

---

ÉCOLE DOCTORALE SCIENCE DE LA VIE & DE LA SANTÉ

IREBS, UMR CNRS 7242

## THÈSE

Présentée par

**Khaled OULD BABAH**

soutenue le : 21 septembre 2012

Pour l'obtention du grade de

**DOCTEUR DE L'UNIVERSITE DE STRASBOURG**

Spécialité : Biophysique et biologie structurale

---

---

### **Analyse biochimique et structurale des interactions multiples des oncoprotéines E6 produites par les papillomavirus**

---

Devant le jury académique

#### **JURY**

M. Gilles TRAVÉ	Dr, IREBS, Strasbourg	Directeur de thèse
M. André ZAPUN	Dr, Grenoble	Rapporteur
M. Yves JACOB	Dr, I. Pasteur, Paris	Rapporteur
M. Ahmed HOUMEIDA	Dr, FST, Nouakchott	Examinateur
M. Marc RUFF	Dr, IGBMC, Strasbourg	Examinateur

**Mr Khaled OULD BABAH** a été membre de la promotion **Rosa Parks** du **Collège doctoral européen** de l'Université de Strasbourg pendant la préparation de sa thèse de **2008 à 2012**. Il a suivi un enseignement hebdomadaire sur les affaires européennes dispensé par des spécialistes internationaux. Ses travaux de recherche ont été effectués dans le cadre d'une collaboration entre l'Université de Strasbourg (IREBS, UMR CNRS 7242) et la FST de l'Université de Nouakchott - Mauritanie (Laboratoire de Biochimie). **Mr OULD BABAH** a bénéficié du soutien financier de l'équipe d'accueil, d'une bourse de 12 mois de la coopération Française et d'un soutien de 6 mois du CDE.

*A vous, mes chers parents !  
Je dédie cette thèse avec tout l'amour,  
le respect et la gratitude que je vous porte.  
Ce que je suis aujourd'hui et ce que je deviendrai est grâce à vous.  
Ma pauvre plume n'arrive pas à exprimer le trémolo sentimental  
que je subis en ce moment. Le journal de mes souvenirs proches et lointains passe devant mes  
yeux me rappelant les efforts énormes que vous faites tous les jours pour ma personne.  
Que Allah vous protège et vous accorde une longue vie mes chers  
Babah et Mariama.*

*« Trouver n'est rien, c'est le plan qui est difficile. »*

*Fiodor Dostoïevski*

*« Observer attentivement, c'est se rappeler distinctement. »*

*Edgar Allan Poe*

*Voilà à mon avis deux qualités nécessaires pour bien réussir un travail scientifique :*

*Perspicacité d'observation et sens de planification.*

---

# Remerciements

Ces travaux de thèse ont été réalisés dans les locaux de l'IREBS, UMR 7242 "Biotechnologie et signalisation cellulaire (BSC)" sous la direction de **Dr Jean-luc Galzi**, au sein de l'équipe "Oncoprotéines" dirigée par **Dr Gilles TRAVÉ**, **Dr Georges ORFANOUDAKIS** et **Dr Murielle MASSON**.

Tout d'abord je voudrais exprimer très particulièrement mes meilleurs remerciements et ma profonde gratitude à mon directeur de thèse **Gilles TRAVÉ** qui m'a accueilli, formé et encadré pertinemment et gentiment pendant les quatre années de cette thèse. Cher **Gilles**, je te remercie profondément pour m'avoir aidé à décortiquer les subtilités de ce projet de recherche et de m'avoir offert l'excellente opportunité de travailler sous ta direction et d'acquérir ta précieuse confiance, ce qui m'a permis de développer mes compétences critiques et constructives, mais aussi pour ta sympathie naturelle, ta sincérité, ton ouverture d'esprit et de coeur, ton exceptionnelle humanité et ta patience. Je te remercie encore une fois, pour les livres et les idées que nous avons échangées dans les domaines de la poésie, de la philosophie, de la politique et de l'histoire.

Je tiens à exprimer ma reconnaissance au membre du jury qui ont bien voulu accepter de consacrer de leurs temps pour juger ce travail : **Dr Ahmed HOUMEIDA**, **Dr. André ZAPUN**, **Dr. Marc RUFF** et **Dr. Yves JACOB**.

Je remercie particulièrement **Georges ORFANOUDAKIS** de m'avoir accordé de son temps pour m'apprendre la technique 'Phage display' malgré ses nombreuses occupations administratives. Merci également à **Murielle MASSON**, **François DERYCKERE** et **Yves NOMINE** pour leur gentillesse et leurs conseils et orientations scientifiques.

Je voudrais également remercier infiniment **Katia ZANIER**, parce que ce travail n'aurait jamais pu être réalisé sans ses conseils et son encadrement. Merci **Katia** pour toutes les com-

## Remerciements

---

pétences et les connaissances que j'ai apprises grâce à ton aide.

Un grand merci à **Abdellahi Ould M'hamed Ould sidi** pour avoir facilité mon intégration depuis mon arrivée en France, et pour m'avoir initié à la paillasse dès mes premiers jours au laboratoire.

Mes remerciements vont également à **Katja LUCK** et **Sébastien CHARBOIER** avec qui j'ai eu le plaisir de partager le bureau 204 (Vielen Dank Frau Kaja und Herr Seb). Je vous remercie pour vos conseils méthodologiques et scientifiques mais aussi pour les moments sympathiques que nous avons passés ensemble, en essayant de cohabiter la raison théologique et la raison scientifique, le cyclisme et l'équitation ou la boxe thai et le kung fu.

Merci également à tous mes collègues de laboratoire : **Anaëlle, Anne, Anissa, Isa, Juan, Sadek** et **Xav**.

J'aimerais également remercier du fond du coeur, mes collègues du **Laboratoire de Biochimie** à la **FST** de **Nouakchott** : **Dr. Fatou Mint Veten** et **Dr. Ghlana Meiloud**.

Je remercie cordialement **Mr Gunter Stier** de l'université de Heidelberg pour les compétences que j'ai acquises grâce à lui dans le domaine du clonage et production des protéines recombinantes.

Je tiens aussi à remercier vivement mes frères mauritaniens à Strasbourg :

**Bah ould Beich, Abdellahi Sweidehmed, Bebbe Ould Elman, Khaled Ould Kar, Moustapha Moud, Moustapha ould Isselkou (Kalkami), Bah ould Ova, Saleck Ould Jdide** et **Sid'ahmed Chouaib**.

Quant à mes **parents, sœurs, frères, nièces** et **neveux** je vous dis tous simplement : cette thèse est la vôtre, sans vous tout cela n'a aucun goût.

Je remercie également les institutions qui m'ont soutenues financièrement durant la période de mes études doctorales : Notre équipe "Oncoprotéines", la Coopération Française (CF) et le Collège doctoral européen (CDE), je remercie particulièrement Madame **Céline Montibeller** du CDE.

---

# Abréviations

## Abréviations

---

- **ADN** : Acide Désoxyribonucléique
  - **ARN** : Acide Ribonucléique
  - **ARNm** : ARN messenger
  - **ATP** : Adénosine 5'-triphosphate
  - **BD** : Bromodomaine
  - **BPV** : Bovine Papillomavirus
  - **Brd4** : Bromodomain containing protein 4
  - **BRM** : Bas Risque Muqueux
  - **CDK** : Cyclin Dependent Kinase
  - **CIS** : Carcinome In Situ
  - **CMH** : Complexe Majeur d'Histocompatibilité
  - **CmPV** : Chelonia mydas PV
  - **CRPV** : Cottontail Rabbit Papillomavirus
  - **DBD** : DNA Binding Domain
  - **DPV** : Deer Papillomavirus
  - **E** : Early region
  - **E6AP** : E6-Associated Protein
  - **E6APp** : Peptide de E6AP
  - **EDTA** : Ethylendiamintetraacetic acid
  - **EGF** : Epidermal Growth Factor
  - **EGFR** : EGF Receptor
  - **FLPV** : Francolinus leucoscepus papillomavirus
  - **GST** : Glutathione-S-Transferase
  - **GTP** : Guanosine triphosphate
  - **HPV** : Human Papillomavirus
  - **HRM** : Haut Risque Muqueux
  - **HSQC** : Heteronuclear Single Quantum Coherence
  - **ICTV** : International Committee on Taxonomy of Viruses
  - **IPTG** : Isopropyl-1- $\beta$ -D-thiogalactopyranoside
  - **ITC** : Isothermal Titration Calorimetry
  - **Kd** : Constante de dissociation
  - **KDa** : Kilo-Dalton
  - **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
  - **NF $\kappa$ B** : Nuclear Factor kappa B
  - **NIE** : Néoplasies Intraépithéliales
  - **NLS** : Nuclear Localisation Signal
  - **ORF** : Open Reading Frame
  - **Ori** : Origine de réplication
  - **PAGE** : Polyacrylamide Gel Electrophoresis ;
  - **Pax-p** : Peptide de Paxilline
  - **PCR** : Polymerase Chain Reaction
  - **PML** : Promyelocytic leukemia protein
  - **Poly A1** : Site de Polyadénylation précoce
  - **Poly A2** : Site de Polyadénylation tardive
  - **pRb** : Protéine du Rétinoblastome
  - **PV** : Papillomavirus
  - **RAR** : Retinoic acid receptor
  - **Rb** : Retinoblastoma protein
  - **RMN** : Résonance magnétique nucléaire
  - **SDS** : Sodium dodecyl sulphate
  - **siRNA** : Small interfering RNA
  - **SPR** : Surface Plasmon Resonance
  - **TAD** : Transactivation domain
  - **TERT** : Telomerase Reverse Transcriptase
  - **TEV** : Tobacco Etch Virus protease
  - **TEVs** : Site de reconnaissance de TEV
  - **TGF** : Transforming Growth Factor
  - **TNF** : Tumour Necrosis Factor
  - **URR** : Upstream Regulatory Region
  - **VIH** : Virus de l'Immunodéficience Humaine
  - **VLP** : Virus-Like-Particles
  - **WT** : Wild type
- 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.

---

# Table des matières

Table des figures	ix
Liste des tableaux	xi
<b>INTRODUCTION GENERALE</b>	<b>1</b>
<b>I Les Papillomavirus et leurs Pathologies</b>	<b>3</b>
1 Les Papillomavirus . . . . .	3
1.1 Définition et historique . . . . .	3
1.2 Classification . . . . .	4
1.3 Structure du virus . . . . .	8
1.3.1 Génome . . . . .	8
1.3.2 Capside . . . . .	9
1.4 L'infection virale et ses conséquences . . . . .	11
1.4.1 Transmission et entrée du virus . . . . .	11
1.4.2 Maintien du génome viral . . . . .	12
1.4.3 Prolifération cellulaire . . . . .	13
1.4.4 Amplification du génome viral . . . . .	13
1.4.5 Synthèse du virus . . . . .	14
1.4.6 Un évènement rare : l'intégration de l'ADN viral dans le génome cellulaire . . . . .	14
2 Pathologies liées aux HPV . . . . .	15
2.1 Néoplasie Intraépithéliale et Cancer du col de l'utérus . . . . .	15



2.2	Autres infections des muqueuses anogénitales . . . . .	18
2.2.1	Condylomes ano-génitaux . . . . .	18
2.3	Infections des muqueuses extra-génitales . . . . .	18
2.3.1	Papillomes buccaux . . . . .	18
2.3.2	Papillomatose orale floride . . . . .	18
2.4	Infections cutanées . . . . .	18
2.4.1	Verrues . . . . .	18
2.4.2	Epidermodysplasie verruciforme et Carcinome épidermoïde . . . . .	19
2.5	Diagnostic et thérapie . . . . .	19
2.5.1	Dépistage . . . . .	19
2.5.2	Traitements . . . . .	20
2.5.3	Vaccinations . . . . .	20
<b>II Caractéristiques structurales et fonctionnelles des protéines des HPV</b>		<b>21</b>
1	Les protéines tardives : L1 et L2 . . . . .	21
1.1	La protéine L1 . . . . .	21
1.2	La protéine L2 . . . . .	24
2	Les protéines précoces non transformatrices : E1, E2 et E4 . . . . .	25
2.1	La protéine E1 . . . . .	25
2.2	La protéine E2 . . . . .	27
2.3	La protéine E4 . . . . .	31
3	Les protéines précoces transformatrices : E5, E6 et E7 . . . . .	31
3.1	La protéine E5 . . . . .	31
3.2	La protéine E7 . . . . .	32
3.2.1	Caractéristiques biochimiques et structurales de E7 . . . . .	32
3.2.2	Activité biologique et partenaires cellulaires . . . . .	34
3.3	La protéine E6 . . . . .	36
3.3.1	Caractéristiques biochimiques et structurales de E6 . . . . .	36
3.3.2	Activité biologique et partenaires cellulaires . . . . .	38
<b>III Introduction au projet de recherche</b>		<b>47</b>
<b>RESULTATS ET DISCUSSION</b>		<b>51</b>

