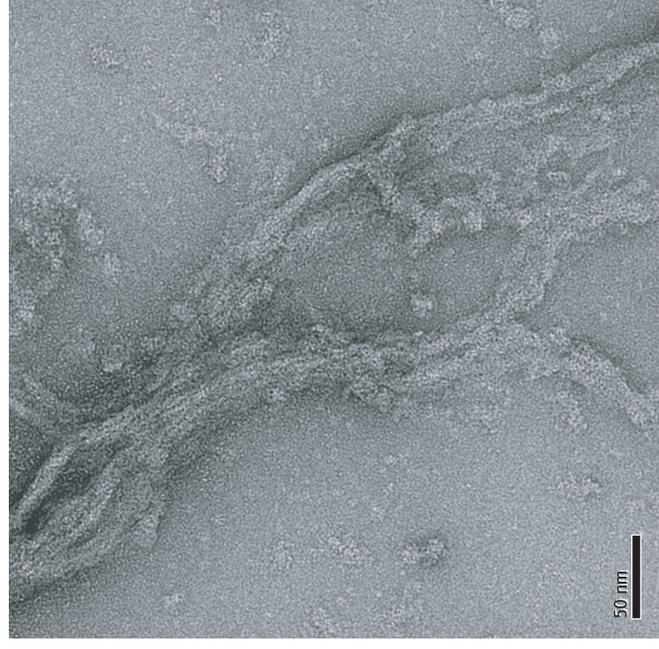
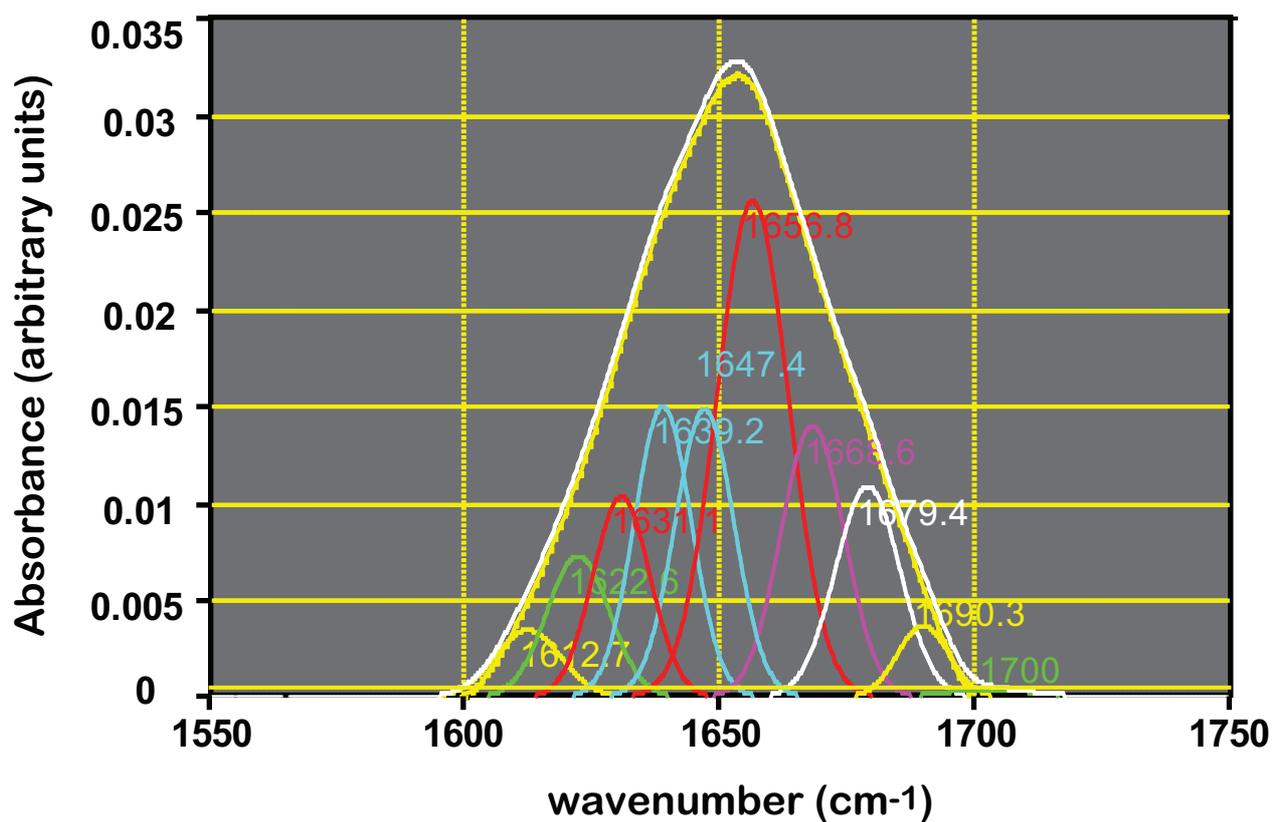
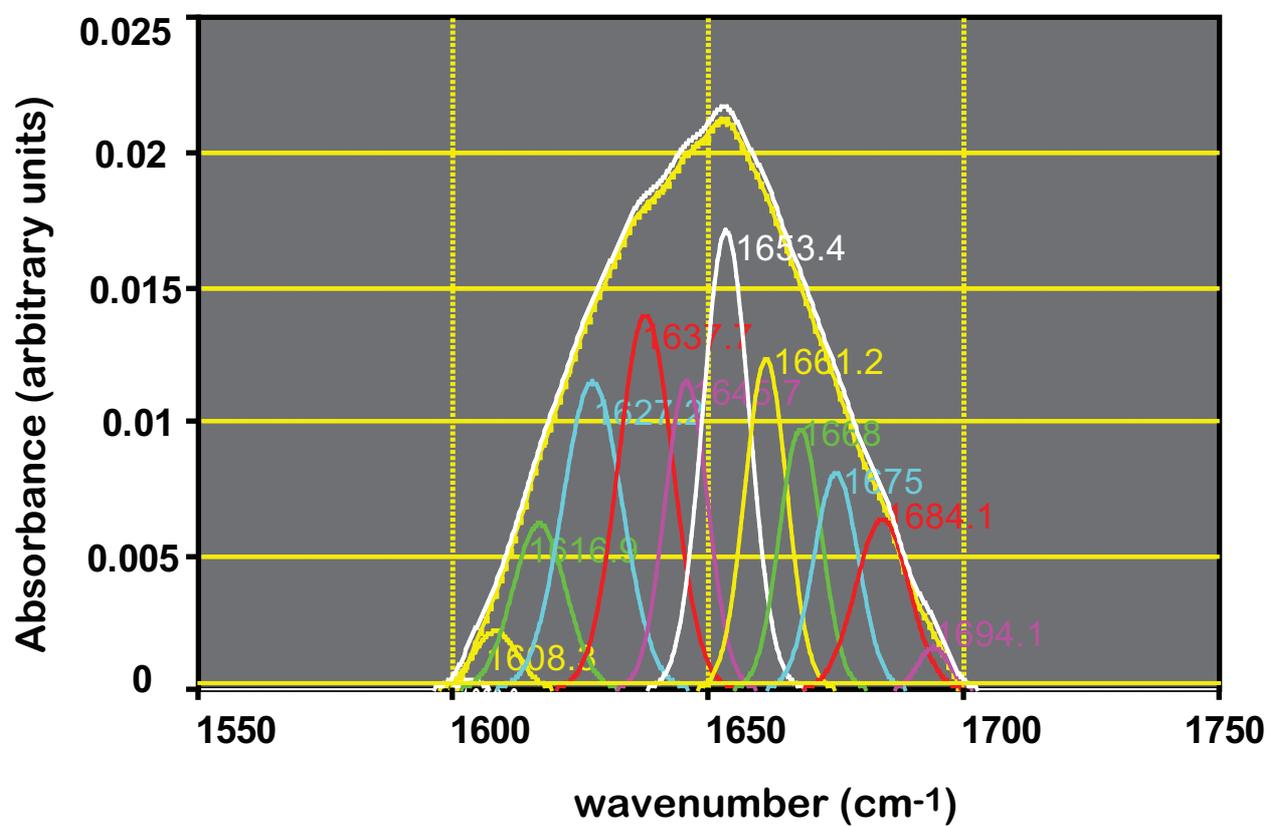


Fig. S1

A**B****Fig. S2**

DISCUSSION GENERALE

DISCUSSION GENERALE

L'obtention d'un échantillon de qualité et quantité suffisantes est la condition sine qua non pour une étude structurale. Mais l'échec systématique de toutes les tentatives d'expression et la purification des oncoprotéines E6 de papillomavirus avait comme résultats un manque évident de données structurales dans la littérature, contrairement au volet biologique largement étudié et discuté.

Les résultats que nous avons montrés dans cette étude constituent à notre connaissance les premières ouvertures vers l'étude structurale des protéines E6 entières de HPV, BPV et de papillomavirus aviaires.

L'expression et la purification de E6 sont deux étapes cruciales, sur lesquelles nous nous sommes attardés en raison des difficultés rencontrées. Nous utilisons des systèmes d'expression bactériens qui sont pourtant largement utilisés, mais les optimisations que nous avons introduites font la différence. La protéine E6 exprimée dans les bactéries a toujours besoin d'une autre protéine porteuse solubilisatrice pour prévenir sa précipitation sous forme de corps d'inclusion. Cependant, la protéine GST largement utilisée comme « facilitatrice » de la purification de la protéine a déjà amélioré la qualité de la protéine E6 exprimée en bactérie (Lechner et Laimins, 1994). Les mêmes auteurs avaient assimilé la GST aux protéines chaperonnes. La protéine de fusion MBP quant à elle, est connue pour son rôle de stabilisateur des protéines mal repliées (Kapust et Waugh, 1999). Nous avons remarqué qu'en ayant cette protéine fusionnée à l'extrémité N-terminale d'une protéine E6, nous gagnons en solubilité. Dans le meilleur des cas nous obtenons principalement des protéines monodisperses, et dans le pire des cas nous obtenons des préparations dominées par des "corps d'inclusion solubles", mais contenant une fraction de MBP-E6 monodisperse plus ou moins importante récupérable dans le surnageant d'une longue étape d'ultracentrifugation. La température d'expression est un facteur important aussi bien que le temps d'induction. L'expression des protéines E6 est effectuée à une température relativement basse ° (15 C), les protéines ainsi exprimées trouveront le temps d'adopter une conformation proche de leur repliement sinon leur repliement original.

Les méthodes de purification d'ores et déjà utilisées en routine dans notre laboratoire, permettent de limiter l'oxydation des protéines E6 riches en résidus cystéines (14 pour HPV16 E6 4C/4S, 14 pour BPV E6 et 5 pour FIPV E6). La purification des protéines E6 fusionnées avec la MBP ou GST ne garantissent pas un repliement correct de E6. La protéine porteuse bien repliée, limite l'association de la protéine E6 (mal repliée) avec d'autres E6, ce qui constitue un état transitoire entre les agrégats et la forme monodisperse.

C'est pour cela que nous ajoutons en général une étape d'ultracentrifugation longue d'une dizaine d'heures pour sédimenter les fractions de MBP-E6 qui se sont proches de l'état agrégé et pour récupérer les fractions encore solubles de la fusion. Toutefois cette étape n'est pas un passage obligatoire pour toutes les constructions que nous avons produites. Les fusions de MBP avec différents domaines (HPV E6-N, HPV-E6-C « sauvage », BE6-N), la protéine FIPV E6 et le complexe covalent LXXLL-BE6 n'ont pas montré de précipitation après l'ultracentrifugation. La force de sédimentation, étant très forte, permet la détection de toute hétérogénéité de populations dans les échantillons et permet d'avoir une première idée du degré de monodispersité, et de la solubilité des échantillons après le clivage entre MBP et E6.

Le repliement des domaines isolés de E6, est largement élucidé dans notre étude. En effet, chaque domaine de E6 se replie isolément. Ceci a été démontré pour le domaine HPV16 E6-N (Lipari et al., 2001), aussi bien que le domaine C-terminal de la même protéine (Nominé et al., 2003). Cependant, aucun de ces auteurs n'étaient parvenus à produire simultanément des échantillons stables des deux domaines utilisables pour l'étude structurale. Au cours de ce travail, nous sommes arrivés à produire les deux domaines de HPV16 E6 sous forme sauvage (E6-N et E6-C). Le domaine N-terminal « sauvage » (1 – 80 aa) que nous avons purifié est une construction plus longue que celle produite par Lipari et al., (2 – 77 aa), et nous avons obtenu les échantillons permettant de résoudre sa structure. Le domaine C-terminal E6-C « sauvage » tout comme le mutant E6-C 4C/4S est soluble. Nous pensons donc que les quatre cystéines non conservées au niveau du domaine C-terminal de HPVE6 16 ne génèrent pas l'agrégation du domaine, mais ne constituent pas non plus des résidus dont la nature est absolument essentielle au bon repliement de la protéine.

Contrairement au HPVE6, à notre connaissance, aucune étude sur les domaines isolés de BE6 n'avait été publiée. Nous avons montré que le domaine BE6-N (par homologie avec HPV16 E6-N) est un domaine soluble, bien replié, alors que les meilleurs échantillons que nous avons obtenu du domaine C-terminal semblent avoir des problèmes de repliement et/ou de solubilité. Les mutations que nous avons introduites dans ce domaine n'ont pas diminué le rapport d'agrégation et la précipitation. Nous avançons donc l'hypothèse que le domaine C-terminal de BE6 est dépendant du domaine N-terminal pour soutenir son repliement.

Une nouvelle séquence de la protéine E6 entière observé chez des papillomavirus infectant d'autres espèces animaux. La protéine E6 de FIPV présente une similarité de séquence avec les deux domaines de HPV16 E6, mais cette protéine ne

contient qu'un seul domaine à zinc. Cette protéine très soluble est monomérique. L'existence d'une protéine E6 "monodomaine" pourrait soutenir l'idée avancée par Danos *et al.*, en 1987 que les deux domaines des protéines E6 proviennent d'une duplication d'un simple domaine (Danos *et al.*, 1987). Les résultats que nous avons obtenus plaident en faveur de cette théorie, et rendent possible la relation entre une forme originale de E6 mono-domaine et la protéine E6 à deux domaines dérivée de l'évolution des virus. L'autre protéine E6 de FPV mono-domaine n'est pas purifiable suivant nos méthodes. L'explication plausible évoque l'implication de plusieurs résidus hydrophobes, probablement exposés à la surface du domaine, qui sont présents dans la séquence de cette protéine par rapport à E6 FIPV où les positions équivalentes sont hydrophiles.

La protéine E7 de FIPV présente des caractéristiques communes aux protéines E7 des papillomavirus infectant les mammifères. L'utilisation des programmes de prédiction de séquences telles que GLOBPLOT et IUPRED nous ont permis de repérer une région N-terminale de E7 comme étant non structurée, ce qui est d'ailleurs caractéristique d'autres protéines E7 issues de différents souches de papillomavirus. A cela s'ajoute la présence du motif conservé *LXCXE* qui sert comme motif de liaison de pRb (Lee *et al.*, 1998; Radulescu *et al.*, 1995). Nous avons produit des échantillons solubles et très concentrés de la région C-terminale de E7 FIPV (les derniers 59 aminoacides). Ce domaine C-terminal purifié est sorti de la gel-filtration sous forme d'un homodimère, ce qui corrèle parfaitement avec ce qui a été décrit dans la littérature concernant le repliement et la structure de la région C-terminal des protéines E7 (Liu *et al.*, 2006; Ohlenschlager *et al.*, 2006)

Les analyses biophysiques de E6 et E7 recombinantes de FIPV ainsi que la localisation de deux protéines dans les compartiments cellulaires évoquent un repliement et une bonne solubilité pour les deux protéines en milieu cellulaire. L'expression abondante et la localisation différentielle de ces protéines suggèrent qu'elles sont fonctionnelles dans leur virus d'origine. En l'absence de toute étude biologique sur ces protéines, nous ne pouvons pas les classer comme des oncoprotéines ; mais le degré de conservation des séquences entre E6 et E7 de FIPV avec leurs orthologues de papillomavirus humains est révélateur d'une certaine ressemblance fonctionnelle.

Les méthodes que nous avons utilisées pour prévenir l'agrégation de la protéine E6 sont loin d'être les mêmes pour toutes les échantillons. Nos résultats montrent que les mutations des cystéines non-conservées dans la séquence de E6 sont largement utiles pour solubiliser la protéine E6 de HPV16, d'où les quatre mutations Cys – Ser (E6 4C/4S). Or, la mutation de deux résidus non-conservés du domaine N-terminal de la

même protéine (E6-N 2C/2S) favorise l'agrégation du domaine. Nous avons également muté les cystéines non-conservées adjacentes de celles liant le zinc (Cys91, Cys128) dans le domaine BE6-C et dans le contexte de la protéine BE6 complète, mais les résultats n'ont pas été conformes à ceux que nous avons obtenus avec E6 de HPV.

La stratégie de co-purification de BE6 avec un peptide issu de son ligand paxilline a permis de produire des excellents échantillons de BE6. Nous avons adopté le même principe afin de tenter de solubiliser la protéine E6 de HPV. Un peptide issu de E6AP ayant été fusionné à l'extrémité N-terminal de E6, la purification du complexe n'a pas donnée l'effet attendu, car cette fusion peptide-protéine, une fois séparée de la protéine MBP précipitait sous forme d'agrégats. Ceci pourrait être lié à la différence remarquable de l'affinité entre paxilline-(LXXLL)/BE6 et E6AP(LXXLL)/E6 HPV. Les constantes de dissociation estimées par BIAcore pour la première et la deuxième interaction sont, respectivement, de l'ordre de 0,1 μM et 4,22 μM (Zanier *et al.*, 2005).

Le domaine N-terminal de HPV16 E6 se dimérise à haute concentration. C'est ce que nous avons remarqué en comparant les spectres de différentes concentrations du domaine E6-N « sauvage ». Le seuil au-dessus duquel la dimérisation commence à dominer est défini à une concentration de 50 μM . La région que nous avons identifiée comme surface probable de dimérisation se révèle bien responsable de ce phénomène. Certains mutants du domaine que nous avons construits ont perdu toute tendance à dimériser même à des concentrations six fois supérieures au seuil préalablement défini.

La force d'ultracentrifugation appliquée pour tracer les profils de sédimentation des différentes populations (monomère, dimère, trimère ...) dans les échantillons de E6-N, a confirmé les résultats obtenus par RMN. Les mutations des résidus que nous avons identifiés « sur la surface de dimérisation » préviennent généralement la formation du dimère, à l'exception de la mutation R48A qui favorise d'avantage la dimérisation du domaine. Les mutants monomères sont également très solubles. La mutation Phe 47 mutée en Arg (F47R) semble avoir un effet très important au niveau de la monomérisation de E6-N aussi bien que la solubilité de la protéine E6 dans sa forme E6 4C/4S F47R.

Le mutant E6 4C/4S F47R, est jusqu'à présent la forme la plus soluble de la protéine E6. Les tests de solubilité attestent de cet état. Les échantillons que nous avons obtenus sont encourageants pour poursuivre les mesures menant à la résolution de la structure. Mais cela est encore difficile, car, la protéine ne supporte pas les concentrations de l'ordre de 500-600 μM (seuil préférable pour la RMN et la cristallographie).

Assurément la dimérisation du domaine N-terminal de E6 serait responsable de l'agrégation de la protéine entière. Etant donné que la concentration de la protéine E6 de HPV16 exprimée dans les cellules semble basse, le phénomène de dimérisation dépendante de la concentration ne doit pas être perceptible dans ce cas.

La dimérisation du domaine N-terminal de HPV E6 a-elle un rôle fonctionnel?

Cette question s'imposait lorsque nous avons remarqué que certaines constructions de E6 portant des mutations favorisant la forme monomérique du domaine N-terminal perdaient certaines activités biologiques, à savoir la dégradation de p53. Nos observations montrent que les mutants qui contiennent le résidu Phe 47 muté, ne dégradent pas le p53 *in vitro* et *in vivo*. Néanmoins, d'autres mutants qui contiennent des mutations ayant empêché la dimérisation de E6-N, dégradent le p53 (par exemple E6 4C/4S D44A, E6 4C/4S H24A). Ce qui signifie que la dimérisation de E6-N ne serait pas nécessaire pour la dégradation de p53. Nous avons également remarqué que l'ensemble de mutants qui portent des mutations dirigés sur la surface de dimérisation de E6-N gardent leur affinité pour les motifs *LXXLL*. L'ensemble donc de ces mutants interagissent *in vitro* avec E6AP.

La dégradation de p53 médiée par E6AP passe d'abord par une interaction E6/E6AP suivi par le recrutement de p53 et la formation d'un complexe trimérique E6/E6AP/p53 (Schwarz *et al.*, 1998). Or, nous avons démontré que les protéines E6 qui contiennent la mutation R47R contribuent à la formation du complexe E6/E6AP *in vitro*, mais ce complexe semble inactif pour le recrutement et/ou la dégradation de p53.

Très récemment, nos collègues Ristiriani *et al.*, ont mené une étude focalisée sur les caractéristiques biologiques d'un simple mutant E6 F47R. En concordance avec nos résultats, ils ont montré que E6 F47R est déficient pour la dégradation de p53. De plus, ce mutant est un dominant négatif capable de freiner la dégradation de p53 (induite par E6 endogène) et restaurer significativement son niveau d'expression dans des cellules HPV positives issues du cancer du col de l'utérus (Ristiriani *et al.*, 2009).

Nous avons mis en évidence des mutants de E6 qui ne reconnaissent pas le motif LXXLL. Le spectre HSQC d'un échantillon de E6 4C/4S a été objet d'une titration avec un peptide minimum synthétisé (non marqué) qui contient la région du motif *LXXLL* de E6AP. Le déplacement chimique traduit par des variations de fréquences (¹⁵N, ¹H) ne concerne normalement que les résidus qui seraient impliqués directement dans cette interaction. Certains mutants ponctuels de E6 4C/4S contenant à chaque fois l'un de ces résidus muté sont plus ou moins solubles que la construction de départ. Dans la plus part de ces mutants, nous avons commuté les résidus concernés par l'acide aminé glutamique (E) pour créer

d'avantage un environnement répulsif au peptide qui contient le motif *LXXLL* de E6AP (chargé négativement). Nos résultats *in vitro* montrent que trois mutants se révèlent non spécifiques pour le peptide *LXXLL* et pour la protéine E6AP entière, ce qui suggère que les résidus L50E, V53E et I128E sont important pour la reconnaissance du motif *LXXLL* et l'interaction E6/E6AP. Raisonnablement, les mêmes mutants (E6 4C/4S L50E et E6 4C/4S V53E) ne dégradent pas la p53 *in vitro*. Ceci est justifié par l'absence de l'interaction E6/E6AP, et donc la suppression de la voie de médiation principale pour la dégradation de p53 à savoir la poly-ubiquitination de celle-ci. Notons que nous n'avons pas obtenus de résultats concluants sur E6 4C/4S I128E, car il est toujours mal traduit dans nos essais *in vitro*, mais, il a été montré qu'un mutant comportant le même résidu muté en alanine (E6 6C/6S I128A) ne dégrade pas le p53 *in vitro* (Nominé *et al.*, 2006).

Néanmoins, nous avons trouvé que ces mutants (E6 4C/4S L50E, E6 4C/4S V53E et E6 4C/4S I128E) ne sont déficients pour la dégradation de p53 que dans le système de dégradation *in vitro*. Les mutants semblent bien dégrader la p53 *in vivo* ce qui contraste à première vue avec nos résultats *in vitro*. Cependant, nous expliquons cela par le fait que la dégradation de p53 *in vitro* dans les systèmes de réticulocytes du lapin dépend uniquement de E6AP, alors que de multiples voies seraient possibles pour la dégradation de p53 *in vivo* (cellules C33A).

Le mutant E6 4C/4S K11E présente une particularité semblable à celle de E6 4C/4S F47 avec bien sûr la différence remarquable de solubilité au bénéfice du dernier mutant. En effet, E6 4C/4S K11E comme E6 4C/4S F47, ne dégrade pas le p53 dans les systèmes *in vitro* et *in vivo*, mais n'empêche pas la formation d'un complexe E6/E6AP. Lorsque nous comparons la détection de E6 dans les cellules C33A transfectées, nous constatons que la localisation nucléaire n'est pas claire pour E6 4C/4S K11E. Contrairement à la protéine E6 sauvage, préférentiellement nucléaire, la E6 4C/4S K11Esemble localisée sans préférence dans le cytoplasme et dans le noyau, bien que le résidu muté (K11) ne se trouve pas à priori dans une séquence NLS (*Nuclear Localisation Signal*) de E6 (Le Roux et Moroianu, 2003).

En somme, nous pensons pour les catégories de mutants qui lient E6AP mais ne ciblent pas le p53, qu'ils sont altérés non pas dans leur interface directe d'interaction avec E6AP, mais pour leur capacité à transduire un changement de conformation nécessaire pour promouvoir l'activité ubiquitine-ligase de E6AP.

Les mutations de cystéines non conservées du C-terminal n'altèrent pas le repliement ni la fonction de la protéine E6. Nous avons réussi à purifier et obtenir des échantillons très solubles et très concentrés du domaine E6-C sous sa forme sauvage, idem

pour le domaine E6-N. Quand ces domaines sont combinés dans la protéine entière, la protéine E6 sauvage est complètement agrégée et difficile à produire, justifiant ainsi le recours vers le mutant E6 4C/4S F47R que nous avons pu produire suffisamment concentré et soluble. Nous avons généré les différentes formes possibles combinant des mutations contre les quatre cystéines non conservé du domaine C-terminale de E6 F47R. L'analyse de la solubilité des différents mutants montrent que le rapport de solubilité est compris entre celui de E6 F47R (sans cystéines mutées) et E6 4C/4S F47R, concluant que ce dernier est le plus soluble. Cette partie est encore préliminaire, parce qu'elle est basée uniquement sur des mini-purifications des échantillons. Côté fonction, E6 4C/4S F47R (quatre cystéines mutée) et E6 F47R (sans cystéines mutées) présentent des résultats similaires pour la non liaison avec E6AP et la dégradation *in vitro* et *in vivo* de p53. Ces résultats sont également observés en utilisant les mutants E6 F47R (Ristriani *et al.*, 2009) et E6 6C/6S F47R qui contiennent les six cystéines non conservées mutées (Nominé *et al.*, 2006). Donc nous pouvons conclure que les mutations des cystéines n'ont pas d'effet contraignant, et que désormais le mutant E6 4C/4S F47R est un bon candidat pour la résolution de la structure finale de E6.

CONCLUSIONS ET PERSPECTIVES

CONCLUSIONS ET PERSPECTIVES

Le pouvoir oncogène des papillomavirus implique essentiellement les deux protéines E6 et E7. D'où le terme « oncoprotéines » communément utilisé pour désigner les protéines incriminées dans le développement du cancer du col de l'utérus associé aux HPV à haut risque. Cependant, à l'exception de l'espèce humaine, les papillomavirus infectent d'autres espèces animales. L'enjeu principal de notre travail focalisé sur le volet structural, se limite à l'étude de la protéine E6, encore peu explorée. Cependant, notre démarche s'élargit à d'autres protéines E6 issues des papillomavirus bovins et aviaires, qui présentent des particularités structurales originales en comparaison à leurs orthologues humains.