---

<b>I</b>	<b>Etude biochimique et structurale de la protéine E6</b>	<b>53</b>
1	Stratégies d'expression et de purification de la protéine E6 issue des Papillomavirus bovins (BPV1 E6) . . . . .	53
2	Structure cristallographique de la protéine E6 des Papillomavirus humains (HPV16) en complexe avec le peptide E6APp . . . . .	64
2.1	Problématique . . . . .	64
2.2	Résultats . . . . .	65
2.2.1	Recherche du meilleur mutant . . . . .	65
2.2.2	Stratégie d'étude de la structure de HPV16 E6 . . . . .	69
2.2.3	Description structurale du cristal HPV16E6/E6APp . . . . .	70
2.3	Discussion . . . . .	71
2.3.1	Repliements E6N et E6C . . . . .	71
2.3.2	Comparaison des structures de E6 bovine et humaine . . . . .	72
2.3.3	Structure de la forme libre de E6 . . . . .	74
2.3.4	Un nouveau mode de reconnaissance des motifs LxxLL . . . . .	74
2.3.5	Conclusion . . . . .	75
<b>II</b>	<b>Analyse du système d'interaction de la protéine E6</b>	<b>77</b>
1	Problématique . . . . .	77
1.1	Préambule . . . . .	77
1.2	Introduction sur la technique de chromatographie Hold-up . . . . .	78
2	Résultats . . . . .	82
2.1	Interactome E6 - LxxLL . . . . .	82
2.1.1	Série 1 : mutants de E6APp . . . . .	83
2.1.2	Série 2 : autres peptides LxxLL d'origine cellulaire . . . . .	88
2.2	Etude d'interface d'interaction E6 - LxxLL . . . . .	93
2.2.1	Complexes HPV16 E6/E6APp et BPV1 E6 - Pax.p (Structures cristallographiques) . . . . .	93
2.2.2	Concernant les autres types de E6 (Modélisation par homologie) . . . . .	94
2.2.3	Détermination de la zone d'interaction . . . . .	95
3	Discussion . . . . .	100
3.1	Modèle de mécanisme d'interaction E6-LxxLL . . . . .	100
3.2	Spécificité et affinité . . . . .	101
3.3	Différences de comportement entre le Peptide et la Protéine . . . . .	102
3.4	Mode compétitif ou enzymatique . . . . .	102

<b>III Mise en évidence d'une nouvelle interaction entre la protéine virale E2 et la protéine cellulaire Brd4</b>	<b>105</b>
1 Problématique . . . . .	105
2 Résultats . . . . .	106
2.1 Interface de reconnaissance . . . . .	106
2.2 Rôle de la phosphorylation et de la salinité du solvant . . . . .	109
3 Discussion . . . . .	114
<b>CONCLUSIONS GENERALES ET PERSPECTIVES</b>	<b>115</b>
<b>MATERIELS ET METHODES</b>	<b>121</b>
<b>Références bibliographiques</b>	<b>131</b>
<b>Annexe</b>	<b>151</b>
1 Manuscrit en cours de rédaction : Structural basis for hijacking of cellular LxxLL motifs by a papillomavirus E6 oncoprotein . . . . .	151
<b>Annexe</b>	<b>187</b>
2 Figures supplémentaires . . . . .	187

---

## Table des figures

I.1	<b>Classification phylogénétique des Papillomavirus</b> . . . . .	7
I.2	<b>Organisation du génome viral</b> . . . . .	9
I.3	<b>La capside virale</b> . . . . .	10
I.4	<b>Cycle de vie du virus HPV</b> . . . . .	12
I.5	<b>Intégration de l'ADN viral dans le génome hôte</b> . . . . .	15
I.6	<b>Contribution relative cumulée des HPV retrouvés dans les tumeurs du Cancer utérin</b> . . . . .	16
I.7	<b>Aspects histologiques des lésions précancéreuses</b> . . . . .	17
II.1	<b>Structure de la protéine L1</b> . . . . .	23
II.2	<b>Structure de la protéine E1</b> . . . . .	26
II.3	<b>Structure du complexe trimérique E1-E2-ADN</b> . . . . .	29
II.4	<b>Interaction E2-Brd4</b> . . . . .	30
II.5	<b>Alignement de séquences et structure de E7</b> . . . . .	33
II.6	<b>Alignements de séquences et structure secondaire de E6</b> . . . . .	37
III.1	<b>Tendance à l'agrégation de E6 sauvage</b> . . . . .	47
III.2	<b>Structure en solution du domaine E6C</b> . . . . .	49
III.3	<b>Structures en solution des deux domaines de HPV16 E6</b> . . . . .	50
I.1	<b>Comparaison de la solubilité des mutants de E6</b> . . . . .	66
I.2	<b>Spectres HSQC des principaux mutants E6</b> . . . . .	68
I.3	<b>Stratégie de production d'une protéine E6 soluble</b> . . . . .	69
I.4	<b>Structure cristallographique de E6 en complexe avec E6APp</b> . . . . .	70

I.5	Topologie structurale de E6 . . . . .	71
I.6	Surface électrique et hydrophobe de E6 . . . . .	72
I.7	Comparaison des deux structures 3D de E6 (humaine et bovine) . . .	73
I.8	Interfaces d'interactions HPV16 E6-E6AP et BPV1 E6-Paxilline . . .	74
II.1	Principe de la technique de Hold-up . . . . .	79
II.2	Technique de Hold-up automatisée . . . . .	80
II.3	Heat map des résultats de Holdup - Série 1 (types confondus) . . . . .	84
II.4	Heat map des résultats de Holdup - Série 1 (types séparés) . . . . .	85
II.5	Résultats de "Holdup" - Série 1 (Histogrammes) . . . . .	86
II.6	Résultats de "Holdup" - Série 2 (Histogrammes) . . . . .	92
II.7	Interfaces d'interactions HPV16 E6-E6AP et BPV1 E6-Paxilline . . .	94
II.8	Alignement des séquences d'interfaces d'interactions de divers types E6 . . . . .	96
II.9	Réseaux de contacts E6-peptide (HPV16 et BPV1) . . . . .	98
II.10	Interface d'interaction des modèles HPV1,3,5,9 et 18 . . . . .	99
III.1	Surface électrique et hydrophobe de E6 . . . . .	106
III.2	Analyse biophysique de l'interaction Brd4(BDII)-E2 (DBD) . . . . .	108
III.3	Interface d'interaction Brd4(BDII)-E2 (DBD) . . . . .	110
III.4	Relation proportionnelle entre la concentration du peptide et les déplacements chimiques des résidus . . . . .	111
III.5	Comparaison d'effets induits par les différents peptides phosphorylés	112
III.6	Robustesse de l'interaction en fonction de la force ionique . . . . .	113

---

## Liste des tableaux

I.1	Classification et pathologies des HPV . . . . .	6
II.1	Liste de partenaires cellulaires de E7 . . . . .	35
II.2	Liste de partenaires cellulaires de E6 . . . . .	39
II.1	Liste de peptides utilisés en Série 2 . . . . .	90
II.2	Résultats de "Holdup" - Série 2 (Tableau) . . . . .	91
II.3	Variations de séquences de la zone d'interaction E6-LxxLL . . . . .	97

## Table des matières

---

---

# INTRODUCTION GENERALE



## INTRODUCTION GENERALE

---

---

---

# Chapitre I

---

## Les Papillomavirus et leurs Pathologies

### 1 Les Papillomavirus

#### 1.1 Définition et historique

Les Papillomavirus (PV) sont des petits virus non enveloppés à ADN bicaténaire et circulaire d'environ 8 kb de taille. Ils infectent mammifères, oiseaux et reptiles, induisant des lésions prolifératives du tissu épithélial (verrues, condylomes) qui dégénèrent parfois en cancers (de Villiers *et al.* [2004]; zur Hausen [2002]).

Les Papillomavirus sont identifiés par leurs hôtes naturels : humains (HPV), Bovins (BPV), Cervidés (DPV), Chéloniens (CmPV) etc. . . (Groff et Lancaster [1985]). Ils sont généralement très spécifiques à leurs hôtes. Cependant, il a été remarqué que certains Papillomavirus des bovins (BPV) peuvent causer des infections chez les chevaux ou chez d'autres espèces outre les bovins (Chambers *et al.* [2003]).

Si des maladies liées aux Papillomavirus (verrues, condylomes . . .) sont connues depuis les ères les plus reculées dans l'histoire, ce n'est qu'en 1933 qu'on a pu observer et décrire pour la première fois un Papillomavirus grâce aux expériences de Shope sur les lapins. En 1950 Strauss et ses collègues ont pu mettre en évidence les particules virales en microscopie électronique (STRAUSS *et al.* [1950]).

Les années 70-80 du vingtième siècle ont connu des progrès majeurs en matière de connaissance des Papillomavirus. Les travaux de Gérard Orth et de son équipe sur le lapin à queue blanche ont permis la caractérisation d'une trentaine de papillomavirus et la démonstration du lien entre la réplication virale, l'expression des oncogènes et la différenciation du kératinocyte dans les tumeurs bénignes ou malignes. En 1978, Orth et Jablonska ont révélé la présence des Papillomavirus dans des lésions malignes de l'épidermodysplasie verruciforme (voir plus loin la section "pathologies") (Orth *et al.* [1977, 1978, 1979]). En 1983 et 1984, zur Hausen et ses collègues ont isolé pour la première fois un Papillomavirus (HPV-16 et HPV-18) directement à

## Chapitre I. Les Papillomavirus et leurs Pathologies

---

partir d'une biopsie du cancer du col de l'utérus, démontrant ainsi l'association entre les HPV et le cancer du col de l'utérus, ce qui lui valut le prix Nobel en 2008 (Boshart *et al.* [1984]; Durst *et al.* [1983]).

### 1.2 Classification

Les Papillomavirus ont été classés en 1962 par Melnick sous la famille des Papovaviridea qui englobait outre les Papillomavirus, les polyomavirus et le virus vacuolisant du singe (SV40) sur la base de caractéristiques communes comme l'absence d'enveloppe et la présentation de la matière génomique en ADN double brin. De plus, le gène de E1 des Papillomavirus présente un domaine hélicase retrouvé dans le T-antigène des polyomavirus (Clertant et Seif [1984]).

En 2002 l'ICTV (International Committee on Taxonomy of Viruses) a annoncé dans son 7ème rapport la scission de la famille des Papovaviridae en deux familles distinctes sur la base des différences de taille du virion, de la longueur d'ADN et du faible score de similitude entre les deux génomes. Les deux nouvelles familles sont : les Polyomaviridae pour les Polyomavirus et les Papillomaviridae pour les Papillomavirus (de Villiers *et al.* [2004]).

A cause de l'inadaptation des Papillomavirus aux techniques classiques de culture en raison de son cycle de vie atypique, on ne pouvait pas établir une taxonomie basée sur les sérotypes, ce qui a amené des scientifiques à proposer une méthode de classification basée principalement sur le pourcentage d'identité entre les séquences nucléotidiques du gène L1 de chaque 'type' du virus, prenant aussi en compte certaines caractéristiques biologiques et médicales comme la spécificité d'hôte, le tropisme tissulaire, le pouvoir pathogène et l'organisation génomique. Le gène L1 est le plus conservé chez les Papillomavirus et code pour la protéine L1, nécessaire et principale pour la formation des capsides virales et suffisante par elle seule à produire les VLP (Virus-Like-Particles) immunisantes, utilisés dans les vaccins anti HPV. Ce classement consiste à subdiviser la famille des Papillomaviridea en nombre de genres qui partagent moins de 60% de leurs séquences de L1 et qui sont désignés par une lettre grecque (alpha, bêta, gamma...etc.). Chaque genre est formé par des espèces qui ont entre 60% et 70% d'identité de séquences et ils sont désignés par un chiffre (alpha-1, alpha-2, bêta-1. . .etc.). Chaque espèce est divisée en plusieurs types (HPV16, BPV1. . .etc.) présentant une identité de séquence de 71% à 89%. Enfin au sein d'un même type on considère que les virus appartiennent à des sous-types différents s'ils ont entre 90% et 98% d'identité (Bernard *et al.* [2010]; Cruz *et al.* [2004]; de Villiers *et al.* [2004]; Fauquet et Fargette [2005]; Kammer *et al.* [2002]). Cette méthode a permis la fondation d'une nomenclature formelle des Papillomavirus, largement utilisée par la communauté scientifique et adoptée globalement par l'ICTV dans son 8ème rapport publié en 2005 (Fauquet et Fargette

[2005]).

Actuellement, grâce à cette nomenclature, on a pu identifier un arbre phylogénique de 29 genres formés par 189 ‘types’ distribués selon leurs hôtes naturels en :120 humains (5 genres : alpha, bêta, gamma, mu et nu), 64 autres mammifères, 3 oiseaux et 2 reptiles.(Bernard *et al.* [2010]; de Villiers *et al.* [2004])(Figure I.1).

- **Tropisme tissulaire des HPV**

Les HPV sont classiquement distingués en deux catégories selon leurs tropisme tissulaires : les muqueux dont la majorité appartiennent aux genres alpha et les cutanés qui appartiennent aux genres bêta, mu et nu. Cependant, il y a des HPV qui ont un tropisme mixte comme les types 3, 7 et 10 (Tableau I.1).