La structure de E6 a résisté pendant vingt ans à toute analyse, car la protéine était très difficile à produire par voie recombinante sous une forme repliée. Tout au début de notre travail, nous avons commencé à améliorer les conditions d'expression et purification de la protéine E6, ce qui a permis de lever les derniers obstacles empêchant la production de la protéine recombinante. La structure du domaine E6-N est résolue par RMN. De plus, nous avons démontré la présence d'un phénomène d'auto-dimérisation du domaine E6-N dépendant de la concentration ; la dimérisation se déclenche spontanément au-delà d'un seuil de concentration de 50 μ M. Des mesures RMN ont révélé l'implication directe de certains résidus comme surface de dimérisation, ensuite nous avons muté ponctuellement les résidus de cette surface. Les mutants qui en résultent sont analysés par ultracentrifugation analytique mettant en évidence certaines formes du domaine E6-N très solubles et résistants à la dimérisation. Les mêmes mutations introduites dans le contexte de la protéine complète, n'ont pas été moins intéressantes: nous disposons désormais d'échantillons très solubles de E6. Nous avons également pu produire le domaine C-terminal sous sa forme sauvage.

La construction E64C/4S F47R est un mutant de E6 combinant le domaine E6-N F47R et le domaine E6-C 4C/4S. Ce mutant est choisi comme un candidat potentiel pour la résolution de la structure de E6. Nous avons déjà obtenu des spectres HSQC (15 N, 1 H) qui constituent les premiers spectres d'un échantillon de E6, et argumentent la qualité du repliement de cet échantillon. De plus, une série des mutants incluant la mutation du résidu F47 ne dégradent pas le p53 alors qu'ils contribuent à la formation d'un complexe avec E6AP. Le mécanisme de la défaillance de dégradation de p53 par ces mutants n'est pas encore élucidé.

Nous avons repéré quelques résidus qui sont essentiels pour l'interaction de la protéine E6 avec son ligand principal E6AP. Ces résidus sont essentiellement K11, L50, V53 et I128. La mutation de l'un de ces résidus conduit à la perte de l'affinité entre les deux partenaires. Les mutants résultant d'une commutation de l'un de ces résidus sont déficients pour la dégradation de p53, action majeure de E6 pour la transformation des cellules.

Parallèlement, nous avons cherché à généraliser notre étude sur les modèles plus ou moins proches de la protéine E6 du HPV.

Tout d'abord, nous avons cloné et produit les protéines E6 et E7 recombinantes issues d'un nouveau papillomavirus infectant des oiseaux (FIPV). Les échantillons marqués ^{15}N nous ont permis d'avoir des spectres HSQC de deux protéines qui se révèlent très solubles. La protéine E6 de FIPV mono-domaine est un monomère, contrairement à ce que nous attendions, mettant en évidence un nouveau repliement de la protéine E6. La partie C-terminale de la protéine E7 de FIPV que nous avons purifiée, présente une caractéristique commune entre les domaines C-terminaux des protéines E7, à savoir la dimérisation obligatoire de cette région. Ensuite, nous avons cloné les séquences codantes (ORF) des deux protéines fusionnées avec des séquences « marqueurs » pour la localisation cellulaire. Les deux protéines E6 et E7 de FIPV se localisent d'une façon étalée dans le noyau et le cytoplasme, dans différentes souches cellulaires (QT6 de cailles, et HeLa).

Nous avons également travaillé sur la protéine E6 issue du papillomavirus BPV1 qui, tout comme HPV16 E6 contient deux domaines à zinc. Tout au début, nous avons rencontré les mêmes problèmes d'agrégation qu'avec HPV16 E6. Après la séparation de deux domaines de la protéine, le domaine N-terminal (BE6-N) était très soluble contrairement au domaine BE6-C. La mutagenèse des cystéines non conservées du domaine BE6 n'a pas amélioré sa solubilité, ce qui suggère qu'il présente, sous sa forme isolée, soit un défaut de repliement soit des surfaces hydrophobes exposées provoquant son agrégation. Pour contourner le problème d'agrégation de la protéine BE6, nous nous sommes intéressés à l'interaction de BE6 entière avec un peptide issu de la paxilline, une protéine cellulaire connue pour sa forte liaison à BE6. Ce peptide contient le motif consensus *LXXLL* reconnu par la plupart des protéines E6. Par des mesures BIAcore, nous avons d'abord démontré la haute affinité (K_d de l'ordre de $0,1\mu\text{M}$) de BE6 pour le peptide *LXXLL* de la paxilline. Ainsi née l'idée d'une exploration de différentes approches visant à stabiliser la protéine BE6 par le peptide *LXXLL*. Alors que la protéine BE6 non fusionnée s'avérait difficile à purifier, nous avons réussi à préparer des échantillons marqués ^{15}N d'une protéine de

fusion *LXXLL*-BE6 ainsi qu'un complexe non covalent *LXXLL*/BE6, tous deux suffisamment concentrés (300 μ M) pour l'analyse RMN. En jouant sur ces stratégies complémentaires, il est désormais possible d'analyser la structure de BPV-E6 liée à son peptide-cible.

En conclusion, la difficulté de l'expression et de la purification des formes recombinantes, est l'apanage commun des protéines E6 toutes souches virales confondues. Les travaux que nous avons menés durant cette thèse ont permis de compléter l'étude structurale et fonctionnelle de la protéine E6 de HPV initiée il y a 15 ans, et d'aborder des études comparables pour ses orthologues de différents virus animaux. De telles données constituent un progrès considérable attendu depuis le début du projet. L'optimisation des protocoles, l'exploration des domaines, la mutagenèse dirigée et la solubilisation des protéines à l'aide de leurs ligands, se sont avérées essentielles pour aboutir à l'étude structurale. Remarquablement, pour produire et solubiliser des protéines relativement proches, il nous a fallu adopter différentes stratégies. D'autre part, la maîtrise de la production et de la qualité de ces protéines nous a permis d'étudier le rôle biologique de certains phénomènes (dimérisation, liaison à *LXXLL* ...), par une approche de mutagenèse respectueuse du bon repliement des protéines.

Les perspectives qui sont ouvertes après ce travail sont essentiellement liées à la résolution de la structure de la protéine E6. Bien que nous avons pu purifier des échantillons marqués très solubles du mutant E6 4C/4S F47R, nous ne pouvons pas pour le moment atteindre des concentrations très élevées (au-delà de 600 μ M). De telles concentrations pourraient être requises pour compléter certaines mesures supplémentaires pour les calculs menant à la résolution de la structure par RMN. Toutefois, le mutant E6 4C/4S F47R a perdu une fonction essentielle: la dégradation de p53. Une construction comme le mutant E6 4C/4S D44R, qui résout le problème d'agrégation tout en gardant la capacité de dégrader p53, pourrait s'avérer un candidat préférentiel pour la résolution de la structure.

Dans un proche avenir, la nouvelle approche de co-purification que nous avons adoptée permettra de résoudre la structure du complexe non covalent *LXXLL*/BE6 en solution par RMN. De plus une optimisation de la co-purification de la protéine E6 de HPV avec un peptide issu de E6AP pourrait permettre la production des échantillons du complexe analysables par RMN ou cristallographie. Dans le cas contraire, les structures à haute résolution du complexe *LXXLL*/BE6 et de la forme libre d'un échantillon très soluble de E6 HPV seront combinés avec les données de titration HSQC de E6 HPV avec le peptide E6AP, afin de modéliser la structure du complexe *LXXLL*/E6 HPV.

À l'issue de ce travail, la structure du domaine N-terminal de E6 de HPV 16, est d'ores et déjà résolue par notre laboratoire qui avait préalablement publié celle du domaine C-terminal. A cela s'ajoute la structure de E6 mono-domaine de FIPV qui sera en mesure d'être résolue très prochainement. Un travail supplémentaire devient nécessaire pour l'analyse de l'architecture des structures de ces différents domaines à travers d'alignements structuraux, corroborés avec des superpositions tridimensionnelles des trois structures. Ensuite, les résidus et motifs qui s'avèreraient important pour le repliement seront utilisés comme des « hits » pour cribler les séquences de l'ensemble d'une centaine de protéines E6 de différents papillomavirus connus à ce jour. Au fur et à mesure nous disposerons progressivement d'une banque de données d'informations structurales sur les protéines E6. La modélisation par homologie des architectures structurales facilitera certainement la compréhension de la logique du repliement des protéines E6 en corrélation avec leurs fonctions.

A long terme, la structure de la protéine E6 de HPV libre, et en complexe avec le peptide *LXXLL*, une fois résolues, pourraient ouvrir la voie vers le criblage *in silico* de molécules chimiques interférant avec l'interaction E6/E6AP. Les molécules ainsi sélectionnés seront criblés de nouveau *in vitro* et *in vivo* pour choisir de potentiels inhibiteurs de cette interaction. De tels inhibiteurs pourraient s'avérer de futurs candidats pour des médicaments ciblant les lésions CNI 1 et CNI 2 ainsi que les tumeurs du col de l'utérus induits par les HPV à haut risque.

MATERIELS ET METHODES

MATERIELS ET METHODES

Bactéries d'expression:

Nous avons travaillé avec les souches BL21 (DE3) (Invitrogen life technologies), BL21 plyS, et BL21 Codon Plus (fourni par Christophe Romer) pour l'expression des protéines.

Préparation des bactéries compétentes:

Les cellules BL21 électro-compétentes ont été préparées à partir d'une préculture sur la nuit à 37°C. Elle est ensuite diluée au 1/40^{ème} dans 1 litre de milieu de culture L-Broth (LB) et incubée de nouveau à 37°C. Lorsque la densité optique à 600 nm a atteint 0,6, les bactéries sont placées immédiatement sur la glace pendant 30 min et puis sont réparties dans 4 tubes de 250 ml puis centrifugées à 4°C à 4000 rpm pendant 15 minutes. Le froid étant un facteur clé dans la préparation, la suite de manipulation a été effectuée à 4°C. Le culot est lavé deux fois avec de l'eau stérile, d'abord avec 250 ml puis 125 ml. Puis le culot a été lavé avec 5 ml de l'eau bidistillée stérile contenant 10 % de glycérol et un dernier cycle de centrifugation. Finalement, nous ajoutons au culot 500 µl de l'eau contenant 10 % de glycérol, nous homogénéisons le milieu et nous le répartissons dans des tubes Eppendorf 1,5 ml stérile (60 µl/tube). Les tubes sont immédiatement transférés à -80°C.

Transformation par électroporation:

60 µl de bactéries sont transformées par électroporation avec 0,7 µl du vecteur (500ng) codant pour la protéine désirée, en utilisant un appareil Gene Pulser™ (BioRad) réglé à 200 Ω, 2,5 V et 25 µF. Elles sont ensuite reprises dans 1 ml de LB et incubées 30 min à 37°C. Les bactéries transformées sont ensuite diluées dans un volume de 10 ml de LB correspondant à la préculture.

Précultures:

La solution de transformation est diluée directement dans une préculture au 1/200^{ème}. Les précultures destinées au marquage M9 (en milieu minimum M9 N¹⁵) sont incubées dans un volume de 50 ml de LB, alors que les non marquées sont incubées dans un volume de 10 ml de LB. L'incubation se déroule toute une nuit à 37°C avec une agitation de 250 rpm.

Milieu LB (Luria Broth):

Ce milieu est utilisé comme milieu non marqué de culture de bactéries. Un litre de LB (Becton Dickinson Microbiology Systems) est composé de : 10g de digestion pancréatique de caséine, 5g d'extrait de levures, 5g de chlorure de sodium. Pour l'expression des protéines E6, le milieu est ajusté avec une concentration de 100 μM de ZnSO_4 au moment de l'induction de cultures.

Milieu minimum M9:

J'ai utilisé un milieu spécifique pour l'expression et le marquage des protéines, c'est le milieu minimum dont la composition est la suivante: 1 L du milieu contient: 8 g Na_2HPO_4 (2 H_2O), 3 g KH_2PO_4 , 0.5 g NaCl , 0.5 g NH_4Cl , 0.85 mg ZnSO_4 , 8.4 g FeCl_3 (6 H_2O), 0.10 mg CuCl_2 , 0.1 mg CoCl_2 (6 H_2O), 11 mg MnCl_2 (2 H_2O), 18 mM EDTA, 1 mM MgSO_4 , 0.33 mM CaCl_2 , 4% D-glucose, 1 mg Biotine, 1 mg Thiamine (pH 7.5). Pour un marquage M9, $\text{N}^{14}\text{H}_4\text{Cl}$ est substitué par $\text{N}^{15}\text{NH}_4\text{Cl}$ (M9 ^{15}N) tandis que pour un double marquage N^{14} et C^{13} , le glucose C^{12} a été substitué par un glucose C^{13} marqué (M9 N^{15} C^{13}). Il faut aussi signaler que pour l'expression des protéines E6, l'utilisation de l'EDTA est strictement déconseillée, pour ne pas chélater les ions du zinc nécessaires pour le bon repliement de la protéine.

Expression des protéines E6 :***Expression des échantillons non marqués***

Les précultures laissées sur la nuit à 37 $^\circ\text{C}$, sont diluées au 1/500^{ème} dans un volume 500 ml de LB qui contient la concentration adéquate d'antibiotique suivant le vecteur d'expression et la souche bactérienne utilisée (15 $\mu\text{g/ml}$ Kanamycine, 35 $\mu\text{g/ml}$ Chloramphénicol, 34 $\mu\text{g/ml}$ Spectinomycine ou 100 $\mu\text{g/ml}$ ampicilline). La culture est incubée sur un agitateur à 37 $^\circ\text{C}$ jusqu'à que la DO_{260} (densité optique à 600 nm) atteint une valeur de 0,8. L'induction de l'expression se fait grâce à l'IPTG à une concentration de 0,5 mM final, simultanément avec l'induction le zinc est ajouté au milieu sous forme de $\text{Zn}^{2+}\text{SO}_4^{2-}$ à 100 μM final. Ensuite les cultures sont incubées à 15 $^\circ\text{C}$ toujours en agitation de 250 rpm pendant une nuit (environ 15 h). Les cultures sont centrifugées dans des pots de 500 ml dans un rotor J.L.A. 10.500 monté sur une centrifugeuse Beckman (5000 rpm). Les culots sont prêts pour l'utilisation, ou bien pour le stockage dans un congélateur à -80°C .