- **Potentiel pathogène des HPV**

Les données épidémiologiques et cliniques cumulés depuis les années 1930 ont permis l’établissement d’un classement des HPV muqueux selon leur risque à développer un cancer en deux grands groupes (McMurray *et al.* [2001]; Munoz *et al.* [2006]) (Tableau I.1) :

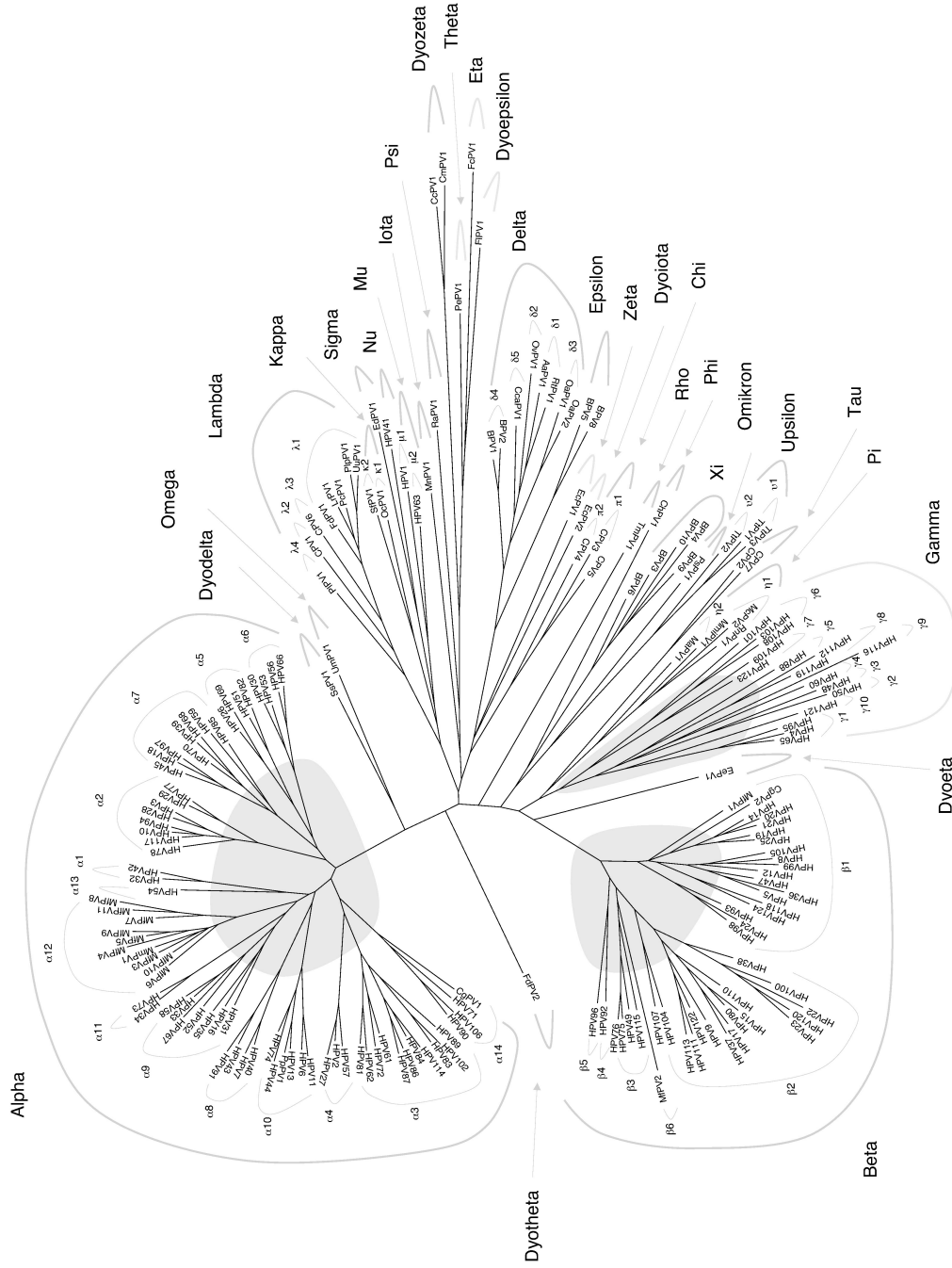
- **HPV à Haut risque muqueux (HPV à HRM)** : Il s’agit des types 16,18, 31, 33, 35, 39, 45, 51, 52, 56, 58 et 59. En plus de ces 12 types on peut ajouter 6 autres types qui sont probablement à HRM : 26, 53, 66, 68, 73 et 82
- **HPV à Bas risque muqueux (HPV à BRM)** : Il s’agit des types 6, 11, 13, 40, 42, 43, 44, 54, 61, 70, 72, 81 et 89.

## Chapitre I. Les Papillomavirus et leurs Pathologies

Genre	Espèce	Type	Risque/Pathologie
<b>Alpha-Papillomavirus</b>	01	HPV 32, 42	BRM, lésions orales ou génitales
	02	HPV 3, 10, 28, 29, 77, 78, 94, 117	BRM, lésions cutanées, parfois muqueuses
	03	HPV 61, 62, 72, 81, 83, 84, 86, 87, 89, 102, 114	BRM, lésions muqueuses
	04	HPV 2, 27, 57	Verrues vulgaires
	05	HPV 26, 51, 69, 82	HRM, lésions muqueuses
	06	HPV 30, 53, 56, 66	HRM, lésions muqueuses
	07	HPV 18, 39, 45, 59, 68, 70, 85, 97	HRM, lésions muqueuses
	08	HPV 7, 40, 43, 91	BRM, lésions cutanées et muqueuses
	09	HPV 16, 31, 33, 35, 52, 58,67	HRM, lésions muqueuses
	10	HPV 6, 11, 13, 44, 74	BRM, codylomes acuminés
	11	HPV 34, 73	HRM, lésions muqueuses
	13	HPV 54	BRM, lésions muqueuses
	14	HPV 71, 90, 106	BRM, lésions muqueuses
	<b>Bêta-Papillomavirus</b>	01	HPV 5, 8, 12, 14, 19, 20, 21, 25, 36, 47, 93, 98, 99, 105, 118, 124
02		HPV 9, 15, 17, 22, 23, 37, 38, 80, 100, 104, 107, 110, 111, 113, 120, 122	lésions cutanées bénignes, parfois malignes
03		HPV 49, 75,76, 115	lésions cutanées bénignes
04		HPV 92	Cancer et lésions cutanées
05		HPV 96	Cancer et lésions cutanées
<b>Gamma-Papillomavirus</b>	01	HPV 4, 65, 95	lésions cutanées
	02	HPV 50	lésions cutanées
	03	HPV 48	lésions cutanées
	04	HPV 60	lésions cutanées
	05	HPV 88	lésions cutanées
	06	HPV 101, 103, 108	lésions cutanées
	07	HPV 109, 123	lésions cutanées
	08	HPV 112, 119	lésions cutanées
	09	HPV 116	lésions cutanées
	10	HPV 121	lésions cutanées
<b>Mu-Papillomavirus</b>	01	HPV 1	Verrues vulgaires, plantaires
	02	HPV 63	Verrues vulgaires, plantaires
<b>Nu-Papillomavirus</b>	01	HPV 41	Cancer et lésions cutanées

**Tableau I.1 – Classification et pathologies des HPV**

Ce tableau montre la classification des Papillomavirus humains. les HPV sont présentés ici selon leurs genres, espèces, types et leur risques à développer un cancer ; Haut risque muqueux (HRM en rouge) ou Bas Risque muqueux (BRM). la classification HRM et BRM ne considère que les HPV des muqueuses. La dernière colonne montre la nature des pathologies liées aux HPV (Cancer, lésions cutanées, lésions muqueuses... ) (d'après Bernard *et al.* [2010]).



**Figure I.1 – Classification phylogénétique des Papillomavirus**  
 Arbre phylogénétique des 29 genres des Papillomavirus qui englobent 189 ‘types’ (120 humains, 64 autres mammifères, 3 oiseaux et 2 reptiles), classés selon la divergence des séquences de leurs protéines structurales L1, (d’après [Bernard et al. \[2010\]](#)).

### 1.3 Structure du virus

#### 1.3.1 Génome

Tous les Papillomavirus contiennent un ADN génomique circulaire, bicaténaire avec un seul brin codant, d'environ 8 kb de taille, avec 8 à 10 phases ouvertes de lectures (Open Reading Fragment [ORF]). La transcription des ORF se fait d'une manière bi, tri ou même polycistronique sous le contrôle de deux promoteurs principaux. Les ARN transcrits viraux suivent un processus de polyadénylation et d'épissage alternatifs avant leur traduction (Zheng et Baker [2006]). Le système de régulation des promoteurs ordonne l'expression des protéines virales sur deux phases : précoce et tardive. Selon cette chronologie d'expression on peut diviser le génome des Papillomavirus en deux régions codantes plus une Longue Région de Contrôle non codante (LCR), les trois régions sont séparées par deux sites de Polyadénylation (précoce :PolyA1 et tardif : PolyA2) (Figure I.2).

**Région précoce (Early [E]) :** Cette région traduite précocement couvre plus de 50% du génome viral et se trouve entre l'origine de répllication (Ori) et la région tardive. Elle comporte six ORF exprimés en protéines non structurales E1, E2, E4, E5, E6 et E7 (Danos *et al.* [1982]). Au début, il a été attribué à cette région deux autres ORF : E3 et E8, mais à partir de ces deux ORF on a seulement prouvé l'expression d'une seule protéine de fusion entre E8 et E2C chez deux types de Papillomavirus, BPV1(Choe *et al.* [1989]; Lambert *et al.* [1987]) et HPV31 (Stubenrauch *et al.* [2000, 2001]) potentiellement impliquée dans la régulation négative de la transcription et la répllication virale(Han *et al.* [1998]; Harry et Wettstein [1996]). La région précoce est sous le contrôle du promoteur précoce qui se trouve en amont de l'ORF de E6, au nucléotide 97 chez HPV16 (p97).

**Région tardive (Late [L]) :** Cette région traduite tardivement occupe environ 40% du génome viral et se trouve en aval de la région précoce, Elle comporte les ORF des protéines de structure L1 (protéine majeure de capsid) et L2 (protéine mineure de capsid). La région tardive est sous le contrôle du promoteur tardif qui se trouve au sein de l'ORF de E7, au nucléotide 670 chez HPV16 (p670).

**Région non codante (Long Control Region [LCR]) :** La région LCR autrement appelé URR (Upstream Regulatory Region) ou NCR (Non Coding Region) couvre environ de 850 pb (10% du génome) et se trouve en amont de la région précoce. Comme son nom l'indique la LCR ne porte pas de séquences codant pour des protéines, mais elle porte l'origine de répllication

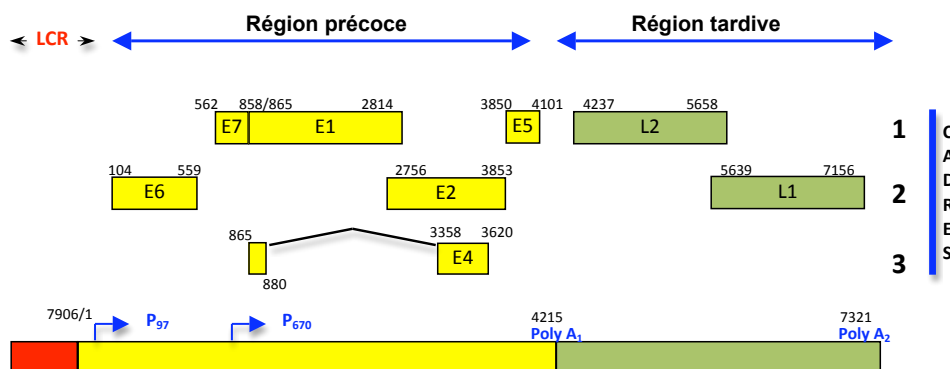


Figure I.2 – Organisation du génome viral

Cette figure montre la structure linéaire du génome viral (HPV16), les promoteurs précoce (P<sub>97</sub>) et tardif (P<sub>670</sub>), les régions de polyadénylation précoce (Poly A<sub>1</sub>) et tardif (Poly A<sub>2</sub>), les ORFs des protéines précoces (E1, E2, E4, E5, E6 et E7) et tardives (L1 et L2) et la région non codante (LCR). Notons que l'ORF de E4 regroupe deux exons et que l'ORF de E5 est absent chez les  $\beta$ -Papillomavirus. (adapté d'après [Zheng et Baker \[2006\]](#)).

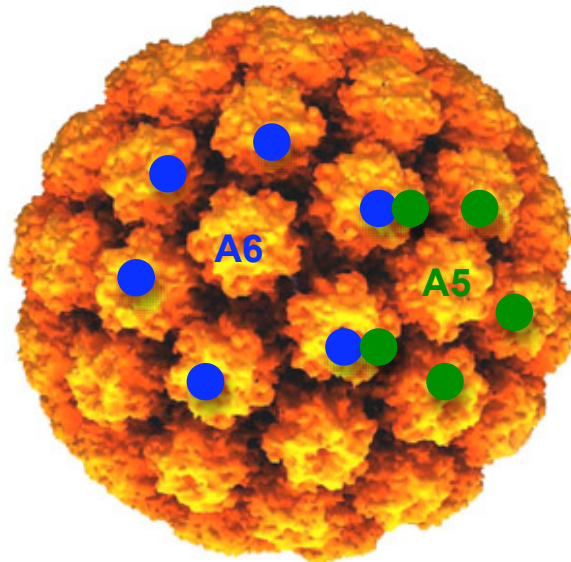
(Ori), ainsi que des sites de fixation des facteurs de transcription importants pour la régulation de l'initiation de transcription par ARN polymérase II au niveau de deux promoteurs précoce et tardif. La région LCR comporte aussi des sites de fixations pour les protéines virales E1 et E2 ([Bernard \[2002\]](#); [Kanaya \*et al.\* \[1997\]](#); [Kisseljov \[2000\]](#); [Morris \*et al.\* \[1993b\]](#)).

### 1.3.2 Capside

Les papillomavirus forment des virions non enveloppés, délimités par une capsidie de 600 Å de diamètre, édifiée principalement par la protéine majeure de capsidie L1 nécessaire et suffisante à elle seule pour construire une capsidie synthétique identique à la capsidie naturelle ([Bishop \*et al.\* \[2007\]](#); [Kirnbauer \*et al.\* \[1992\]](#)). Cependant la protéine mineure de capsidie L2 participe aussi à l'assemblage des capsomères mais avec une faible représentation (12 copies par virion vs 360 copies de L1) et servent surtout à former une sorte de pont entre la capsidie et le génome ([Baker \*et al.\* \[1991\]](#); [Buck \*et al.\* \[2008\]](#)). Les données structurales basées sur des analyses de microscopie Electronique sur les virions des Papillomavirus des lapins, des bovins (BPV1) et des humains (HPV1) montrent qu'ils ont globalement la même structure générale ([Belnap \*et al.\* \[1996\]](#); [Hagensee \*et al.\* \[1994\]](#); [Trus \*et al.\* \[1997\]](#)).

En 2000 Chen et ses collègues ont publié une structure atomique de la protéine L1 ([Chen \*et al.\* \[2000\]](#)) à partir d'un cristal de 12 pentamères assemblés (Small Virus Like Particle [VLP]) stables à pH faible. Deux ans plus tard, Modis et al. ([Modis \*et al.\* \[2002\]](#)) ont proposé un modèle





**Figure I.3 – La capsid virale**

Structure de la capsid des Papillomavirus composée par 72 pentamères de L1 (360 copies) : 12 pentamères sont situés sur un axe de symétrie 5 ([A5] axes indiqués par taches vertes), les 60 autres pentamères sont situés sur un axe de symétrie 6 ([A6] axes indiqués par taches bleues). Cette structure est basée sur une reconstruction par cryo-électromicroscopie.(adapté d'après [Modis \*et al.\* \[2002\]](#)).

qui présente les différentes interactions inter et intra capsomère et ont décrit l'architecture générale de la capsid. Le squelette icosaédrique de la capsid est composé par 360 copies de L1 assemblées sous forme de 72 pentamères dont 60 pentamères sont positionnés sur des axes de symétrie 6 (A6) et 12 pentamères sur des axes de symétrie 5 (A5) (Figure I.3).

Malgré les différences entre les Papillomavirus et les polyomavirus en matière de taille du virion (600Å vs 500Å) ou en matière de taille du génome (8kb pour 5kb) [Modis et al. \[2002\]](#) et malgré l'absence d'une similarité significative en terme de séquence au niveau de la protéine L1, la comparaison de la capsid des Papillomavirus avec celle des polyomavirus montre une forte similarité entre les deux structures dont le principal trait est l'assemblage des pentamères L1 chez les Papillomavirus et l'assemblage des pentamères de son homologue chez les polyomavirus (VP1 chez SV40) ([Chen \*et al.\* \[2000\]](#); [Rayment \*et al.\* \[1982\]](#)).

### 1.4 L'infection virale et ses conséquences

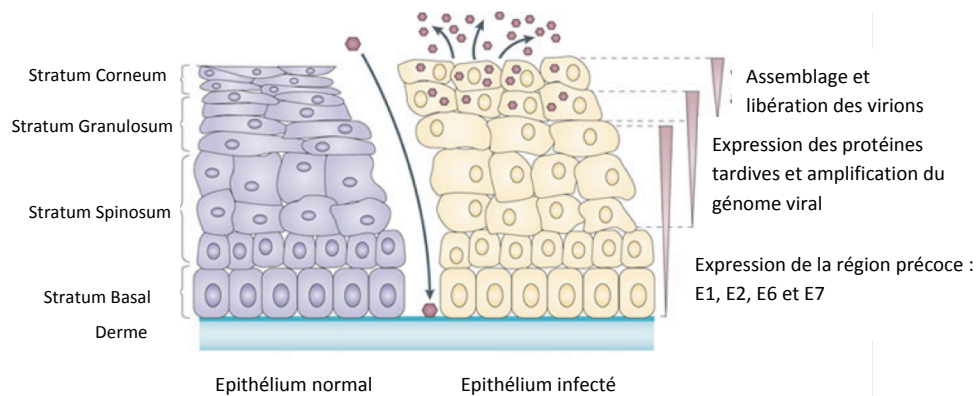
#### 1.4.1 Transmission et entrée du virus

##### **Transmission virale :**

La transmission des virus HPV se fait souvent par contact direct à travers les excoriations de l'épiderme ou des muqueuses d'une personne à une autre (hétéro-inoculation) ou du sujet à lui-même favorisée par le grattage des verrues par exemple (auto-inoculation) ce qui explique le caractère multifocal des lésions causés par les Papillomavirus. La transmission indirecte se fait vraisemblablement aussi par l'intermédiaire des objets ou des surfaces contaminées. L'infection de la région anale et génitale par les Papillomavirus est souvent associée aux maladies sexuellement transmissibles, dont elle présente aussi certaines caractéristiques épidémiologiques (Hildesheim *et al.* [1991]; Mardh *et al.* [1998]; Smith *et al.* [2002]). La transmission de la mère à l'enfant est possible par contact direct lors du passage à travers les muqueuses génitales (Kaye *et al.* [1996]), ou même par voie sanguine (Tseng *et al.* [1992]). La phase d'incubation varie d'une semaine à plusieurs mois. Parfois, le virus effectue une infection latente où l'ADN viral reste à l'état plasmidique en attendant l'effet d'autres facteurs capables de faire évoluer cette infection latente vers une infection productive (immunodépression par exemple) (Oh *et al.* [2004]).

##### **Entrée du virus :**

Les Papillomavirus ont une particularité remarquable par le fait qu'ils ont un tropisme tissulaire spécifique pour les épithéliums pluristratifiés (la peau et les muqueuses génitale ou buccale). On sait que le virus pénètre dans les cellules de la couche basale à la faveur d'une microlésion, mais le mécanisme de la pénétration du virus n'a pas encore formellement été élucidé à cause – peut-être – de la nature des cellules cibles de ces virus qui sont en différenciation constante. Cependant les intégrines  $\alpha 6 \beta 4$  semblent constituer un récepteur de ces virus permettant l'endocytose des virions dans des vésicules à clathrine. Les glycosaminoglycanes superficiels des cellules semblent aussi être impliqués dans l'infection virale. Dans le cytoplasme, le virus migre vers le noyau à l'aide du réseau de microtubules et de filaments d'actine du cytosquelette (Selinka *et al.* [2002]). Pour passer la barrière de la membrane nucléaire une décapsidation virale se fait après la rupture des ponts intra-capsomériques sous la pression de l'environnement cytoplasmique réducteur (Li *et al.* [1998]). La présence des séquences de localisation nucléaire dans les protéines L1 et L2 favorise le transport du génome viral dans le noyau (Selinka *et al.* [2002]). En cas d'une infection productive, l'éventuel processus qui suit l'entrée du virus est constitué de quatre étapes : maintien du génome viral, phase de prolifération, amplification du génome et synthèse du virus (Doorbar [2005])(Figure I.4).



**Figure I.4 – Cycle de vie du virus HPV**

Représentation schématisée de l'épithélium sain (à gauche) et infecté (à droite), reflétant les étapes clés dans l'infection : l'entrée du virus, l'expression des protéines précoces, l'expression des protéines tardives, l'amplification du l'ADN viral et la synthèse du virus (adapté d'après [Moody et Laimins \[2010\]](#)).

### 1.4.2 Maintien du génome viral

Après l'entrée du virus dans la couche basale, le génome viral se maintient dans les cellules de cette couche à l'état épisomal sous la forme d'un petit nombre de copies sous le contrôle des protéines virales précoces E1 et E2 exprimées déjà à ce stade ([Wilson \*et al.\* \[2002\]](#)). Ces deux protéines assurent la ségrégation équitable du génome viral pendant la division cellulaire ([You \*et al.\* \[2004\]](#)). Certains auteurs évoquent une expression éventuelle à ce stade des protéines virales "transformatrices" (capables de provoquer un processus de transformation cellulaire) E6 et E7 dans la couche basale ([Crum \*et al.\* \[1988\]](#)). En effet, il semble que l'infection initiale soit suivie d'une prolifération cellulaire accrue responsable d'une augmentation du nombre de cellules porteuses du génome viral. On pense donc que toutes les quatre protéines virales précoces sont exprimées à un niveau faible et que le génome viral est maintenu à un nombre de copies de l'ordre de 10 à 200 par cellule ([De Geest \*et al.\* \[1993\]](#); [Stanley \*et al.\* \[1989\]](#)). Cependant, d'autres études mettent en doute la contribution de E6 et E7 à la prolifération des cellules basales en suggérant que la seule expression des deux protéines E1 et E2 pourrait assurer le maintien de l'ADN épisomique du virus ([Zhang \*et al.\* \[1999\]](#)).

### 1.4.3 Prolifération cellulaire

A l'état normal, les cellules épithéliales arrêtent leur cycle cellulaire et commencent le processus de différenciation terminale dès la fin de leur migration aux couches suprabasales (Madison [2003]). En cas d'infection par HPV ces cellules continuent anormalement leur cycle cellulaire sous l'effet des oncoprotéines virales E6 et E7 (Sherman *et al.* [1997]). Pour réaliser cette fonction, il semble que ces deux protéines agissent d'une façon coopérative. Chez les HPV à HRM ces deux protéines sont exprimées à partir d'un ARNm bicistronique (Stacey *et al.* [2000]). E6 comme E7 sont capables, individuellement, de stimuler la progression du cycle cellulaire et de s'associer avec des protéines cellulaires impliquées dans la régulation du cycle cellulaire (Munger *et al.* [2001]). L'association de E7 avec le régulateur négatif du cycle cellulaire pRb empêche la liaison de ce dernier avec le facteur de transcription E2F permettant ainsi l'expression des protéines nécessaires à la réplication de l'ADN et levant le blocage de la phase S résultant préalablement de l'association pRb-E2F. L'interaction E7-pRb permet aussi l'expression de la cycline E, protéine nécessaire pour l'entrée en phase S. La protéine E7 est également capable de s'associer avec d'autres protéines cellulaires impliquées dans la régulation du cycle cellulaire comme le complexe de transcription AP-1 (Antinore *et al.* [1996]), les inhibiteurs des CDK (kinases cycline-dépendantes) p21 et p27 (Funk *et al.* [1997]) et la déacétylase d'histone HDAC (Longworth et Laimins [2004b]). La protéine E6 complète le rôle de E7 en agissant contre le processus d'apoptose déclenchée en réponse de l'entrée en phase S provoquée par la protéine E7. E6 interagit avec et dégrade la protéine p53 via un triple complexe dont le troisième partenaire est la protéine ubiquitine ligase (E6AP). E6 est aussi capable d'agir contre l'apoptose via d'autres protéines pro-apoptotiques comme Bak (Thomas et Banks [1998]) et Bax (Li et Dou [2000]). Notons également que les protéines E6 de HPV à HRM possèdent à leur extrémité C-terminale un motif conservé de liaison aux domaines PDZ qui leur permet de recruter des protéines à domaine PDZ impliquées dans le contrôle de l'adhésion cellulaire et de stimuler par ce biais la prolifération cellulaire indépendamment de E7 (Favre-Bonvin *et al.* [2005]; Gardiol *et al.* [1999]; Hampson *et al.* [2004]; Jing *et al.* [2007]; Nakagawa et Huibregtse [2000]).

### 1.4.4 Amplification du génome viral

L'amplification du génome viral requiert l'expression de toutes les protéines précoces (E1, E2, E4, E5, E6 et E7) (Genther *et al.* [2003]; Peh *et al.* [2004]). Le complexe d'initiation de réplication virale est composé de deux protéines E1/E2 qui jouent un rôle similaire aux protéines cellulaires cdc6/MCMs. La première étape est l'interaction de E2 sous sa forme dimérique avec la région LCR. Le dimère de E2 recrute ensuite un dimère de E1 au niveau de l'origine

## Chapitre I. Les Papillomavirus et leurs Pathologies

---

de réplication virale via une interaction entre les domaine hélicase de E1 et le domaine de transactivation de E2 (TAD). Après le recrutement du dimère de E1 sur l'ADN viral, le dimère de E2 se détache et le dimère de E1 prend une nouvelle forme hexamérique qui constitue sa forme fonctionnelle marquant ainsi le début de l'amplification du génome viral (Liu *et al.* [1998]).