Expression des échantillons marqués

Les précultures (50 ml) sont récupérées et centrifugées dans des flacons de 50 ml à 4000 rpm pendant 20 min. Après avoir éliminé le surnageant, le culot ré-suspendu dans 10

ml de LB pour être ensuite dilué dans un volume de 500 ml de M9. Les cultures sont incubées à 15 C°; la suite est la même pour une expression non marquée.

Précautions particulières pour la purification de E6:

Avant de commencer de purifier la protéine E6, il est bien nécessaire de suivre quelques conseils. La température ne doit pas connaître des fluctuations lors de la purification, pour cela, un bac de glace sera un bon compagnon sur la paillasse, et en déplacement.

Les tampons utilisés doivent avoir été bien filtrés, dégazés et avoir passé par un système délaissant ses dernières bulles d'oxygène au bénéfice de l'Argon. La sonication est une étape critique, l'échantillon est exposé à la fois au surchauffage et à la production de mousse. Il faut donc absolument être très minutieux à cette étape, l'échantillon doit rester au froid pendant la sonication qui doit être courte et efficace. L'amplitude des sondes doit être raisonnablement choisie en fonction de la taille de la sonde et du volume à soniquer. Après la lyse des cellules, une concentration minimale de DTT (2 mM) est requise pour garantir un environnement réducteur, celle-ci doit accompagner l'échantillon tout au long de la purification, au stockage. En effet, la présence du DTT dans les échantillons est sujette d'une oxydation probable suite aux différents aléas (un tube qui tombe, les bulles d'air, une hausse de température ...), Il est donc très conseillé de garder un œil sur ce point et d'ajuster soigneusement la concentration du DTT chaque fois que cela s'avérerait nécessaire.

Au terme de la purification, l'échantillon de E6 est entre nos mains. D'abord, il ne faut qu'il soit trop concentré, la concentration du NaCl doit être maintenue à 400 mM, ensuite, nous pouvons garder notre échantillon dans les réfrigérateurs (4 °C). Dans ce cas, il faut penser à prévoir un petit moment (parfois très long, selon la solubilité d'échantillon) avant de commencer les mesures sur la protéine, pour que l'échantillon soit emmené à la concentration nécessaire. En plus, tout doit être vérifié: la concentration, l'état de stabilité, l'oxydation du DTT. Ceci prévient que les échantillons ne se dégradent pas, surtout pour ceux qui sont moins solubles et ne supportent pas la concentration très élevée.

Ces points (entre autres) sont réellement d'importance capitale, le fait de les avoir comme réflexes garantit la qualité des protéines E6 au bout du périple.

Lyse des bactéries:

Le tampon de lyse des bactéries est composé de : Tris-HCl 50 mM (pH 6,8), 400 mM NaCl, 2 mM DTT, 2,5 µg/ml de DNase, 2,5 µg/ml de RNase, 5% de glycérol, ainsi qu'une tablette pour 50 ml de tampon d'un cocktail sans EDTA d'inhibiteur de protéases

(Roche). En moyenne, un litre de culture bactérienne est repris dans 50 ml de tampon de lyse.

La sonication est effectuée sur des volumes de 25 à 75 ml successifs avec une sonde 50 mm (BRANSON, Digital Sonifier®). Une amplitude de 20 à 50% est appliquée pendant 2 minutes par cycles de 1 sec d'ultrasons et 3 sec de pause. Puis l'échantillon est replacé sur glace pour éviter un échauffement trop important. Cette étape est répétée généralement 3 à 4 fois jusqu'à ce que la solution passe d'un état opaque à une phase trouble mais légèrement translucide. Les extraits bruts sont ensuite ultracentrifugés pendant 1 heure dans des tubes spéciales à 38 000 rpm (80 000 x g au milieu du tube) dans un rotor Ti 70 monté dans une centrifugeuse L-70 (Beckman).

Les surnageants sont filtrés sur des filtres à 0,22 µm, puis déposés sur une colonne amylose, résine spécifique de la protéine MBP (Maltose Binding Protein). La colonne a été préalablement équilibrée dans un tampon composé de : 50 mM Tris-HCl pH 6,8, 400 mM NaCl, 2 DTT (tampon A). La colonne est ensuite lavée avec 8 à 10 volumes de tampon B (A + cocktail d'inhibiteurs de protéases). La quantité d'anti-protéase décroît par pas successifs de 100 à 10 % de la concentration recommandée par le fabricant. Le décrochage de la fusion est assuré par l'ajout de maltose à une concentration de 10 mM.

Elimination des molécules agrégées: l'ultracentrifugation:

Les éluas sont placés dans des tubes UltraClear® de 12,5 ml dans un rotor SW41 Ti à 38 000 rpm (200 000 x g) sur une centrifugeuse L-70 (Beckman) réfrigérée à 4°C. La durée minimale expérimentalement déterminée doit être supérieure ou égale à 15 heures. Par précaution, les 2 ml séparant le surnageant et le culot ne sont pas prélevés avec la solution qui contient notre protéine.

Coupure de la protéine de fusion, séparation de MBP et de la protéine E6:

La digestion à la TEV se déroule entre 8 et 20°C jusqu'à coupure complète (entre 2 et 24 heures selon la température et le stock de la protéase dont l'activité est variable entre chaque préparation). La solution est ensuite concentrée jusqu'à un volume légèrement supérieur à 2 ml puis chargée sur une colonne gel-filtration HiLoad 16/60 Superdex 75 (gamme de séparation entre 3 et 70 kDa) (Amersham Biosciences) préalablement équilibrée avec 250 ml d'un nouveau tampon: 20 mM Sodium Phosphate pH 6,8, 200 mM NaCl, 2 mM DTT. L'élution de la protéine conduit à un pic unique qui révèle l'état monomérique ou un état d'oligomérisation (le cas échéant) d'après la calibration de la colonne. Les échantillons sont ensuite concentrés dans des concentrateurs Vivaspin 15 filter de 5 kDa (VIASCIENCE, 5/15) jusqu'à la concentration voulue pour les mesures.

Mini-purification de la protéine E6:

Le culot d'expression de 25ml de culture de la protéine E6 fusionnée avec MBP ou la GST est repris dans le tampon de lyse (Tris-HCl 50 mM (pH 6,8), 400 mM NaCl, 2 mM DTT, 2,5 µg/ml de DNase, 2,5 µg/ml de RNase, 5% de glycérol, une tablette pour 50 ml de tampon d'un cocktail sans EDTA d'inhibiteur de proteases (Roche)). La sonication se fait dans un tube Eppendorf de 2ml par une petite sonde de 3 mm (Vibracell 72412, Bioblock) pour deux cycles de 2 min chacun, l'amplitude du sonicateur est de 8 – 10 % (1 sec d'ultrasons et 1 sec de repos). Les tubes sont centrifugés dans une centrifugeuse réfrigérée (Eppendorf), 4 C°, 30 min à 13000 rpm. Le surnageant est ensuite incubé pendant 30 min avec 300 µl d'une résine amylose (New England Biolabs) ou bien 200 µl d'une résine de glutathione sepharose (Amersham Biosciences) pré-équilibrée du tampon d'équilibration (Tris-HCl 50 mM (pH 6,8), 400 mM NaCl, 2 mM DTT). L'incubation se déroule sur un agitateur à 4 C°. La résine est lavée par 3 cycles de lavages séparés à chaque fois d'une incubation de 3 min sur l'agitateur (12 ml du tampon d'équilibration avec 10% du cocktail d'inhibiteurs de protéases pour chaque cycle). L'élution de la fusion se fait par l'équivalent de 1,5 volume de résine (le tampon d'élution est le même tampon de lavage avec 20 mM de maltose (résine amylose) ou bien 20 mM de glutathione réduit (résine glutathione); une incubation de 5 min sur la glace avec des petits coups de vortex chaque minute environ. Enfin la résine suspendue avec le tampon d'élution est déposée dans un puits d'une plaque de 96 puits avec des filtres de 0,45 µm (Millipore). La fusion MBP-E6 ou GST-E6 est récupérée par centrifugation des plaques pendant 5 min et gardée au froid (4 °C).

Test de solubilité des protéines E6:

50 µl de la fusion MBP-E6 ou GST-E6 muni-purifiée est incubée dans un tube Eppendorf 1,5 pendant une nuit (15 heures à 8 °C) en présence de 5 µl de la TEV recombinante purifiée. Après le clivage complet entre la protéine E6 et la protéine porteuse, l'échantillon est transféré dans un petit tube de Eppendorf 0,5 et centrifugé 30 min à 13000 rpm. Environ 50 µl de surnageant à récupérer dans nouveau tube Eppendorf 0,5 et le culot suspendu délicatement dans un volume de 50 µl du tampon. Les échantillons (total non coupé (T), surnageant (S) et Culots (C)) sont ensuite mélangés avec du bleu de dépôt, et séparés sur un gel SDS-PAGE ; le gel est ensuite coloré au bleu de coomassie.

Co-expression de la protéine E6 avec les peptides:

Les bactéries compétentes (BL21) sont transformées simultanément avec deux plasmides: un vecteur d'expression (pACYC-11b) qui contient la séquence de E6 non

fusionnée, et un autre vecteur (pETM) portant la séquence de GST-*LXXLL*. Les vecteurs sont introduits dans les bactéries par électroporation. Les précultures et les conditions de la co-expression sont celles d'une expression normale en présence de deux antibiotiques Spectinomycine 34 µg/ml et Kanamycine 15 µg/ml.

Tests de la co-purification de E6 avec les peptides:

La co-purification de E6 avec les peptides *LXXLL* est dépendante de la co-expression, c'est-à-dire que le mélange entre les deux partenaires se fait au niveau des culots d'expression.

Deux types de résines d'affinité sont utilisés:

- co-purification sur résine d'amylose: 20 ml de culot d'expression de la protéine MBP-E6 est repris avec 4 ml de GST-peptide.
- co-purification sur une résine glutathione: 20 ml de culot d'expression de la protéine MBP-E6 est repris avec 40 ml de GST-peptide.

La résuspension se fait dans un tube Eppendorf de 2 ml qui contient le tampon: Tris-HCl 50 mM (pH 6,8), 400 mM NaCl, 2 mM DTT. La sonication se fait dans le même tube par une petite sonde de 3 mm (Vibracell 72412, Bioblock) pour deux cycles de 2 min chacun, l'amplitude du sonicateur est de 8 – 10 % (1 sec d'ultrasons et 1 sec de repos). Les tubes sont centrifugés dans une centrifugeuse réfrigérée (Eppendorf), 4 °C, 30 min à 13000 rpm. La suite de la co-purification proprement dite est comme celle décrite pour la mini-purification.

REFERENCES BIBLIOGRAPHIQUES

REFERENCES BIBLIOGRAPHIQUES

Abbate EA, Berger JM, Botchan MR. The X-ray structure of the papillomavirus helicase in complex with its molecular matchmaker E2. *Genes Dev.* (2004) Aug 15;18(16):1981-96.

Alphs HH, Gambhira R, Karanam B, Roberts JN, Jagu S, Schiller JT, Zeng W, Jackson DC, Roden RB. Protection against heterologous human papillomavirus challenge by a synthetic lipopeptide vaccine containing a broadly cross-neutralizing epitope of L2. *Proc Natl Acad Sci U S A.* (2008) Apr 15;105(15):5850-5.

Androphy EJ, Hubbert NL, Schiller JT, Lowy DR. Identification of the HPV-16 E6 protein from transformed mouse cells and human cervical carcinoma cell lines. *EMBO J.* (1987) Apr;6(4):989-92.

Antinore MJ, Birrer MJ, Patel D, Nader L, McCance DJ. The human papillomavirus type 16 E7 gene product interacts with and trans-activates the AP1 family of transcription factors. *EMBO J.* (1996) Apr 15;15(8):1950-60.

Antson AA, Burns JE, Moroz OV, Scott DJ, Sanders CM, Bronstein IB, Dodson GG, Wilson KS, Maitland NJ. Structure of the intact transactivation domain of the human papillomavirus E2 protein. *Nature.* (2000) Feb 17;403(6771):805-9.

Ashrafi GH, Haghshenas M, Marchetti B, Campo MS. E5 protein of human papillomavirus 16 downregulates HLA class I and interacts with the heavy chain via its first hydrophobic domain. *Int J Cancer.* (2006) Nov 1;119(9):2105-12.

Askew DS, Ashmun RA, Simmons BC, Cleveland JL. Constitutive c-myc expression in an IL-3-dependent myeloid cell line suppresses cell cycle arrest and accelerates apoptosis. *Oncogene.* (1991) Oct;6(10):1915-22.

Baker TS, Newcomb WW, Olson NH, Cowser LM, Olson C, Brown JC. Structures of bovine and human papillomaviruses. Analysis by cryoelectron microscopy and three-dimensional image reconstruction. *Biophys J.* (1991) Dec; 60(6):1445-56.

Banks L, Edmonds C, Vousden KH. Ability of the HPV16 E7 protein to bind RB and induce DNA synthesis is not sufficient for efficient transforming activity in NIH3T3 cells. *Oncogene*. (1990) Sep;5(9):1383-9.