### 1.4.5 Synthèse du virus

Les protéines structurales L1 et L2 nécessaires à la synthèse de la capside virale sont exprimées tardivement dans les cellules des couches supérieures après l'achèvement de l'amplification du génome viral (Ozbun et Meyers [1998]). La protéine mineure de capside L2 exprimée avant la protéine majeure L1 s'accumule sur les corps PML (Promyelocytic Leukemia nuclear body) et recrute L1 sur les PML. Il a été suggéré que ces corps PML représentent le site de la réplication d'ADN viral (Day *et al.* [1998]; Swindle *et al.* [1999]) et que l'accumulation des protéines structurales à ce site facilite la capsidation virale. Cependant la protéine L1 est capable d'effectuer un assemblage de particules virales similaire à la capside naturelle en absence de L2. Ceci suggère que le rôle de L2 est plutôt complémentaire, pour améliorer l'assemblage des particules (Stauffer *et al.* [1998]; Zhou *et al.* [1993]) et favoriser l'infectiosité du virus (Rodén *et al.* [1997]). Après l'assemblage des particules (voir paragraphe 'Capside'), les virions sont libérés à l'extérieur des cellules pour recommencer le cycle de vie viral. La libération des virions ne se fait qu'après que la cellule porteuse ait mis fin à sa migration et soit arrivée à la surface de l'épithélium (Rodén *et al.* [1997]). On pense que la protéine E4 exprimée simultanément avec L1 et L2 dans les mêmes cellules (Doorbar *et al.* [1997]), contribue à la libération des virus en perturbant l'intégrité de la kératine (Wang *et al.* [2004]) et en affectant l'assemblage de l'enveloppe cornée (Lehr *et al.* [2004]).

### 1.4.6 Un évènement rare : l'intégration de l'ADN viral dans le génome cellulaire

Dans les infections produites par les HPV à HRM, il arrive parfois que le génome viral s'intègre au sein du génome de la cellule hôte. Cet évènement semble être très important pour le processus de cancérogenèse. La linéarisation du génome viral permettant cette intégration se fait grâce à une rupture au niveau du gène E2 (Hebner et Laimins [2006]; Romanczuk et Howley [1992]). L'intégration et la rupture du gène E2 sont suivies par une augmentation du taux des oncoprotéines E6 et E7, une diminution des protéines structurales (L1 et L2) et des autres protéines précoces (E4 et E5). L'abondance des protéines transformatrices (E6 et E7) provoquée par cet évènement favorise la prolifération tumorale et le développement des tumeurs

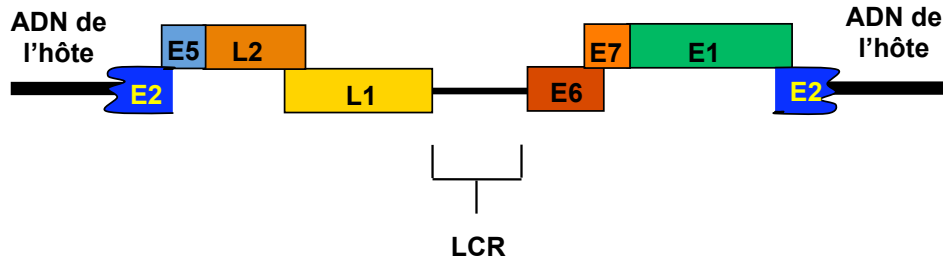


Figure I.5 – Intégration de l'ADN viral dans le génome hôte

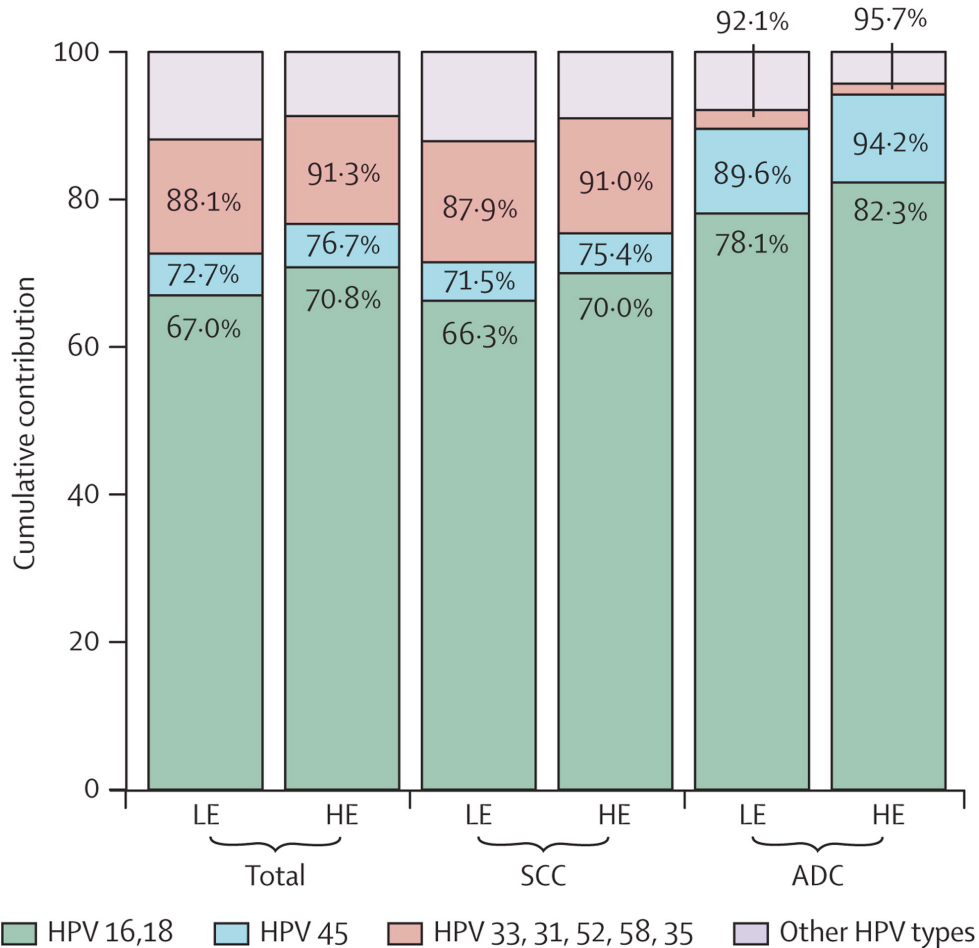
Schéma représentatif du génome linéaire de l'HPV (à HRM) en cas d'intégration dans le génome hôte, la rupture se fait dans des régions préférentielles sur le gène E2, (adapté d'après [Dell et Gaston \[2001\]](#)).

cancéreuses ([Munoz \*et al.\* \[2006\]](#))(Figure I.5).

## 2 Pathologies liées aux HPV

### 2.1 Néoplasie Intraépithéliale et Cancer du col de l'utérus

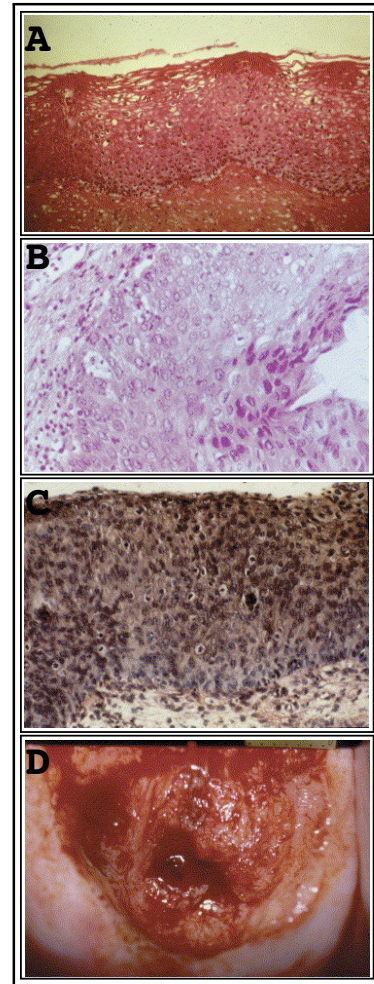
Le cancer du col de l'utérus représente la pathologie la plus grave parmi les maladies causées par les Papillomavirus. Il occupe la deuxième place sur la liste des cancers de femmes les plus fréquents dans le monde entier (plus de 500 000 nouveaux cas en 2005). L'incidence de ce cancer est très élevée dans les pays en voie de développement. Environ 83% des nouveaux cas de ce cancer surviennent dans ces pays. Dans ces régions la mortalité causée par ce cancer atteint des chiffres énormes. Sur 260 000 décès annuels associés au cancer du col de l'utérus près de 95% ont lieu dans les pays en voie de développement ([Ferlay J \[2004\]](#)). Le facteur étiologique principal du cancer du col de l'utérus est l'infection par les Papillomavirus : environ 99% des cas sont associé aux HPV à HRM ([McLaughlin-Drubin et Munger \[2009\]](#)), principalement les types HPV16 et HPV18 pour 70% des cas et HPV 45, 33, 31, 52, 58, 35 pour 20% des cas ([de Sanjose \*et al.\* \[2010\]](#)) (Figure I.6).



**Figure I.6 – Contribution relative cumulée des HPV retrouvés dans les tumeurs du Cancer utérin**

Dans cette figure : l'estimation minimale (LE) et maximale (HE) de la contribution des types HPV retrouvés dans les lésions de deux types du cancer du col utérins, l'ADC : Adénocarcinome (épithélium glandulaire) et le SCC : Carcinome épidermoïde (épithélium malpighien stratifié). Environ 90% des cas sont associés aux 8 types HPV (16, 18, 45, 33, 31, 52, 58, 35) dont 70% est représenté par 16 et 18. Dans les 10% qui reste on trouve les types 26, 30, 61, 67, 69, 82 et 91(Other HPV types). Ces statistiques sont effectuées à une échelle mondiale ( 22661 échantillons obtenues à partir de 14249 femmes venues de 38 pays) (d'après [de Sanjose et al. \[2010\]](#)).

Les lésions précancéreuses de Néoplasies intraépithéliales (NIE) qui constituent les précurseurs du cancer du col de l'utérus se classent en trois grades selon l'ampleur de la région touchée de l'épithélium : Le stade I (NIE I) touche seulement le tiers basal de l'épiderme, capable d'évoluer pour devenir un stade II (NIE II) qui touche deux tiers de l'épiderme. Ce dernier à son tour peut devenir un (NIE III) qui affecte l'intégralité de l'épithélium (Figure I.7). La Néoplasie interépithéliale peut se transformer à échéance en cancer (Carcinome in situ : CIS) ou régresser pour revenir à l'état épithélial normal. Le risque de progression d'une lésion vers un Carcinome est de 7% en cas de présence d'une infection persistante des HPV à HRM, alors qu'il est quasi nul en cas d'une infection transitoire (Ho *et al.* [1995]; Nobbenhuis *et al.* [1999]; zur Hausen [2000]). L'évolution des lésions utérines vers un cancer est généralement lente : 10 à 20 ans après l'infection causative (Barron et Richart [1981]). Bien que le virus HPV est nécessaire pour causer ce cancer, il n'est pas suffisant : d'autres facteurs étiologiques sont aussi nécessaires pour le développement d'une infection virale vers un cancer. Il s'agit du tabac, du nombre élevé des partenaires sexuels, de l'utilisation prolongée des contraceptions hormonales, de l'immunodépression, de la prédisposition génétique et de la présence d'une co-infection avec certaines autres maladies sexuellement transmissibles comme le VIH (Munoz *et al.* [2006]). Chez l'homme, les cancers génitaux sont beaucoup moins fréquents que chez la femme, mais certaines infections HPV à HRM peuvent provoquer un Cancer in situ génital masculin (Solivan *et al.* [1990]).



**Figure I.7 – Aspects histologiques des lésions précancéreuses**

Cette figure présente les lésions précancéreuses susceptibles à développer un cancer du col de l'utérus : **A** : NIE de grade I, **B** : NIE de grade II, **C** : NIE de grade III, **D** : Cancer invasif (Douvrier S [2004]).



### 2.2 Autres infections des muqueuses anogénitales

#### 2.2.1 Condylomes ano-génitaux

Les condylomes sont des excroissances anogénitales bénignes très fréquentes avec un aspect digitiforme ou en forme de chou-fleur (**Condylomes extra-génitaux acuminés**) ou un aspect plan (**Condylomes plans**).

### 2.3 Infections des muqueuses extra-génitales

#### 2.3.1 Papillomes buccaux

Ce sont des proliférations de la muqueuse buccale. Elles peuvent être à la base, des verrues vulgaires ou des condylomes acuminés transmis par autoinoculation ou par transmission sexuelle.

#### 2.3.2 Papillomatose orale floride

C'est un carcinome verruqueux au niveau des muqueuses linguale, jugale ou gingivale. Les types HPV responsables sont HPV 6, 11 et 16.

### 2.4 Infections cutanées

#### 2.4.1 Verrues

**Verrues vulgaires :** Ce sont des verrues très fréquentes et siègent préférentiellement sur le dos des mains et des doigts. Elles sont dues le plus souvent aux HPV02 et parfois au HPV1 ou HPV4. Certaines formes de verrues vulgaires observées exclusivement chez les bouchers sont provoquées par les virus HPV07.

**Verrues plantaires :** Elles se distinguent en deux types : les Myrmécies causées par HPV1 et les verrues en mosaïque causées par HPV02.