Banks L, Spence P, Androphy E, Hubbert N, Matlashewski G, Murray A, Crawford L. Identification of human papillomavirus type 18 E6 polypeptide in cells derived from human cervical carcinomas. *J Gen Virol*. (1987) May;68 (Pt 5):1351-9.

Barbosa MS, Edmonds C, Fisher C, Schiller JT, Lowy DR, Vousden KH. The region of the HPV E7 oncoprotein homologous to adenovirus E1a and Sv40 large T antigen contains separate domains for Rb binding and casein kinase II phosphorylation. *EMBO J*. (1990) Jan;9(1):153-60.

Barron BA, Richart RM. Screening protocols for cervical neoplastic disease. *Gynecol Oncol*. (1981) Oct;12(2 Pt 2):S156-67.

Ben-Saadon R, Fajerman I, Ziv T, Hellman U, Schwartz AL, Ciechanover A. The tumor suppressor protein p16(INK4a) and the human papillomavirus oncoprotein-58 E7 are naturally occurring lysine-less proteins that are degraded by the ubiquitin system. Direct evidence for ubiquitination at the N-terminal residue. *J Biol Chem*. (2004) Oct 1;279(40):41414-21.

Berezutskaya E, Bagchi S. The human papillomavirus E7 oncoprotein functionally interacts with the S4 subunit of the 26 S proteasome. *J Biol Chem*. (1997) Nov 28;272(48):30135-40.

Bishop B, Dasgupta J, Klein M, Garcea RL, Christensen ND, Zhao R, Chen XS. Crystal structures of four types of human papillomavirus L1 capsid proteins: understanding the specificity of neutralizing monoclonal antibodies. *J Biol Chem*. (2007) Oct 26;282(43):31803-11.

Bodaghi S, Jia R, Zheng ZM. Human papillomavirus type 16 E2 and E6 are RNA-binding proteins and inhibit in vitro splicing of pre-mRNAs with suboptimal splice sites. *Virology*. (2009) Mar 30;386(1):32-43.

Bonnez W, Rose RC, Borkhuis C, Da Rin C, Reichman RC. Evaluation of temperature sensitivity of human papillomavirus type 11 by using the human xenograft severe combined immunodeficiency mouse model. *J Clin Microbiol*. (1994) Jun;32(6):1575-7.

Bosch FX. Epidemiology of human papillomavirus infections: new options for cervical cancer prevention. *Salud Publica Mex.* (2003);45 Suppl 3:S326-39.

Braspenning J, Marchini A, Albarani V, Levy L, Ciccolini F, Cremonesi C, Ralston R, Gissmann L, Tommasino M. The CXXC Zn binding motifs of the human papillomavirus type 16 E7 oncoprotein are not required for its in vitro transforming activity in rodent cells. *Oncogene.* (1998) Feb 26;16(8):1085-9.

Brehm A, Nielsen SJ, Miska EA, McCance DJ, Reid JL, Bannister AJ, Kouzarides T. The E7 oncoprotein associates with Mi2 and histone deacetylase activity to promote cell growth. *EMBO J.* (1999) May 4;18(9):2449-58.

Brown DR, Kitchin D, Qadadri B, Neptune N, Batteiger T, Ermel A. The human papillomavirus type 11 E1--E4 protein is a transglutaminase 3 substrate and induces abnormalities of the cornified cell envelope. *Virology.* (2006) Feb 5;345(1):290-8.

Bubb V, McCance DJ, Schlegel R. DNA sequence of the HPV-16 E5 ORF and the structural conservation of its encoded protein. *Virology.* (1988) Mar;163(1):243-6.

Buck CB, Cheng N, Thompson CD, Lowy DR, Steven AC, Schiller JT, Trus BL. Arrangement of L2 within the papillomavirus capsid. *J Virol.* (2008) Jun;82(11):5190-7.

Burkhardt A, DiMaio D, Schlegel R. Genetic and biochemical definition of the bovine papillomavirus E5 transforming protein. *EMBO J.* (1987) Aug;6(8):2381-5.

Chambers G, Ellsmore VA, O'Brien PM, Reid SW, Love S, Campo MS, Nasir L. Association of bovine papillomavirus with the equine sarcoid. *J Gen Virol.* (2003) May; 84(Pt 5):1055-62. Review.

Chan SY, Bernard HU, Ratterree M, Birkebak TA, Faras AJ, Ostrow RS. Genomic diversity and evolution of papillomaviruses in rhesus monkeys. *J Virol.* 1997 Jul;71(7):4938-43.

Chan SY, Delius H, Halpern AL, Bernard HU. Analysis of genomic sequences of 95 papillomavirus types: uniting typing, phylogeny, and taxonomy. *J Virol.* 1995 May; 69(5):3074-83.

Chen JJ, Reid CE, Band V, Androphy EJ. Interaction of papillomavirus E6 oncoproteins with a putative calcium-binding protein. *Science*. (1995) Jul 28;269(5223):529-31.

Chen XS, Garcea RL, Goldberg I, Casini G, Harrison SC. Structure of small virus-like particles assembled from the L1 protein of human papillomavirus 16. *Mol Cell*. (2000) Mar;5(3):557-67.

Chow LT, Reilly SS, Broker TR, Taichman LB. Identification and mapping of human papillomavirus type 1 RNA transcripts recovered from plantar warts and infected epithelial cell cultures. *J Virol*. (1987) Jun;61(6):1913-8.

Christensen ND, Reed CA, Cladel NM. Induction of neutralizing antibodies to papillomaviruses by anti-idiotypic antibodies. *Virology*. (1995) Jul 10;210(2):292-301.

Cicero DO, Nadra AD, Eliseo T, Dellarole M, Paci M, de Prat-Gay G. Structural and thermodynamic basis for the enhanced transcriptional control by the human papillomavirus strain-16 E2 protein. *Biochemistry*. (2006) May 30;45(21):6551-60.

Clemens KE, Brent R, Gyuris J, Münger K. Dimerization of the human papillomavirus E7 oncoprotein in vivo. *Virology*. (1995) Dec 1;214(1):289-93.

Cole ST, Danos O. Nucleotide sequence and comparative analysis of the human papillomavirus type 18 genome. Phylogeny of papillomaviruses and repeated structure of the E6 and E7 gene products. *J Mol Biol*. (1987) Feb 20;193(4):599-608.

Conrad M, Bubb VJ, Schlegel R. The human papillomavirus type 6 and 16 E5 proteins are membrane-associated proteins which associate with the 16-kilodalton pore-forming protein. *J Virol*. (1993) Oct;67(10):6170-8.

Cooper K, Herrington CS, Evans MF, Gatter KC, McGee JO. p53 antigen in cervical condylomata, intraepithelial neoplasia, and carcinoma: relationship to HPV infection and integration. *J Pathol*. (1993) Sep;171(1):27-34.

Couturier J, Sastre-Garau X, Schneider-Maunoury S, Labib A, Orth G. Integration of papillomavirus DNA near myc genes in genital carcinomas and its consequences for proto-oncogene expression. *J Virol*. (1991) Aug;65(8):4534-8.

- Crook T, Fisher C, Masterson PJ, Vousden KH.** Modulation of transcriptional regulatory properties of p53 by HPV E6. *Oncogene*. (1994) Apr;9(4):1225-30.
- Crook T, Tidy JA, Vousden KH.** Degradation of p53 can be targeted by HPV E6 sequences distinct from those required for p53 binding and trans-activation. *Cell*. (1991) Nov 1;67(3):547-56.
- Cruz MR, Cerqueira DM, Cruz WB, Camara GN, Brigido MM, Silva EO, Carvalho LG, Martins CR.** Prevalence of human papillomavirus type 16 variants in the Federal District, Central Brazil. *Mem Inst Oswaldo Cruz* 2004;99:281-282.
- Daniels PR, Sanders CM, Coulson P, Maitland NJ.** Molecular analysis of the interaction between HPV type 16 E6 and human E6-associated protein. *FEBS Lett*. (1997) Oct 13;416(1):6-10.
- de Villiers EM, Fauquet C, Broker TR, Bernard HU, zur Hausen H.** Classification of papillomaviruses. *Virology*. (2004) Jun 20;324(1):17-27. Review.
- Dell G, Gaston K.** Human papillomaviruses and their role in cervical cancer. *Cell Mol Life Sci*. (2001) Nov;58(12-13):1923-42.
- DeMasi J, Huh KW, Nakatani Y, Münger K, Howley PM.** Bovine papillomavirus E7 transformation function correlates with cellular p600 protein binding. *Proc Natl Acad Sci U S A*. (2005) Aug 9;102(32):11486-91.
- Demeret C, Desaintes C, Yaniv M, Thierry F.** Different mechanisms contribute to the E2-mediated transcriptional repression of human papillomavirus type 18 viral oncogenes. *J Virol*. (1997) Dec;71(12):9343-9.
- Dong G, Broker TR, Chow LT.** Human papillomavirus type 11 E2 proteins repress the homologous E6 promoter by interfering with the binding of host transcription factors to adjacent elements. *J Virol*. (1994) Feb;68(2):1115-27.
- Doorbar J, Campbell D, Grand RJ, Gallimore PH.** Identification of the human papilloma virus-1a E4 gene products. *EMBO J*. (1986) Feb;5(2):355-62.

Doorbar J, Coneron I, Gallimore PH. Sequence divergence yet conserved physical characteristics among the E4 proteins of cutaneous human papillomaviruses. *Virology* (1989) ;172:51-62.

Doorbar J, Ely S, Sterling J, McLean C, Crawford L. Specific interaction between HPV-16 E1-E4 and cytokeratins results in collapse of the epithelial cell intermediate filament network. *Nature*. (1991) Aug 29;352(6338):824-7.

Doorbar J, Gallimore PH. Identification of proteins encoded by the L1 and L2 open reading frames of human papillomavirus 1a. *J Virol*. (1987) Sep;61(9):2793-9.

Doorbar J, Parton A, Hartley K, Banks L, Crook T, Stanley M, Crawford L. Detection of novel splicing patterns in a HPV16-containing keratinocyte cell line. *Virology*. (1990) Sep;178(1):254-62.

Doorbar J. The papillomavirus life cycle. *J Clin Virol* (2005) ;32S:S7-S15.

Doyle DA, Lee A, Lewis J, Kim E, Sheng M, MacKinnon R. Crystal structures of a complexed and peptide-free membrane protein-binding domain: molecular basis of peptide recognition by PDZ. *Cell*. (1996) Jun 28;85(7):1067-76.

Dyson N, Howley PM, Münger K, Harlow E. The human papilloma virus-16 E7 oncoprotein is able to bind to the retinoblastoma gene product. *Science*. (1989) Feb 17;243(4893):934-7.

El Mehdaoui S, Touzé A, Laurent S, Sizaret PY, Rasschaert D, Coursaget P. Gene transfer using recombinant rabbit hemorrhagic disease virus capsids with genetically modified DNA encapsidation capacity by addition of packaging sequences from the L1 or L2 protein of human papillomavirus type 16. *J Virol*. (2000) Nov;74(22):10332-40.

Enemark EJ, Chen G, Vaughn DE, Stenlund A, Joshua-Tor L. Crystal structure of the DNA binding domain of the replication initiation protein E1 from papillomavirus. *Mol Cell*. (2000) Jul;6(1):149-58.

Enemark EJ, Joshua-Tor L. Mechanism of DNA translocation in a replicative hexameric helicase. *Nature*. (2006) Jul 20;442(7100):270-5.

Enemark EJ, Stenlund A, Joshua-Tor L. Crystal structures of two intermediates in the assembly of the papillomavirus replication initiation complex. *EMBO J.* (2002) Mar 15;21(6):1487-96.

Evander M, Frazer IH, Payne E, Qi YM, Hengst K, McMillan NA. Identification of the alpha6 integrin as a candidate receptor for papillomaviruses. *J Virol.* (1997) Mar;71(3):2449-56.

Figge J, Webster T, Smith TF, Paucha E. Prediction of similar transforming regions in simian virus 40 large T, adenovirus E1A, and myc oncoproteins. *J Virol.* (1988) May;62(5):1814-8.

Filippova M, Filippov VA, Kagoda M, Garnett T, Fodor N, Duerksen-Hughes PJ. Complexes of human papillomavirus type 16 E6 proteins form pseudo-death-inducing signaling complex structures during tumor necrosis factor-mediated apoptosis. *J Virol.* (2009) Jan;83(1):210-27.

Filippova M, Song H, Connolly JL, Dermody TS, Duerksen-Hughes PJ. The human papillomavirus 16 E6 protein binds to tumor necrosis factor (TNF) R1 and protects cells from TNF-induced apoptosis. *J Biol Chem.* (2002) Jun 14;277(24):21730-9.

Firzlaff JM, Galloway DA, Eisenman RN, Lüscher B. The E7 protein of human papillomavirus type 16 is phosphorylated by casein kinase II. *New Biol.* (1989) Oct;1(1):44-53.

Flores ER, Allen-Hoffmann BL, Lee D, Lambert PF. The human papillomavirus type 16 E7 oncogene is required for the productive stage of the viral life cycle. *J Virol.* (2000) Jul;74(14):6622-31.

Florin L, Becker KA, Lambert C, Nowak T, Sapp C, Strand D, Streeck RE, Sapp M. Identification of a dynein interacting domain in the papillomavirus minor capsid protein L2. *J Virol.* (2006) Jul;80(13):6691-6.

Fouts ET, Yu X, Egelman EH, Botchan MR. Biochemical and electron microscopic image analysis of the hexameric E1 helicase. *J Biol Chem.* (1999) Feb 12;274(7):4447-58.

Fox JD, Kapust RB, Waugh DS. Single amino acid substitutions on the surface of Escherichia coli maltose-binding protein can have a profound impact on the solubility of fusion proteins. *Protein Sci.* (2001) Mar;10(3):622-30.