**Verrues filiformes :** Elles sont à la base des verrues vulgaires mais logées au niveau de la région du visage et du cou dotées d'un aspect filiforme.

**Verrues planes :** Les verrues planes sont des verrues à surface lisse, qui se trouvent préférentiellement sur le visage, le dos des mains, les avant-bras et sur les genoux. Elles sont causées par les HPV 3, 10 ou 23.

### 2.4.2 Epidermodysplasie verruciforme et Carcinome épidermoïde

Epidermodysplasie verruciforme est une maladie rare également appelée maladie de 'Lutz-Lewandowsky'. C'est une Génodermatose (affectation cutanée transmise génétiquement) caractérisée par un déficit immunitaire envers une vingtaine de types de HPV (HPV5, 8, 20... etc). Cette maladie peut se développer en cancer surtout dans les régions du corps exposées aux radiations solaires, ce qui évoque le rôle des radiations UV dans ce processus cancérogène. Cette coopération UV-HPV est observée aussi chez certains cancers épidermoïdes et basocellulaires dans lesquels l'ADN du type HPV 16 à été mis en évidence ([Douvier S \[2004\]](#)).

## 2.5 Diagnostic et thérapie

### 2.5.1 Dépistage

La mise en évidence d'une infection par HPV est possible via plusieurs techniques appliquées sur des échelles histologiques, cellulaires ou moléculaires.

#### **Diagnostic cyto-histologique**

L'étude cyto-histologique des frottis ou biopsies des lésions soupçonnées par colposcopie visent à dévoiler des aspects caractéristiques des cellules infectées par HPV. Il s'agit des cellules témoins de l'activité virale avec chromatine irrégulière et vacuole volumineuse intra-cytoplasmique (Koïlocytes), ou des cellules bi- ou pluri-nuclées, témoins de multiplication continue et perte de capacité de différenciation.

#### **Détection des virions**

En cas d'une importante réplication, les virions HPV sont visualisables via la microscopie électronique, mais cette méthode ne permet pas la distinction entre les types viraux.

#### **Examen sérologique**

L'analyse de la réponse immunitaire contre les pseudoparticules virales assemblées à partir des protéines structurales L1 et L2 (VLP) représente une bonne méthode de diagnostic qui permet l'identification spécifique malgré la présence de réactions croisées entre génotypes proches.

#### **Détection de l'ADN viral**

Cette technique est indispensable en cas de l'intégration du génome viral dans le génome hôte, ou de l'infection latente (absence de multiplication virale). Il s'agit en fait de techniques d'hybridation moléculaires dont le principe repose sur l'utilisation d'une amorce marquée et complémentaire d'une séquence génomique virale en vue d'une détection quantitative et qualitative de l'ADN viral (PCR quantitative, Southern-Blot, Dot blot ...).

### 2.5.2 Traitements

Les techniques de traitements des tumeurs bénignes ou malignes causées par les HPV sont aussi diversifiées que les pathologies visées. Il s'agit des traitements locaux par antiviraux, antimitotiques ou même par des interférons  $\alpha$  ou  $\beta$  et des traitements chirurgicaux à l'aide de la cryothérapie, l'électrocoagulation ou l'exérèse... Parfois le recours à des traitements systémiques (à dissémination généralisée dans tout l'organisme) s'avère nécessaire, en particulier chez les immunodéprimés en cas d'infections récidivantes et diffuses. On peut utiliser dans ces cas les rétinoïdes caractérisés par leur capacité à favoriser la différenciation et à réduire la prolifération cellulaire.

### 2.5.3 Vaccinations

Les deux vaccins prophylactiques utilisés actuellement contre les HPV ont été développés à partir de pseudoparticules de la protéine L1 (VLP-L1). La conception de ces virus est inspirée du fait que les pseudo-virus dépourvus de matériel génétique sont à la fois non nocifs pour l'organisme et aussi immunogènes que les virus naturels. Le premier vaccin est commercialisé par la société américaine Merck et nommé Gardasil. Il est dirigé contre les types HPV 6, 11, 16 et 18 (Villa *et al.* [2005]). Le deuxième vaccin nommé Cervarix est commercialisé par la société anglaise GlaxoSmithKline et offre une protection contre les types HPV 16 et 18. L'efficacité de ces vaccins est d'environ 100% en cas de non infection avant l'immunisation vaccinale. Ils sont destinés principalement aux jeunes filles entre 11 et 14 ans (avant l'activité sexuelle). La vaccination chez les hommes reste à discuter (Begue *et al.* [2007]). En terme de santé publique une telle vaccination prendrait tout son sens puisque les hommes, qui sont porteurs sains des Papillomavirus, ont plus de partenaires sexuels en moyenne que les femmes et constituent donc le vecteur majeur de l'épidémie. Certaines expérimentations ont été faites dans le but de développer une vaccination thérapeutique pour agir sur l'immunité locale des cellules attaquées par HPV contre les protéines virales E6 et E7. Ces recherches sont encore préliminaires (Brinkman *et al.* [2007]; Grasso *et al.* [2012]).

---

---

# Chapitre II

---

## Caractéristiques structurales et fonctionnelles des protéines des HPV

### 1 Les protéines tardives : L1 et L2

#### 1.1 La protéine L1

La protéine L1 est une petite protéine de 55 KDa qui s'auto-assemble en pentamères pour former les capsomères de la capsid virale.

**Monomère de L1 :** La structure cristallographique de L1 du type HPV16, montre que l'architecture monomérique de cette protéine est composée de trois zones de repliement distinctes. Le core (cœur) de la protéine adopte un repliement en forme de sandwich bêta (Jelly roll) (Figure II.1). Ce repliement est bâti par les 362 résidus de la région N-terminale (20-382) sur un total de 504 résidus. La partie C-terminale (383-474), majoritairement en hélice, est repliée sous forme de deux courtes hélices alpha (h2 et h3) suivies par une autre plus longue (h4). Cette deuxième zone constitue l'interface d'interaction avec les autres monomères. Le court brin bêta (J) et l'hélice alpha (h5) qui se trouvent en aval, rejoignent le core de la protéine pour consolider le repliement en sandwich bêta ; le brin (J) s'ajoute aux feuillets bêta (CHEF), et l'hélice (h5) se loge dans la cavité hydrophobe que forme le pentamère et interagit avec les résidus du feuillet bêta (BIDG). La troisième zone est constituée par les 31 derniers résidus. C'est une région de désordre, riche en acides aminés basiques et orientée à l'intérieur de la particule (Chen *et al.* [2000]). Ce segment qui contient le Signal de Localisation Nucléaire (NLS), interagit probablement avec l'ADN viral (Zhou *et al.* [1991]).

**Interactions intra-pentamériques de L1 :** La structure pentamérique de L1 présente une forme conique, semblable à un petit cratère dont le contour est formé par les cinq copies de L1, le

## Chapitre II. Caractéristiques structurales et fonctionnelles des protéines des HPV

---

cratère a un diamètre d'environ 14 Å. Les liaisons intra-pentamériques sont dues à des contacts directs entre les chaînes principales des polypeptides. Le brin bêta extérieur (G) du feuillet (BIDG) (indiqué par G1 rouge : Figure I.9) vient s'allonger le long du feuillet bêta (CHEF) du monomère voisin. De plus le chevauchement des boucles participe aussi à l'interaction, la boucle (H-I) s'intercalant entre les deux boucles (F-G et E-F) du monomère voisin ([Chen \*et al.\* \[2000\]](#)) (Figure II.1).

**Interactions inter-pentamériques de L1 :** La similarité entre la structure du virion des HPV et celui des Polyomavirus, que nous avons évoquée dans le chapitre précédent, se retrouve à l'échelle monomérique et pentamérique. L'interaction entre les pentamères se fait grâce aux liaisons hydrophobes entre l'hélice alpha (h4) d'un pentamère et les hélices alpha (h2 et h3) de son voisin ([Bishop \*et al.\* \[2007\]](#)) (Figure II.1).

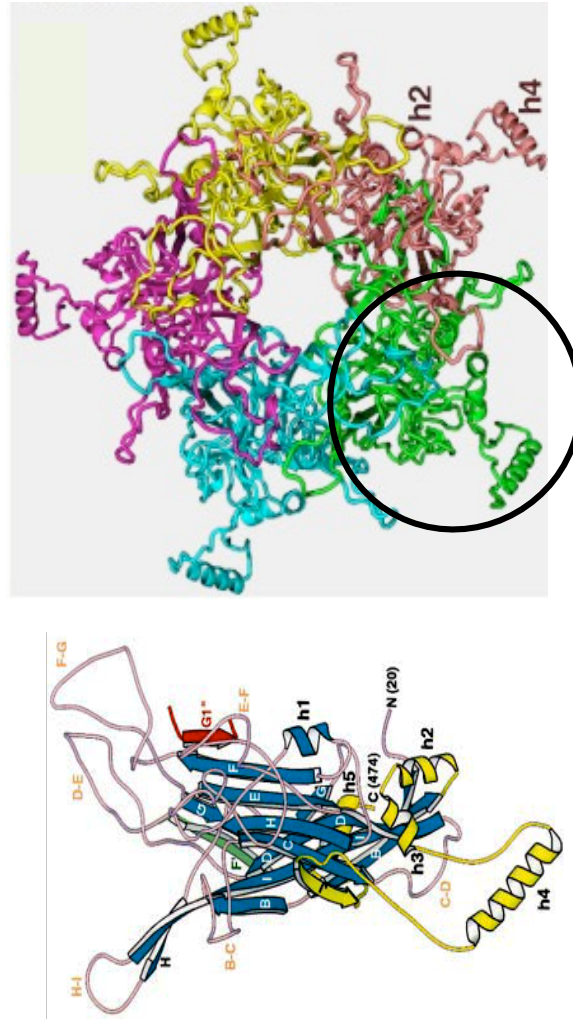


Figure II.1 – Structure de la protéine L1

A gauche : Structure monomérique de L1 : repliement en Jelly roll, les 5 hélices  $\alpha$ , (h1-h5) les 9 feuillets  $\beta$  (BCDEFGHIJ), et les boucles (B-C, C-D, D-E, E-F, F-G, H-I). A droite : Structure pentamérique de L1 (d'après [Chen et al. \[2000\]](#)).

La fonction principale de cette protéine - comme déjà vu dans le chapitre précédent - est l'encapsulation du virus, à l'aide de la protéine L2, mais elle seule est nécessaire et suffisante pour l'assemblage des pseudoparticules virales. Certains auteurs ont fait des expérimentations qui montrent un rôle important de L1 dans l'étape de l'entrée du virus via l'attachement de ses particules sur les protéoglycanes de sulfates d'héparane (HSPGs) (Rommel *et al.* [2005]; Shafti-Keramat *et al.* [2003]). Selon une étude parue en 2009 ces polysaccharides jouent un rôle similaire chez un autre modèle viral. Elles constituent en fait, sur les spermatozoïdes, un récepteur du virus VIH (Ceballos *et al.* [2009]).

### 1.2 La protéine L2

La protéine mineure de capsid L2 est une protéine de 75 KDa. Elle joue un rôle complémentaire dans l'assemblage de la capsid virale (12 copies L2 vs 72 pentamères de L1(360 copies)). Il a été suggéré que L2 constitue un pont entre la capsid et le génome viral, favorisant l'assemblage de l'ADN viral dans la capsid (El Mehdaoui *et al.* [2000]). Certaines études montrent un rôle présumé de L2 dans le transport du virus pendant ou après l'infection via l'interaction de L2 avec l'actine  $\beta$  (Yang *et al.* [2003]) ou avec la dynéine (complexe protéique associé aux microtubules) (Florin *et al.* [2006]). Le fait que les épitopes de L2 sont plus conservés que ceux de L1 a amené certains chercheurs à proposer l'ajout de cette protéine dans les vaccins L1-VLP, avec pour but d'offrir une immunisation plus large en terme de types HPV visés (Alphs *et al.* [2008]).

## 2 Les protéines précoces non transformatrices : E1, E2 et E4

Ces trois protéines n'ont pas la capacité à induire l'immortalisation des cellules infectées.

### 2.1 La protéine E1

La protéine E1 (environ 70 KDa) constitue avec la protéine E2 le complexe d'initiation de réplication de l'ADN viral. Cette protéine est composée de trois domaines fonctionnels : le domaine N-terminal dont le rôle est peu étudié, le domaine central qui constitue le domaine de reconnaissance de l'ADN (Origin Binding Domain [OBD]) et le domaine C-terminal qui assure une fonction Hélicase/ATPase et porte aussi une zone d'oligomérisation. Pour remplir sa fonction critique dans la réplication virale, E1 nécessite une coopération de la part de la protéine E2. Comme déjà expliqué dans le chapitre précédent, le dimère de E2 interagit dans une première étape à la fois avec l'ADN et avec E1 permettant ainsi la fixation de cette dernière sur l'origine de réplication de l'ADN viral. Lors d'une deuxième étape E1 s'homo-dimérise et effectue une première hydrolyse d'ATP (Figure II.2, A). Le dimère de E2 se détache après l'hydrolyse de l'ATP, E1 recrute des protéines chaperonnes (Hsp70, Hsp40) (Lin *et al.* [2002]) et adopte une forme fonctionnelle hexamérique qui catalyse la séparation de deux brins ADN à partir de l'origine de réplication dans le sens 3'-5' (Fouts *et al.* [1999]; White *et al.* [2001]) (Figure II.2, B et C). L'énergie nécessaire pour cette activité est à nouveau fournie par l'activité ATPase du domaine C-terminal (Seo *et al.* [1993]).