Franco EL, Villa LL, Sobrinho JP, Prado JM, Rousseau MC, Désy M, Rohan TE. Epidemiology of acquisition and clearance of cervical human papillomavirus infection in women from a high-risk area for cervical cancer. *J Infect Dis.* (1999) Nov;180(5):1415-23.

Funk JO, Waga S, Harry JB, Espling E, Stillman B, Galloway DA. Inhibition of CDK activity and PCNA-dependent DNA replication by p21 is blocked by interaction with the HPV-16 E7 oncoprotein. *Genes Dev.* (1997) Aug 15;11(16):2090-100.

Gao Q, Kumar A, Singh L, Huibregtse JM, Beaudenon S, Srinivasan S, Wazer DE, Band H, Band V. Human papillomavirus E6-induced degradation of E6TP1 is mediated by E6AP ubiquitin ligase. *Cancer Res.* (2002) Jun 1;62(11):3315-21.

Garcia-Alai MM, Gallo M, Salame M, Wetzler DE, McBride AA, Paci M, Cicero DO, de Prat-Gay G. Molecular basis for phosphorylation-dependent, PEST-mediated protein turnover. *Structure.* (2006) Feb;14(2):309-19.

García-Vallvé S, Alonso A, Bravo IG. Papillomaviruses: different genes have different histories. *Trends Microbiol.* (2005) Nov; 13(11): 514-21

Genther SM, Sterling S, Duensing S, Münger K, Sattler C, Lambert PF. Quantitative role of the human papillomavirus type 16 E5 gene during the productive stage of the viral life cycle. *J Virol.* (2003) Mar;77(5):2832-42.

Giroglou T, Florin L, Schäfer F, Streeck RE, Sapp M. Human papillomavirus infection requires cell surface heparan sulfate. *J Virol.* (2001) Feb;75(3):1565-70.

Gonzalez SL, Stremlau M, He X, Basile JR, Münger K. Degradation of the retinoblastoma tumor suppressor by the human papillomavirus type 16 E7 oncoprotein is important for functional inactivation and is separable from proteasomal degradation of E7. *J Virol.* (2001) Aug;75(16):7583-91.

Gottschling M, Stamatakis A, Nindl I, Stockfleth E, Alonso A, Bravo IG. Multiple evolutionary mechanisms drive papillomavirus diversification. *Mol Biol Evol.* (2007) May; 24(5):1242-58.

Greider CW. Telomerase activation. One step on the road to cancer? *Trends Genet.* (1999) Mar;15(3):109-12.

Groff DE, Lancaster WD. Molecular cloning and nucleotide sequence of deer papillomavirus. *J Virol.* (1985) Oct;56(1):85-91.

Grossman SR, Laimins LA. E6 protein of human papillomavirus type 18 binds zinc. *Oncogene.* (1989) Sep;4(9):1089-93.

Grossman SR, Mora R, Laimins LA. Intracellular localization and DNA-binding properties of human papillomavirus type 18 E6 protein expressed with a baculovirus vector. *J Virol.* (1989) Jan;63(1):366-74.

Gross-Mesilaty S, Reinstein E, Bercovich B, Tobias KE, Schwartz AL, Kahana C, Ciechanover A. Basal and human papillomavirus E6 oncoprotein-induced degradation of Myc proteins by the ubiquitin pathway. *Proc Natl Acad Sci U S A.* (1998) Jul 7;95(14):8058-63.

Harris SF, Botchan MR. Crystal structure of the human papillomavirus type 18 E2 activation domain. *Science.* (1999) Jun 4;284(5420):1673-7.

Hebner CM, Laimins LA. Human papillomaviruses: basic mechanisms of pathogenesis and oncogenicity. *Rev Med Virol.* (2006) Mar-Apr;16(2):83-97.

Hegde RS, Androphy EJ. Crystal structure of the E2 DNA-binding domain from human papillomavirus type 16: implications for its DNA binding-site selection mechanism. *J Mol Biol.* (1998) Dec 18;284(5):1479-89.

Hegde RS, Grossman SR, Laimins LA, Sigler PB. Crystal structure at 1.7 Å of the bovine papillomavirus-1 E2 DNA-binding domain bound to its DNA target. *Nature.* (1992) Oct 8;359(6395):505-12.

Hegde RS. The papillomavirus E2 proteins: structure, function, and biology. *Annu Rev Biophys Biomol Struct.* (2002); 31:343-60.

Hellberg D, Stendahl U. The biological role of smoking, oral contraceptive use and endogenous sexual steroid hormones in invasive squamous epithelial cervical cancer. *Anticancer Res.* (2005) Jul-Aug;25(4):3041-6.

Helt AM, Galloway DA. Destabilization of the retinoblastoma tumor suppressor by human papillomavirus type 16 E7 is not sufficient to overcome cell cycle arrest in human keratinocytes. *J Virol.* (2001) Aug;75(15):6737-47.

Herbst LH, Lenz J, Van Doorslaer K, Chen Z, Stacy BA, Wellehan JF Jr, Manire CA, Burk RD. Genomic characterization of two novel reptilian papillomaviruses, *Chelonia mydas* papillomavirus 1 and *Caretta caretta* papillomavirus 1. *Virology.* 2009 Jan 5;383(1):131-5.

Holt SE, Shay JW, Wright WE. Refining the telomere-telomerase hypothesis of aging and cancer. *Nat Biotechnol.* (1996) Jul;14(7):836-9.

Hooley E, Fairweather V, Clarke AR, Gaston K, Brady RL. The recognition of local DNA conformation by the human papillomavirus type 6 E2 protein. *Nucleic Acids Res.* (2006) ;34(14):3897-908.

Howley PM. Warts, cancer and ubiquitylation: lessons from the papillomaviruses. *Trans Am Clin Climatol Assoc.* (2006) ;117:113-26

Huh K, Zhou X, Hayakawa H, Cho JY, Libermann TA, Jin J, Harper JW, Munger K. Human papillomavirus type 16 E7 oncoprotein associates with the cullin 2 ubiquitin ligase complex, which contributes to degradation of the retinoblastoma tumor suppressor. *J Virol.* (2007) Sep;81(18):9737-47.

Huh KW, DeMasi J, Ogawa H, Nakatani Y, Howley PM, Münger K. Association of the human papillomavirus type 16 E7 oncoprotein with the 600-kDa retinoblastoma protein-associated factor, p600. *Proc Natl Acad Sci U S A.* (2005) Aug 9;102(32):11492-7.

Imai Y, Tsunokawa Y, Sugimura T, Terada M. Purification and DNA-binding properties of human papillomavirus type 16 E6 protein expressed in *Escherichia coli*. *Biochem Biophys Res Commun.* (1989) Nov 15;164(3):1402-10.

Jackson S, Harwood C, Thomas M, Banks L, Storey A. Role of Bak in UV-induced apoptosis in skin cancer and abrogation by HPV E6 proteins. *Genes Dev.* (2000) Dec 1;14(23):3065-73.

Jewers RJ, Hildebrandt P, Ludlow JW, Kell B, McCance DJ. Regions of human papillomavirus type 16 E7 oncoprotein required for immortalization of human keratinocytes. *J Virol.* (1992) Mar;66(3):1329-35.

Joyce JG, Tung JS, Przysiecki CT, Cook JC, Lehman ED, Sands JA, Jansen KU, Keller PM. The L1 major capsid protein of human papillomavirus type 11 recombinant virus-like particles interacts with heparin and cell-surface glycosaminoglycans on human keratinocytes. *J Biol Chem.* (1999) Feb 26;274(9):5810-22.

Kämmer C, Tommasino M, Syrjänen S, Delius H, Hebling U, Warthorst U, Pfister H, Zehbe I. Variants of the long control region and the E6 oncogene in European human papillomavirus type 16 isolates: implications for cervical disease. *Br J Cancer.* (2002) Jan 21;86(2):269-73.

Kanaya T, Kyo S, Laimins LA. The 5' region of the human papillomavirus type 31 upstream regulatory region acts as an enhancer which augments viral early expression through the action of YY1. *Virology.* (1997) Oct 13;237(1):159-69.

Kanda T, Watanabe S, Zanma S, Sato H, Furuno A, Yoshiike K. Human papillomavirus type 16 E6 proteins with glycine substitution for cysteine in the metal-binding motif. *Virology.* (1991) Dec;185(2):536-43.

Kapust RB, Waugh DS. Escherichia coli maltose-binding protein is uncommonly effective at promoting the solubility of polypeptides to which it is fused. *Protein Sci.* (1999) Aug;8(8):1668-74.

Kirnbauer R, Booy F, Cheng N, Lowy DR, Schiller JT. Papillomavirus L1 major capsid protein self-assembles into virus-like particles that are highly immunogenic. *Proc Natl Acad Sci U S A.* (1992) Dec 15; 89(24):12180-4.

Kisseljov FL. Virus-associated human tumors: cervical carcinomas and papillomaviruses. *Biochemistry (Mosc).* (2000) Jan;65(1):68-77.

Kiyono T, Hiraiwa A, Fujita M, Hayashi Y, Akiyama T, Ishibashi M. Binding of high-risk human papillomavirus E6 oncoproteins to the human homologue of the Drosophila discs large tumor suppressor protein. *Proc Natl Acad Sci U S A.* (1997) Oct 14;94(21):11612-6.

Kjellberg L, Hallmans G, Ahren AM, Johansson R, Bergman F, Wadell G, Angström T, Dillner J. Smoking, diet, pregnancy and oral contraceptive use as risk factors for cervical intra-epithelial neoplasia in relation to human papillomavirus infection. *Br J Cancer.* (2000) Apr;82(7):1332-8.

Klingelhutz AJ, Foster SA, McDougall JK. Telomerase activation by the E6 gene product of human papillomavirus type 16. *Nature*. (1996) Mar 7;380(6569):79-82.

Knight GL, Grainger JR, Gallimore PH, Roberts S. Cooperation between different forms of the human papillomavirus type 1 E4 protein to block cell cycle progression and cellular DNA synthesis. *J Virol*. (2004) Dec;78(24):13920-33.

Koushik A, Ghosh A, Duarte-Franco E, Forest P, Voyer H, Matlashewski G, Coutlée F, Franco EL; Biomarkers of Cervical Cancer Risk (BCCR) Study Team. The p53 codon 72 polymorphism and risk of high-grade cervical intraepithelial neoplasia. *Cancer Detect Prev*. (2005) 29(4):307-16.

Krawczyk E, Supryniewicz FA, Liu X, Dai Y, Hartmann DP, Hanover J, Schlegel R. Koilocytosis: a cooperative interaction between the human papillomavirus E5 and E6 oncoproteins. *Am J Pathol*. (2008) Sep;173(3):682-8.

Kumar S, Hedges SB. A molecular timescale for vertebrate evolution. *Nature*. (1998) Apr 30;392(6679):917-20.

Le Roux LG, Moroianu J. Nuclear entry of high-risk human papillomavirus type 16 E6 oncoprotein occurs via several pathways. *J Virol*. 2003 Feb;77(4):2330-7.

Lechner MS, Laimins LA. Inhibition of p53 DNA binding by human papillomavirus E6 proteins. *J Virol*. (1994) Jul;68(7):4262-73.

Lee JE, Kim CY, Giaccia AJ, Giffard RG. The E6 and E7 genes of human papilloma virus-type 16 protect primary astrocyte cultures from injury. *Brain Res*. (1998) Jun 8;795(1-2):10-6.

Lee SS, Glaunsinger B, Mantovani F, Banks L, Javier RT. Multi-PDZ domain protein MUPP1 is a cellular target for both adenovirus E4-ORF1 and high-risk papillomavirus type 18 E6 oncoproteins. *J Virol*. (2000) Oct;74(20):9680-93.

Li X, Coffino P. High-risk human papillomavirus E6 protein has two distinct binding sites within p53, of which only one determines degradation. *J Virol*. (1996) Jul;70(7):4509-16.

Liang H, Petros AM, Meadows RP, Yoon HS, Egan DA, Walter K, Holzman TF, Robins T, Fesik SW. Solution structure of the DNA-binding domain of a human papillomavirus E2

protein: evidence for flexible DNA-binding regions. *Biochemistry*. (1996) Feb 20;35(7):2095-103.

Lie AK, Skarsvag S, Skomedal H, Haugen OA, Holm R. Expression of p53, MDM2, and p21 proteins in high-grade cervical intraepithelial neoplasia and relationship to human papillomavirus infection. *Int J Gynecol Pathol*. (1999) Jan;18(1):5-11.

Lin BY, Makhov AM, Griffith JD, Broker TR, Chow LT. Chaperone proteins abrogate inhibition of the human papillomavirus (HPV) E1 replicative helicase by the HPV E2 protein. *Mol Cell Biol*. (2002) Sep;22(18):6592-604.

Lipari F, McGibbon GA, Wardrop E, Cordingley MG. Purification and biophysical characterization of a minimal functional domain and of an N-terminal Zn²⁺-binding fragment from the human papillomavirus type 16 E6 protein. *Biochemistry*. (2001) Feb 6;40(5):1196-204.

Liu JS, Kuo SR, Makhov AM, Cyr DM, Griffith JD, Broker TR, Chow LT. Human Hsp70 and Hsp40 chaperone proteins facilitate human papillomavirus-11 E1 protein binding to the origin and stimulate cell-free DNA replication. *J Biol Chem*. (1998) Nov 13;273(46):30704-12.

Longworth MS, Laimins LA. Pathogenesis of human papillomaviruses in differentiating epithelia. *Microbiol Mol Biol Rev*. (2004) Jun;68(2):362-72. Review.

Majewski S, Jablonska S. Possible involvement of epidermodysplasia verruciformis human papillomaviruses in the immunopathogenesis of psoriasis: a proposed hypothesis. *Exp Dermatol*. (2003) Dec;12(6):721-8.