**Structure cristallographique de E1 :** La protéine E1 est très similaire à l'antigène grand-T des polyomavirus, en terme de séquence, structure et fonction (Enemark *et al.* [2002]; Schuck *et Stenlund* [2005]). La structure cristallographique de l'hexamère E1 (Enemark *et al.* [2002]) montre que les protomères de cette dernière forment un tube de 13 Å de diamètre entourant un des deux brins d'ADN, l'autre brin étant encerclé par un deuxième hexamère. La structure cristallographique du domaine central (OBD) (Enemark *et al.* [2000]), montre que celui-ci est composé de 4 hélices  $\alpha$  responsables des interactions non spécifiques avec l'ADN, et d'une boucle qui porte les résidus (conservés chez tous les PVs) partenaires de la zone spécifique de reconnaissance de E1 sur l'ADN (Figure II.2).



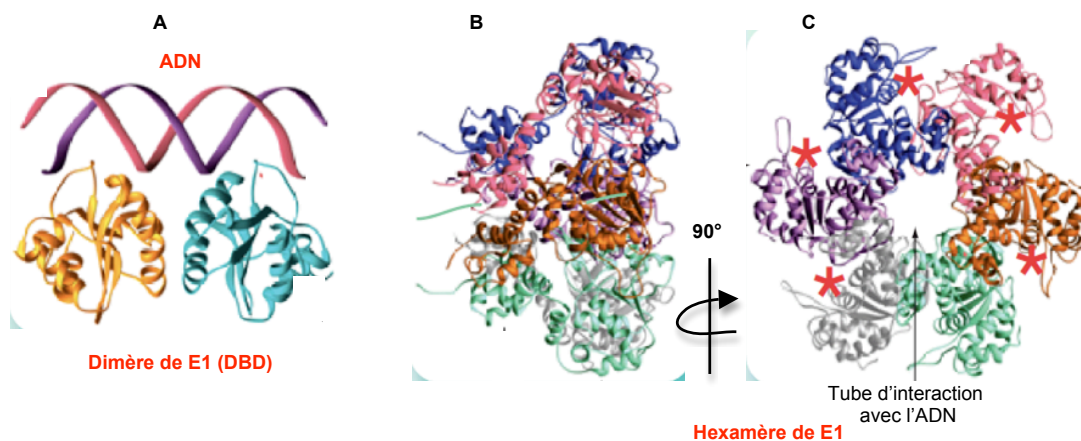


Figure II.2 – Structure de la protéine E1

A : Structure montrant les deux sous-unités du dimère du domaine OBD E1 (orange et cyan) attachées à l'ADN bicaténaire (rouge et violet) via leurs boucles de reconnaissance. B et C : Deux vues différentes de la structure fonctionnelle (hexamérique) de E1 (séparée par  $90^\circ$  par rapport à l'axe Y), montrant le tube de fixation de l'ADN, et les locus de fixation d'ATP (astérisques rouges) (d'après Hegde [2006]).

### 2.2 La protéine E2

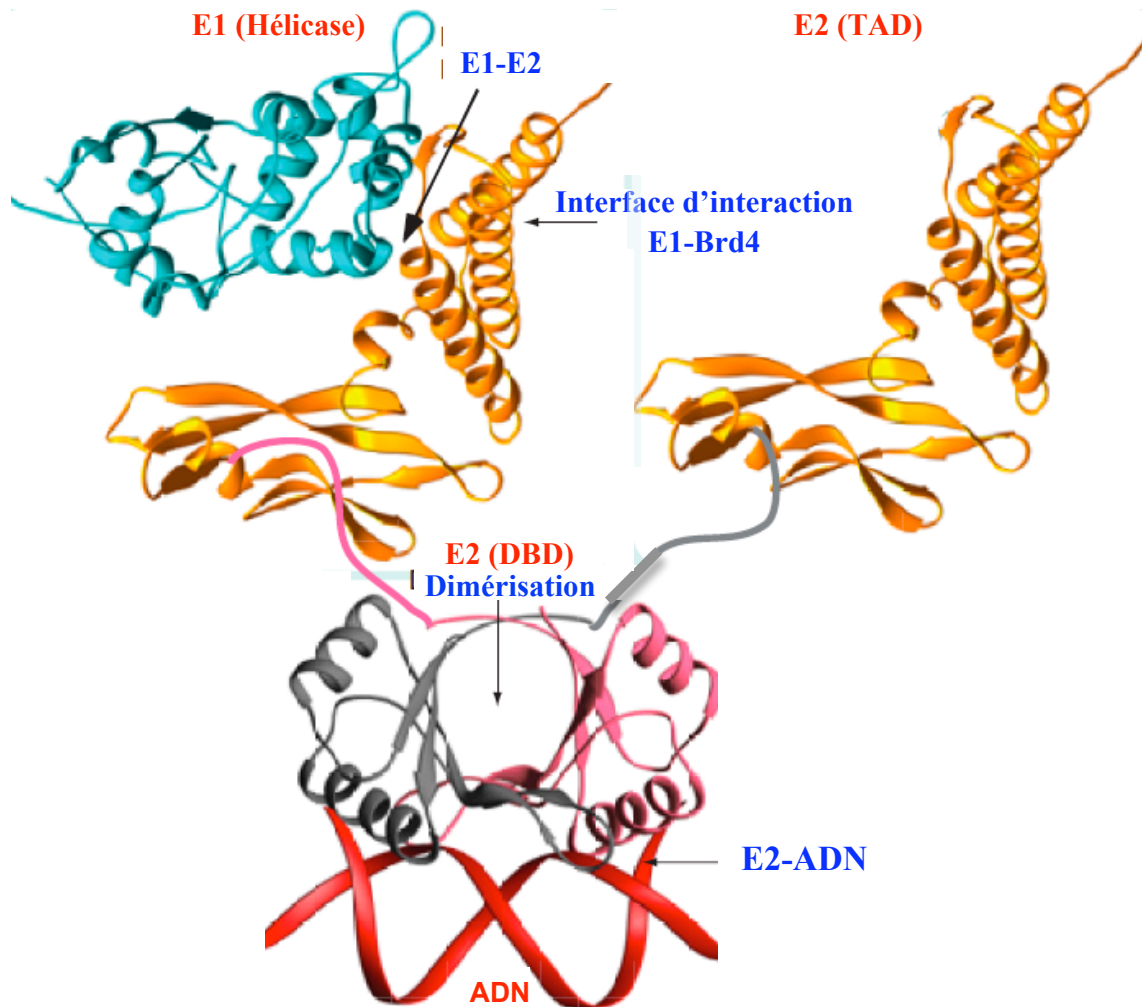
La protéine E2 (50 KDa, 365 résidus) est l'une des protéines de HPV les plus étudiées, elle est le facteur de transcription viral régulant le cycle des HPV. Elle régule la transcription, la réplication et la ségrégation de l'ADN viral. Il a été suggéré que cette protéine contrôle la prolifération cellulaire par régulation négative de E6 et E7 ainsi que des protéines tardives (Nishimura *et al.* [2000]). En cas de faible expression, E2 agit comme un activateur de transcription en interagissant avec une forte affinité sur une première catégorie de sites de l'ADN viral et en cas de forte expression elle agit comme un répresseur de transcription en interagissant avec une moindre affinité avec une deuxième catégorie de sites, et en bloquant ainsi la reconnaissance de ces sites par certains facteurs de transcription (Bouvard *et al.* [1994]; Steger et Corbach [1997]). E2 est composée de deux domaines structuraux et fonctionnels séparés par une zone charnière flexible. Le domaine N-terminal (200 résidus) représente le domaine de trans-activation (TAD) tandis que le domaine C-terminal (90 résidus) médie l'interaction avec l'ADN (DBD) (Kurg *et al.* [2006]) (Figure II.3).

**Domaine de trans-activation (TAD) :** Le domaine TAD comporte en lui même deux régions structurales; la partie N-terminale est formée par trois hélices  $\alpha$  tandis que la partie C-terminale est repliée sous forme de trois feuillets  $\beta$ . L'orientation de ces deux composantes structurales peut varier d'un type HPV à un autre (Antson *et al.* [2000]). Nous avons vu plus haut que la protéine E2 interagit avec E1 pour former le complexe d'initiation de réplication. Cette interaction se fait précisément entre la partie C-terminale de E1 (hélicase/ATPase) et le domaine N-terminal de E2 (TAD) (Abbate *et al.* [2004]). Ce même domaine est aussi capable de lier le domaine C-terminal de la protéine Brd4 associée aux fuseau mitotique, ce qui permet la ségrégation équitable du génome viral et le maintien des formes épisomales de l'ADN viral (Figure II.4) (You *et al.* [2004]).

**Domaine d'interaction avec l'ADN (DBD) :** Le domaine DBD (environ 90 résidus) reconnaît sous sa forme dimérisée l'ADN via le consensus ACCN(6)GGT (Androphy *et al.* [1987b]). Les structures de ce domaine en complexe avec l'ADN chez plusieurs types de Papillomavirus ont été publiées et montrent toutes le même repliement du domaine dimérisé. Il s'agit - pour chaque protomère - de quatre régions  $\beta$  antiparallèles et deux hélices  $\alpha$  dont une est impliquée dans un contact direct avec l'ADN (Harris et Botchan [1999]; Hegde et Androphy [1998]; Hegde *et al.* [1992]; Hooley *et al.* [2006]; Liang *et al.* [1996]).

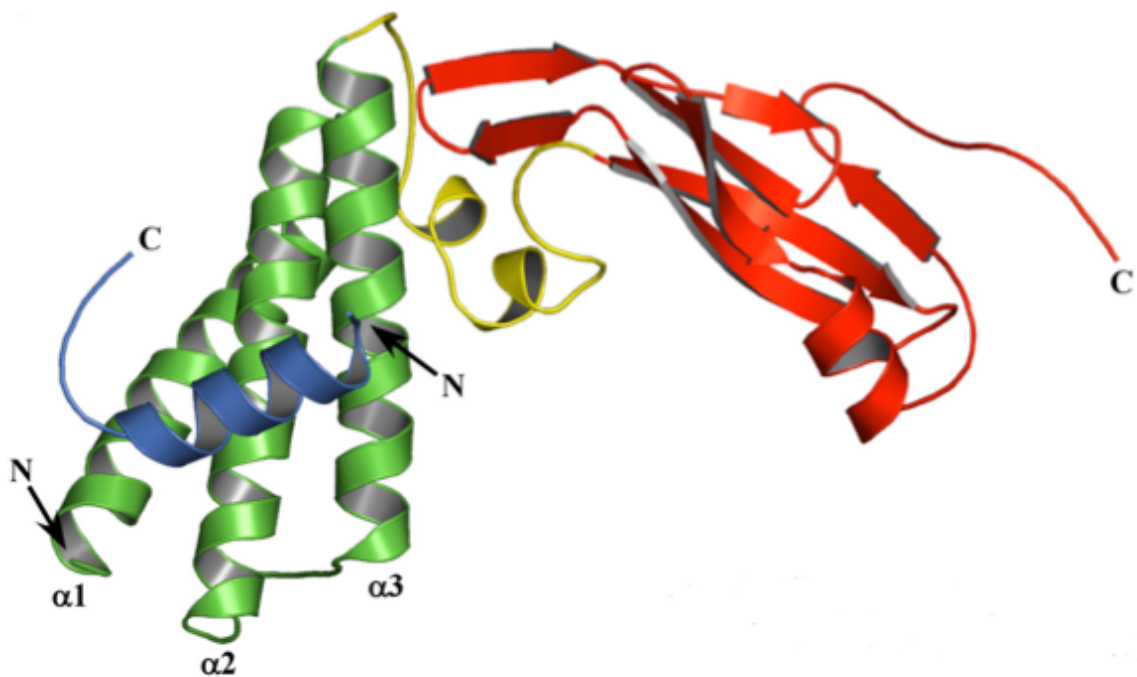
**Zone inter-domaines :** La zone inter-domaines de E2 est identifiée comme une zone non repliée, malgré que certains auteurs proposent que cette région a un repliement plutôt hélicoïdal, modulable par phosphorylation ([Garcia-Alai \*et al.\* \[2006\]](#)). Cette région de E2 constitue en effet une zone de phosphorylation riche en résidus P, E, S, T. Les régions possédant une telle composition constituent parfois un signal de dégradation, éventuellement dépendante de la phosphorylation, par la voie du protéasome, ce qui semble se produire pour E2 ([Penrose \*et al.\* \[2004\]](#)) (Figure II.3).

## II.2 Les protéines précoces non transformatrices : E1, E2 et E4



**Figure II.3 – Structure du complexe trimérique E1-E2-ADN**

Une représentation "cartoon" de l'interaction du domaine DBD-E2 dimérisé avec l'ADN (rouge) via ses deux sous-unités (gris et magenta). Un des 2 domaines de transactivation (TAD-E2) du dimère E2 (orange) interagit avec le domaine hélicase de E1 (cyan). Cette figure montre aussi la région d'interaction de E2 avec son partenaire cellulaire Brd4 (d'après [Chen \*et al.\* \[2000\]](#)).