Mansky KC, Batiza A, Lambert PF. Bovine papillomavirus type 1 E1 and simian virus 40 large T antigen share regions of sequence similarity required for multiple functions. *J Virol*. (1997) Oct;71(10):7600-8.

Massimi P, Pim D, Banks L. Human papillomavirus type 16 E7 binds to the conserved carboxy-terminal region of the TATA box binding protein and this contributes to E7 transforming activity. *J Gen Virol*. (1997) Oct;78 (Pt 10):2607-13.

Masson M, Hindelang C, Sibler AP, Schwalbach G, Travé G, Weiss E. Preferential nuclear localization of the human papillomavirus type 16 E6 oncoprotein in cervical carcinoma cells. *J Gen Virol*. (2003) Aug;84(Pt 8):2099-104.

Matlashewski GJ, Tuck S, Pim D, Lamb P, Schneider J, Crawford LV. Primary structure polymorphism at amino acid residue 72 of human p53. *Mol Cell Biol.* (1987) Feb;7(2):961-3.

Matorras R, Ariceta JM, Rementeria A, Corral J, Gutierrez de Terán G, Diez J, Montoya F, Rodriguez-Escudero FJ. Human immunodeficiency virus-induced immunosuppression: a risk factor for human papillomavirus infection. *Am J Obstet Gynecol.* (1991) Jan;164(1 Pt 1):42-4.

Mavromatis KO, Jones DL, Mukherjee R, Yee C, Grace M, Münger K. The carboxyl-terminal zinc-binding domain of the human papillomavirus E7 protein can be functionally replaced by the homologous sequences of the E6 protein. *Virus Res.* (1997) Nov;52(1):109-18.

McBride AA, Byrne JC, Howley PM. E2 polypeptides encoded by bovine papillomavirus type 1 form dimers through the common carboxyl-terminal domain: transactivation is mediated by the conserved amino-terminal domain. *Proc Natl Acad Sci U S A.* (1989) Jan;86(2):510-4.

McBride AA, Schlegel R, Howley PM. The carboxy-terminal domain shared by the bovine papillomavirus E2 transactivator and repressor proteins contains a specific DNA binding activity. *EMBO J.* (1988) Feb;7(2):533-9.

McLaughlin-Drubin M, Münger K. Viruses associated with human cancer. *Biochim Biophys Acta.* (2008) Mar;1782(3):127-50.

McLaughlin-Drubin ME, Huh KW, Münger K. Human papillomavirus type 16 E7 oncoprotein associates with E2F6. *J Virol.* (2008) Sep;82(17):8695-705.

Melillo RM, Helin K, Lowy DR, Schiller JT. Positive and negative regulation of cell proliferation by E2F-1: influence of protein level and human papillomavirus oncoproteins. *Mol Cell Biol.* (1994) Dec;14(12):8241-9.

Modis Y, Trus BL, Harrison SC. Atomic model of the papillomavirus capsid. *EMBO J.* (2002) Sep 16;21(18):4754-62.

Morris JD, Crook T, Bandara LR, Davies R, LaThangue NB, Vousden KH. Human papillomavirus type 16 E7 regulates E2F and contributes to mitogenic signalling. *Oncogene.* (1993) Apr;8(4):893-8.

Morris PJ, Dent CL, Ring CJ, Latchman DS. The octamer binding site in the HPV16 regulatory region produces opposite effects on gene expression in cervical and non-cervical cells. *Nucleic Acids Res.* (1993) Feb 25;21(4):1019-23.

Moscicki AB, Hills N, Shiboski S, Powell K, Jay N, Hanson E, Miller S, Clayton L, Farhat S, Broering J, Darragh T, Palefsky J. Risks for incident human papillomavirus infection and low-grade squamous intraepithelial lesion development in young females. *JAMA.* (2001) Jun 20;285(23):2995-3002.

Münger K, Basile JR, Duensing S, Eichten A, Gonzalez SL, Grace M, Zacny VL. Biological activities and molecular targets of the human papillomavirus E7 oncoprotein. *Oncogene.* (2001) Nov 26;20(54):7888-98.

Munoz N, Castellsagué X, de González AB, Gissmann L. HPV in the etiology of human cancer. *Vaccine* (2006) Aug 31;24 Suppl 3:S3/1-10.

Munoz N, Méndez F, Posso H, Molano M, van den Brule AJ, Ronderos M, Meijer C, Munoz A; Instituto Nacional de Cancerologia HPV Study Group. Incidence, duration, and determinants of cervical human papillomavirus infection in a cohort of Colombian women with normal cytological results. *J Infect Dis.* (2004) Dec 15;190(12):2077-87.

Nakagawa S, Huibregtse JM. Human scribble (Vartul) is targeted for ubiquitin-mediated degradation by the high-risk papillomavirus E6 proteins and the E6AP ubiquitin-protein ligase. *Mol Cell Biol.* (2000) Nov;20(21):8244-53.

Nelson LM, Rose RC, Moroianu J. Nuclear import strategies of high risk HPV16 L1 major capsid protein. *J Biol Chem.* (2002) Jun 28;277(26):23958-64.

Nguyen MM, Nguyen ML, Caruana G, Bernstein A, Lambert PF, Griep AE. Requirement of PDZ-containing proteins for cell cycle regulation and differentiation in the mouse lens epithelium. *Mol Cell Biol.* (2003) Dec;23(24):8970-81.

Nominé Y, Charbonnier S, Miguet L, Potier N, Van Dorsselaer A, Atkinson RA, Travé G, Kieffer B. ¹H and ¹⁵N resonance assignment, secondary structure and dynamic behaviour of the C-terminal domain of human papillomavirus oncoprotein E6. *J Biomol NMR.* (2005) Feb;31(2):129-41.

Nominé Y, Charbonnier S, Ristriani T, Stier G, Masson M, Cavusoglu N, Van Dorselaer A, Weiss E, Kieffer B, Travé G. Domain substructure of HPV E6 oncoprotein: biophysical characterization of the E6 C-terminal DNA-binding domain. *Biochemistry*. (2003) May 6;42(17):4909-17.

Nominé Y, Masson M, Charbonnier S, Zanier K, Ristriani T, Deryckère F, Sibler AP, Desplancq D, Atkinson RA, Weiss E, Orfanoudakis G, Kieffer B, Travé G. Structural and functional analysis of E6 oncoprotein: insights in the molecular pathways of human papillomavirus-mediated pathogenesis. *Mol Cell*. (2006) Mar 3;21(5):665-78.

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)b Apr;14(4):297-305.

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)a Oct;23(1):22-32.

Oetke C, Auvinen E, Pawlita M, Alonso A. Human papillomavirus type 16 E5 protein localizes to the Golgi apparatus but does not grossly affect cellular glycosylation. *Arch Virol*. (2000) 145(10):2183-91.

Oh ST, Longworth MS, Laimins LA. Roles of the E6 and E7 proteins in the life cycle of low-risk human papillomavirus type 11. *J Virol*. (2004) Mar;78(5):2620-6.

Ohlenschläger O, Seiboth T, Zengerling H, Briese L, Marchanka A, Ramachandran R, Baum M, Korbass M, Meyer-Klaucke W, Dürst M, Görlach M. Solution structure of the partially folded high-risk human papilloma virus 45 oncoprotein E7. *Oncogene*. (2006) Sep 28;25(44):5953-9.

Oliveira S, Sousa H, Santos AM, Pinto D, Pinto-Correia AL, Fontoura D, Moutinho J, Medeiros R. The p53 R72P polymorphism does not influence cervical cancer development in a Portuguese population: a study in exfoliated cervical cells. *J Med Virol*. (2008) Mar;80(3):424-9.

Orth G, Favre M, Croissant O. Characterization of a new type of human papillomavirus that causes skin warts. *J Virol*. (1977) Oct;24(1):108-20.

Orth G, Favre M, Majewski S, Jablonska S. Epidermodysplasia verruciformis defines a subset of cutaneous human papillomaviruses. *J Virol.* (2001) May;75(10):4952-3.

Park DS, Selvey LA, Kelsall SR, Frazer IH. Human papillomavirus type 16 E6, E7 and L1 and type 18 E7 proteins produced by recombinant baculoviruses. *J Virol Methods.* (1993) Dec 31;45(3):303-18.

Park JS, Kim EJ, Kwon HJ, Hwang ES, Namkoong SE, Um SJ. Inactivation of interferon regulatory factor-1 tumor suppressor protein by HPV E7 oncoprotein. Implication for the E7-mediated immune evasion mechanism in cervical carcinogenesis. *J Biol Chem.* (2000) Mar 10;275(10):6764-9.

Parkin DM, Bray F, Ferlay J, Pisani P. Global cancer statistics, 2002. *CA Cancer J Clin.* (2005) Mar-Apr;55(2):74-108.

Penrose KJ, Garcia-Alai M, de Prat-Gay G, McBride AA. Casein Kinase II phosphorylation-induced conformational switch triggers degradation of the papillomavirus E2 protein. *J Biol Chem.* (2004) May 21;279(21):22430-9.

Penrose KJ, McBride AA. Proteasome-mediated degradation of the papillomavirus E2-TA protein is regulated by phosphorylation and can modulate viral genome copy number. *J Virol.* (2000) Jul;74(13):6031-8.

Pfister H. Chapter 8: Human papillomavirus and skin cancer. *J Natl Cancer Inst Monogr.* (2003) ;(31):52-6.

Phelps WC, Münger K, Yee CL, Barnes JA, Howley PM. Structure-function analysis of the human papillomavirus type 16 E7 oncoprotein. *J Virol.* (1992) Apr;66(4):2418-27.

Phelps WC, Yee CL, Münger K, Howley PM. The human papillomavirus type 16 E7 gene encodes transactivation and transformation functions similar to those of adenovirus E1A. *Cell.* (1988) May 20;53(4):539-47.

Piersma SJ, Welters MJ, van der Hulst JM, Kloth JN, Kwappenberg KM, Trimbos BJ, Melief CJ, Hellebrekers BW, Fleuren GJ, Kenter GG, Offringa R, van der Burg SH. Human papilloma virus specific T cells infiltrating cervical cancer and draining lymph nodes show remarkably frequent use of HLA-DQ and -DP as a restriction element. *Int J Cancer.* (2008) Feb 1;122(3):486-94.

Pim D, Banks L. HPV-18 E6*1 protein modulates the E6-directed degradation of p53 by binding to full-length HPV-18 E6. *Oncogene*. (1999) Dec 9;18(52):7403-8.

Pim D, Massimi P, Banks L. Alternatively spliced HPV-18 E6* protein inhibits E6 mediated degradation of p53 and suppresses transformed cell growth. *Oncogene*. (1997) Jul 17;15(3):257-64.

Pinto LA, Edwards J, Castle PE, Harro CD, Lowy DR, Schiller JT, Wallace D, Kopp W, Adelsberger JW, Baseler MW, Berzofsky JA, Hildesheim A. Cellular immune responses to human papillomavirus (HPV)-16 L1 in healthy volunteers immunized with recombinant HPV-16 L1 virus-like particles. *J Infect Dis*. (2003) Jul 15;188(2):327-38.

Plug-DeMaggio AW, Sundsvold T, Wurscher MA, Koop JI, Klingelutz AJ, McDougall JK. Telomere erosion and chromosomal instability in cells expressing the HPV oncogene 16E6. *Oncogene*. (2004) Apr 29;23(20):3561-71.

Radulescu RT, Bellitti MR, Ruvo M, Cassani G, Fassina G. Binding of the LXCXE insulin motif to a hexapeptide derived from retinoblastoma protein. *Biochem Biophys Res Commun*. (1995) Jan 5;206(1):97-102.

Rapp L, Chen JJ. The papillomavirus E6 proteins. *Biochim Biophys Acta*. (1998) Aug 19;1378(1):F1-19.

Ristriani T, Fournane S, Orfanoudakis G, Travé G, Masson M. A single-codon mutation converts HPV16 E6 oncoprotein into a potential tumor suppressor, which induces p53-dependent senescence of HPV-positive HeLa cervical cancer cells. *Oncogene*. (2009) Feb 5;28(5):762-72.

Ristriani T, Masson M, Nominé Y, Laurent C, Lefevre JF, Weiss E, Travé G. HPV oncoprotein E6 is a structure-dependent DNA-binding protein that recognizes four-way junctions. *J Mol Biol*. (2000) Mar 10;296(5):1189-203.

Ristriani T, Nominé Y, Masson M, Weiss E, Travé G. Specific recognition of four-way DNA junctions by the C-terminal zinc-binding domain of HPV oncoprotein E6. *J Mol Biol*. (2001) Jan 26;305(4):729-39.

Roden RB, Greenstone HL, Kirnbauer R, Booy FP, Jessie J, Lowy DR, Schiller JT. In vitro generation and type-specific neutralization of a human papillomavirus type 16 virion pseudotype. *J Virol.* (1996) Sep;70(9):5875-83.

Romanczuk H, Howley PM. Disruption of either the E1 or the E2 regulatory gene of human papillomavirus type 16 increases viral immortalization capacity. *Proc Natl Acad Sci U S A.* (1992) Apr 1;89(7):3159-63.

Rommel O, Dillner J, Fligge C, Bergsdorf C, Wang X, Selinka HC, Sapp M. Heparan sulfate proteoglycans interact exclusively with conformationally intact HPV L1 assemblies: basis for a virus-like particle ELISA. *J Med Virol.* (2005) Jan;75(1):114-21.

Sastre-Garau X, Couturier J, Favre M, Orth G. A recurrent human papillomavirus integration site at chromosome region 12q14-q15 in SW756 and SK-v cell lines derived from genital tumors. *C R Acad Sci III.* (1995) Apr;318(4):475-8.

Schlegel R, Wade-Glass M, Rabson MS, Yang YC. The E5 transforming gene of bovine papillomavirus encodes a small, hydrophobic polypeptide. *Science.* (1986) Jul 25;233(4762):464-7.

Schuck S, Stenlund A. Role of papillomavirus E1 initiator dimerization in DNA replication. *J Virol.* (2005) Jul;79(13):8661-4.

Schwarz SE, Rosa JL, Scheffner M. Characterization of human hect domain family members and their interaction with Ubch5 and Ubch7. *J Biol Chem.* (1998) May 15;273(20):12148-54.

Selinka HC, Giroglou T, Sapp M. Analysis of the infectious entry pathway of human papillomavirus type 33 pseudovirions. *Virology.* (2002) Aug 1;299(2):279-287.

Seo YS, Müller F, Lusky M, Gibbs E, Kim HY, Phillips B, Hurwitz J. Bovine papilloma virus (BPV)-encoded E2 protein enhances binding of E1 protein to the BPV replication origin. *Proc Natl Acad Sci U S A.* (1993) Apr 1;90(7):2865-9. Erratum in: *Proc Natl Acad Sci U S A* (1993) Jun 15;90(12):5878.

Shafti-Keramat S, Handisurya A, Kriehuber E, Meneguzzi G, Slupetzky K, Kirnbauer R. Different heparan sulfate proteoglycans serve as cellular receptors for human papillomaviruses. *J Virol.* (2003) Dec;77(24):13125-35.

Sherman L, Itzhaki H, Jackman A, Chen JJ, Koval D, Schlegel R. Inhibition of serum- and calcium-induced terminal differentiation of human keratinocytes by HPV 16 E6: study of the association with p53 degradation, inhibition of p53 transactivation, and binding to E6BP. *Virology*. (2002) Jan 20;292(2):309-20.

Sibbet G, Romero-Graillet C, Meneguzzi G, Saveria Campo M. alpha6 integrin is not the obligatory cell receptor for bovine papillomavirus type 4. *J Gen Virol*. (2000) Jun;81 Pt 6:1629.

Singer A. Proceedings: Promiscuous sexual behavior and its effect on cervical morphology. *Br J Obstet Gynaecol*. (1975) Jul;82(7):588-9.

Sparkowski J, Mense M, Anders J, Schlegel R. E5 oncoprotein transmembrane mutants dissociate fibroblast transforming activity from 16-kilodalton protein binding and platelet-derived growth factor receptor binding and phosphorylation. *J Virol*. (1996) Apr;70(4):2420-30.

Stacey SN, Bartholomew JS, Ghosh A, Stern PL, Mackett M, Arrand JR. Expression of human papillomavirus type 16 E6 protein by recombinant baculovirus and use for detection of anti-E6 antibodies in human sera. *J Gen Virol*. (1992) Sep;73 (Pt 9):2337-45.

Steenbergen RD, de Wilde J, Wilting SM, Brink AA, Snijders PJ, Meijer CJ. HPV-mediated transformation of the anogenital tract. *J Clin Virol*. (2005) Mar;32 Suppl 1:S25-33.

Storey A, Thomas M, Kalita A, Harwood C, Gardiol D, Mantovani F, Breuer J, Leigh IM, Matlashewski G, Banks L. Role of a p53 polymorphism in the development of human papillomavirus-associated cancer. *Nature*. (1998) May 21;393(6682):229-34.

Straight SW, Hinkle PM, Jewers RJ, McCance DJ. The E5 oncoprotein of human papillomavirus type 16 transforms fibroblasts and effects the downregulation of the epidermal growth factor receptor in keratinocytes. *J Virol*. (1993) Aug;67(8):4521-32.

Stubenrauch F, Laimins LA. Human papillomavirus life cycle: active and latent phases. *Semin Cancer Biol*. (1999) Dec;9(6):379-86.

Thomas M, Banks L. Human papillomavirus (HPV) E6 interactions with Bak are conserved amongst E6 proteins from high and low risk HPV types. *J Gen Virol*. (1999) Jun;80 (Pt 6):1513-7.

Thomas M, Laura R, Hepner K, Guccione E, Sawyers C, Lasky L, Banks L. Oncogenic human papillomavirus E6 proteins target the MAGI-2 and MAGI-3 proteins for degradation. *Oncogene*. (2002) Aug 1;21(33):5088-96.

Thomas M, Massimi P, Banks L. HPV-18 E6 inhibits p53 DNA binding activity regardless of the oligomeric state of p53 or the exact p53 recognition sequence. *Oncogene*. (1996) Aug 1;13(3):471-80.

Thorland EC, Myers SL, Persing DH, Sarkar G, McGovern RM, Gostout BS, Smith DI. Human papillomavirus type 16 integrations in cervical tumors frequently occur in common fragile sites. *Cancer Res*. (2000) Nov 1;60(21):5916-21.

Tong X, Howley PM. The bovine papillomavirus E6 oncoprotein interacts with paxillin and disrupts the actin cytoskeleton. *Proc Natl Acad Sci U S A*. (1997) Apr 29;94(9):4412-7.

Touze A, El Mehdaoui S, Sizaret PY, Mouglin C, Munoz N, Coursaget P. The L1 major capsid protein of human papillomavirus type 16 variants affects yield of virus-like particles produced in an insect cell expression system. *J Clin Microbiol*. (1998) Jul;36(7):2046-51.

Tungteakkhun SS, Duerksen-Hughes PJ. Cellular binding partners of the human papillomavirus E6 protein. *Arch Virol*. (2008) ;153(3):397-408.

Ullman CG, Haris PI, Galloway DA, Emery VC, Perkins SJ. Predicted alpha-helix/beta-sheet secondary structures for the zinc-binding motifs of human papillomavirus E7 and E6 proteins by consensus prediction averaging and spectroscopic studies of E7. *Biochem J*. (1996) Oct 1;319 (Pt 1):229-39.

Valle GF, Banks L. The human papillomavirus (HPV)-6 and HPV-16 E5 proteins co-operate with HPV-16 E7 in the transformation of primary rodent cells. *J Gen Virol*. (1995) May;76 (Pt 5):1239-45.

Van Ranst M, Kaplan JB, Burk RD. Phylogenetic classification of human papillomaviruses: correlation with clinical manifestations. *J Gen Virol*. (1992) Oct;73 :2653-60.

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) Jan 8;16(1):43-52.

Villa LL, Costa RL, Petta CA, Andrade RP, Ault KA, Giuliano AR, Wheeler CM, Koutsky LA, Malm C, Lehtinen M, Skjeldestad FE, Olsson SE, Steinwall M, Brown DR, Kurman RJ, Ronnett BM, Stoler MH, Ferenczy A, Harper DM, Tamms GM, Yu J, Lupinacci L, Railkar R, Taddeo FJ, Jansen KU, Esser MT, Sings HL, Saah AJ, Barr E. Prophylactic quadrivalent human papillomavirus (types 6, 11, 16, and 18) L1 virus-like particle vaccine in young women: a randomised double-blind placebo-controlled multicentre phase II efficacy trial. *Lancet Oncol.* (2005) May;6(5):271-8.

Vousden KH, Doniger J, DiPaolo JA, Lowy DR. The E7 open reading frame of human papillomavirus type 16 encodes a transforming gene. *Oncogene Res.* (1988) Sep;3(2):167-75.

Walboomers JM, Jacobs MV, Manos MM, Bosch FX, Kummer JA, Shah KV, Snijders PJ, Peto J, Meijer CJ, Munoz N. Human papillomavirus is a necessary cause of invasive cervical cancer worldwide. *J Pathol.* (1999) Sep;189(1):12-9.

Wang Q, Griffin H, Southern S, Jackson D, Martin A, McIntosh P, Davy C, Masterson PJ, Walker PA, Laskey P, Omary MB, Doorbar J. Functional analysis of the human papillomavirus type 16 E1=E4 protein provides a mechanism for in vivo and in vitro keratin filament reorganization. *J Virol.* (2004) Jan;78(2):821-33.

Weintraub SJ, Chow KN, Luo RX, Zhang SH, He S, Dean DC. Mechanism of active transcriptional repression by the retinoblastoma protein. *Nature.* (1995) Jun 29;375(6534):812-5.

Wentzensen N, Vinokurova S, von Knebel Doeberitz M. Systematic review of genomic integration sites of human papillomavirus genomes in epithelial dysplasia and invasive cancer of the female lower genital tract. *Cancer Res.* (2004) Jun 1;64(11):3878-84.

Werness BA, Levine AJ, Howley PM. Association of human papillomavirus types 16 and 18 E6 proteins with p53. *Science.* (1990) Apr 6;248(4951):76-9.

White PW, Pelletier A, Brault K, Titolo S, Welchner E, Thauvette L, Fazekas M, Cordingley MG, Archambault J. Characterization of recombinant HPV6 and 11 E1 helicases: effect of ATP on the interaction of E1 with E2 and mapping of a minimal helicase domain. *J Biol Chem.* (2001) Jun 22;276(25):22426-38.

Yang L, Mohr I, Fouts E, Lim DA, Nohaile M, Botchan M. The E1 protein of bovine papilloma virus 1 is an ATP-dependent DNA helicase. *Proc Natl Acad Sci U S A.* (1993) Jun 1;90(11):5086-90.

Yang R, Yutzy WH 4th, Viscidi RP, Roden RB. Interaction of L2 with beta-actin directs intracellular transport of papillomavirus and infection. *J Biol Chem.* (2003) Apr 4;278(14):12546-53.

You J, Croyle JL, Nishimura A, Ozato K, Howley PM. Interaction of the bovine papillomavirus E2 protein with Brd4 tethers the viral DNA to host mitotic chromosomes. *Cell.* (2004) Apr 30;117(3):349-60.

Zhang Y, Appleton BA, Wu P, Wiesmann C, Sidhu SS. Structural and functional analysis of the ligand specificity of the HtrA2/Omi PDZ domain. *Protein Sci.* (2007) Aug;16(8):1738-50.

Zheng ZM, Baker CC. Papillomavirus genome structure, expression, and post-transcriptional regulation. *Front Biosci.* 2006 Sep 1;11:2286-302.

Zhou J, Sun XY, Louis K, Frazer IH. Interaction of human papillomavirus (HPV) type 16 capsid proteins with HPV DNA requires an intact L2 N-terminal sequence. *J Virol.* (1994) Feb;68(2):619-25.

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) Aug;73(8):6209-19.

zur Hausen H. Human papillomaviruses and their possible role in squamous cell carcinomas. *Curr Top Microbiol Immunol.* (1977); 78:1-30.

zur Hausen H. Papillomaviruses and cancer: from basic studies to clinical application. *Nat Rev Cancer.* (2002) May;2(5):342-50.

zur Hausen H. Papillomaviruses in human cancers. *Proc Assoc Am Physicians.* (1999) Nov-Dec;111(6):581-7.

Zwerschke W, Jansen-Dürr P. Cell transformation by the E7 oncoprotein of human papillomavirus type 16: interactions with nuclear and cytoplasmic target proteins. *Adv Cancer Res.* (2000);78:1-29.

TITRE : Production, analyse structurale et caractérisation fonctionnelle d'oncoprotéines E6 issues de différents papillomavirus.

RÉSUMÉ : Autrefois rattachés aux *papovaviridae* avec les polyomavirus, les papillomavirus sont individualisés depuis peu et forment l'unique genre de la famille *papillomaviridae*. Les papillomavirus (PVs) induisent chez les mammifères et les oiseaux des lésions prolifératives dans les tissus épithéliaux. La plupart de ces lésions sont bénignes (verrues et condylomes) mais elles dégénèrent parfois en cancers. Les papillomavirus humains (HPV) génitaux à haut risque sont responsables de 99% des cas du cancer du col de l'utérus, deuxième cause de mortalité chez la femme. Certains HPV cutanés à haut risque provoquent également des cancers de la peau. Les oncoprotéines E6 et E7 sont particulièrement impliquées dans le processus de carcinogenèse, en interagissant notamment avec les protéines anti-tumorales p53 et pRb. Cependant, E6 et E7 lient et souvent provoquent la dégradation d'environ 100 autres protéines cellulaires régulant elles-mêmes l'expression, la stabilité ou les modifications post-transcriptionnelles d'autres protéines. La structure de E6 a longtemps résisté à toute analyse, car la protéine était très difficile à produire par voie recombinante sous une forme repliée. Notre équipe a travaillé sur ce problème depuis 1995.

Les travaux que j'ai menés durant cette thèse ont permis de compléter l'étude structurale et fonctionnelle de la protéine E6 du HPV et d'aborder des études comparables pour ses orthologues de différents virus animaux. De telles données constituent un progrès considérable attendu depuis une vingtaine d'années. L'optimisation des protocoles, l'exploration des domaines, la mutagenèse dirigée et la solubilisation des protéines à l'aide de leurs ligands, se sont avérés essentiels pour aboutir à l'étude structurale. Remarquablement, pour produire et solubiliser des protéines relativement proches, il m'a fallu adopter différentes stratégies. D'autre part, la maîtrise de la production et de la qualité de ces protéines m'a permis d'étudier le rôle biologique de certains phénomènes (dimérisation, liaison à *LXXLL* ...), par une approche de mutagenèse respectueuse du bon repliement des protéines.

DISCIPLINE : Biochimie, Biologie Moléculaire

MOTS CLEFS : HPV, oncoprotéine E6, cancer du col de l'utérus, purification de protéines difficiles, solubilité des protéines, interaction protéine/ligand, phylogénie des PVs.

INTITULÉ ET ADRESSE DU LABORATOIRE : IREBS, FRE 3211, Équipe Oncoprotéines, ESBS, 1, Bd Sébastien Brandt ; BP 10413 ; 67412 Illkirch Cedex