**Figure II.4 – Interaction E2-Brd4**

Une représentation "cartoon" de l'interaction du domaine de transactivation (TAD) de la protéine virale E2 via ses trois hélices  $\alpha$  (vert) ( $\alpha1, \alpha2$  et  $\alpha3$ ) avec la région C-terminale de la protéine cellulaire Brd4 (Bleu) (d'après [Abbate et al. \[2006\]](#)).

### 2.3 La protéine E4

E4 est une petite protéine (17 KDa), exprimée sous forme de fusion E1–E4 entre le produit d’expression de l’ORF E4 et les 5 premiers résidus N-terminaux de l’ORF E1. Ses ARNm sont très abondants dans les cellules infectées par HPV ([Doorbar \*et al.\* \[1986\]](#)). La fonction de cette protéine n’est pas encore claire. Il semblerait qu’elle ne soit requise ni pour la persistance de l’ADN viral ni pour la transformation des cellules infectées ([Neary \*et al.\* \[1987\]](#)). Cependant, l’expression de E4 est corrélée avec la réplication de l’ADN viral ce qui suggère qu’elle est peut-être impliquée dans le processus d’amplification génomique ([Nakahara \*et al.\* \[2005\]](#)). De plus, comme déjà mentionné au chapitre premier, il a été proposé que la protéine E4 contribue à la libération des virus en perturbant l’intégrité de la kératine ([Wang \*et al.\* \[2004\]](#)) et en affectant l’assemblage de l’enveloppe cornée ([Lehr \*et al.\* \[2004\]](#)). Il semble aussi que E4 puisse favoriser un arrêt du cycle cellulaire à la phase G2 ([Knight \*et al.\* \[2004\]](#); [Longworth et Laimins \[2004b\]](#)). En cas d’intégration de l’ADN viral dans le génome hôte, l’ORF de E4, qui est incluse sous forme d’un décalage de phase de lecture, dans celle de E2, est perdue au cours de l’étape de linéarisation ([Romanczuk et Howley \[1992\]](#)).

## 3 Les protéines précoces transformatrices : E5, E6 et E7

Contrairement aux protéines E1, E2 et E4 ces trois protéines ont un pouvoir oncogène et sont capables d’induire la multiplication continue et l’immortalisation des cellules infectées. Cependant, il faut noter que E5 a un pouvoir oncogène faible par rapport à E6 et E7 ([Valle et Banks \[1995\]](#)).

### 3.1 La protéine E5

Cette petite protéine (9,4 KDa) très hydrophobe, considérée comme relativement peu oncogène, est composée de deux domaines fonctionnels : le domaine N-terminal est responsable de la localisation de E5 au niveau des membranes des endosomes et dans la membrane plasmique ([Schlegel \*et al.\* \[1986\]](#)). Le domaine fonctionnel C-terminal a un potentiel de transformation des cellules ([Burkhardt \*et al.\* \[1987\]](#); [Sparkowski \*et al.\* \[1996\]](#)). Le potentiel oncogène de E5 est considéré comme faible à cause de l’absence de cette dernière dans des carcinomes lorsque le génome viral est intégré dans le génome hôte ([Schwarz \*et al.\* \[1985\]](#); [Yee \*et al.\* \[1985\]](#)). Cependant, des copies épisomales de l’ADN viral sont parfois observées dans certains cas de carcinomes, et la capacité transformatrice de E5 a été prouvée chez des souris transgéniques avec ou sans traitement d’oestrogène ([Genther Williams \*et al.\* \[2005\]](#); [Maufort \*et al.\* \[2010\]](#)). E5 augmente

la réponse des kératinocytes aux facteurs de croissance épidermiques (EGF), en favorisant la phosphorylation de son récepteur (EGFR) (Straight *et al.* [1993]). Chez les Papillomavirus à HRM, E5 augmente le pH de l'appareil de Golgi, ce qui diminue la réponse immunitaire via l'inhibition de l'expression des molécules du complexe majeur d'histocompatibilité du type I (CMHI) (Ashrafi *et al.* [2006]). Certains auteurs ont montré que la protéine E5 coopère nécessairement avec E6 pour induire les Koilocytes (voir chapitre 1 : Diagnostic cyto-histologique), la délétion des derniers 20 résidus de cette protéine inhibe ces irrégularités cellulaires (Genther Williams *et al.* [2005]). Notons que la protéine E5 n'est pas codée par les génomes des virus  $\beta$ -HPVs (Klingelhutz et Roman [2012]).

### 3.2 La protéine E7

#### 3.2.1 Caractéristiques biochimiques et structurales de E7

La protéine E7 constitue avec E6 l'une des deux protéines les plus importantes dans le processus pathogène des HPV à HR. Ces deux oncoprotéines coopèrent pour immortaliser les cellules des keratinocytes humaines (Howley *et al.* [1989]), pour l'activation de la Télomérase (enzyme qui rallonge les chromosomes par des télomères durant la réplication) (Liu *et al.* [2008]) et pour provoquer l'instabilité génomique induisant le développement des tumeurs malignes (Duensing et Munger [2004]).

E7 est une protéine d'environ 98 résidus formée par 3 régions conservées CR1, CR2 et CR3 successives du côté N-terminal vers C-terminal. Ces trois régions de E7 présentent des similarités (relativement faibles) avec des régions de l'antigène grand T du virus SV40 et de E1A de l'adénovirus. Les régions CR1 et CR2 sont non repliées tandis que la région CR3 se replie sous forme dimérique stable et possède une structure secondaire composée -pour chaque protomère- par 2 hélices  $\alpha$  ( $\alpha_1$  et  $\alpha_2$ ), 2 brins  $\beta$  ( $\beta_1$  et  $\beta_2$ ) et une boucle entre  $\beta_1$  et  $\beta_2$  qui participe avec certains résidus de  $\alpha_2$  à la fixation du zinc. La région CR2 porte le motif LxCxE capable de lier la protéine du rétinoblastome (pRb) et les autres éléments de la famille des protéines à 'pocket domain' (p107 et p130) (Liu *et al.* [2006]) (Figure II.5).

## II.3 Les protéines précoces transformatrices : E5, E6 et E7

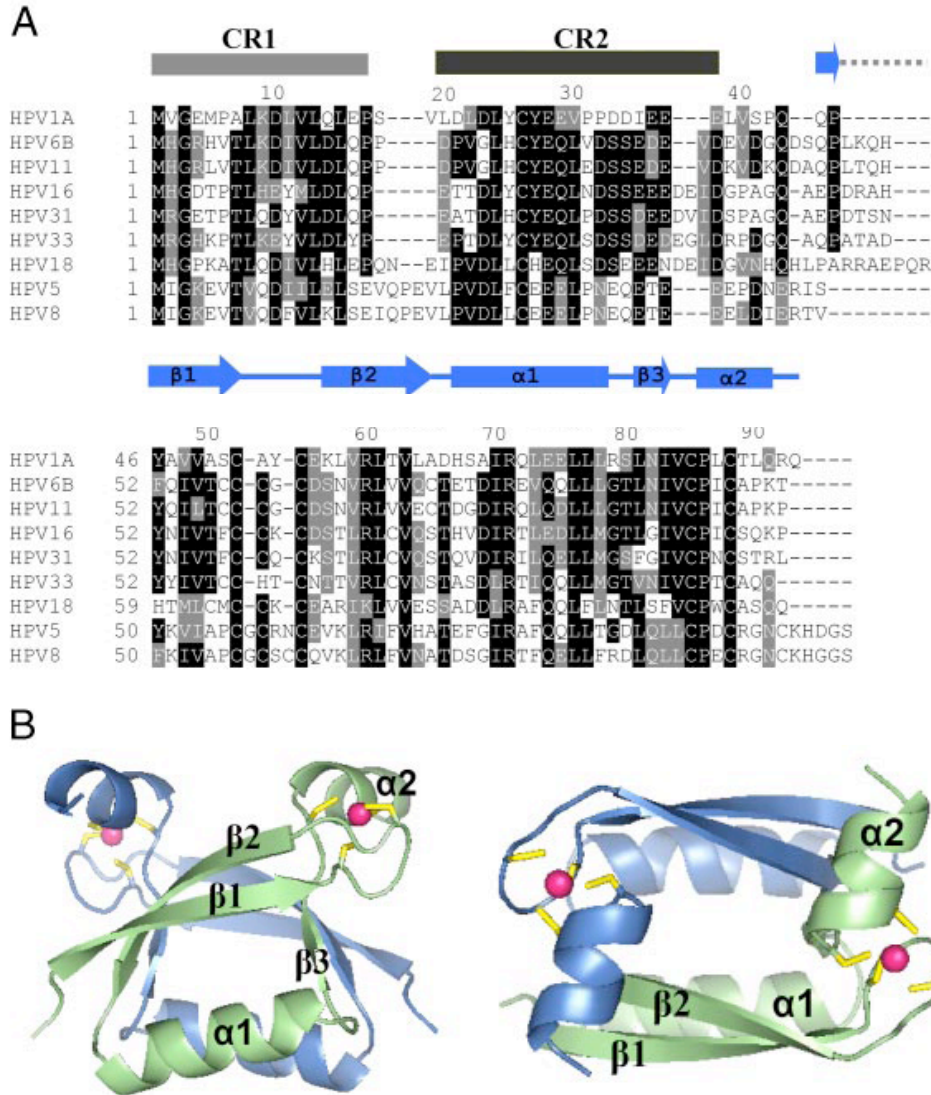


Figure II.5 – Alignement de séquences et structure de E7

**A** : Alignement de séquences et structure secondaire de la protéine E7 de différents types de HPV. L'ombrage noir indique les résidus très conservés et l'ombrage gris indique les résidus partiellement conservés. **B** : Structure cristallographique d'un dimère du domaine C-terminal à zinc de E7 (CR3) avec 2 vues différentes (un protomère coloré en bleu et l'autre en vert). (d'après Liu *et al.* [2006])



### 3.2.2 Activité biologique et partenaires cellulaires

La protéine E7 localisée principalement dans le noyau peut interagir avec plusieurs protéines cellulaires impliquées dans - entre autres- la prolifération et la différenciation cellulaires (Dyson *et al.* [1989]; Phelps *et al.* [1992]). Dans le tableau II.1 sont présentés les ligands de E7 selon l'état actuel de la littérature, répertoriés par domaine d'activité. Parmi les partenaires cellulaires de E7 les plus importants, on compte le facteur de transcription pRB. pRB est reconnu comme déjà dit par la région CR2 de E7. Cependant sa dégradation implique aussi la contribution de la région CR1. Cette association est plus forte chez E7 de HPV à HRM (Flores *et al.* [2000]).

La protéine pRb, exprimée par le gène suppresseur des tumeurs RB1 localisé sur le chromosome 13, a une fonction cruciale dans le cycle cellulaire normal. A l'entrée de la cellule en phase S, pRb subit une phosphorylation suscitée par le complexe cycline D-cdk4 et cycline E-cdk2, ce qui permet la libération d'un facteur de transcription de la famille E2F lié préalablement par pRb hypophosphorylée. La libération de E2F permet l'activation des gènes nécessaires à la prolifération cellulaire. En cas d'une infection à HPV, pRb est liée et dégradée par E7 via la voie de protéasome 26S, ce qui perturbe le contrôle de E2F. La libération de E2F rend le domaine d'activation transcriptionnelle de l'hétérodimère E2F/DP accessible permettant l'activation permanente des gènes nécessaires à l'entrée en phase S (Melillo *et al.* [1994]; Morris *et al.* [1993a]). En réponse à cet effet nocif de E7, la cellule augmente la transcription du gène codant pour p16 (inhibiteur des cdk4 et 6), associé à l'arrêt du cycle cellulaire en phase G1 (Ben-Saadon *et al.* [2004]). Malgré cette augmentation du taux de p16 et les signaux d'arrêt de croissance du TGF (Transforming growth factor), les cellules des tumeurs malignes continuent leur prolifération. Il a été suggéré que cette résistance aux signaux d'arrêt de croissance est due à la capacité de E7 à supprimer l'action inhibitrice des inhibiteurs des cdk, p21<sup>cip1</sup> (Funk *et al.* [1997]) et p27<sup>kip1</sup> (Zerfass-Thome *et al.* [1996]), exercée sur E2F (Munger *et al.* [2001]). D'un autre côté les E7 de HPVs à HRM et BRM sont capables de lier le facteur de transcription E2F6, élément du complexe E2F6/polycomb impliqué dans le mécanisme du 'destin cellulaire' (McLaughlin-Drubin *et al.* [2008]). De plus, E7 interagit avec son propre transcrit ARNm (Chinami *et al.* [1993]). Il a été également montré que les E7 de HPVs à HRM et BRM sont capables d'augmenter l'activité transcriptionnelle de HIF-1 (Hypoxia-Inducible Factor1), qui participe à l'angiogénèse en bloquant la liaison entre les déacétylase d'histone (HDAC) et le facteur de transcription HIF-1 $\alpha$  via une interaction directe entre ce dernier et E7 (Bodily *et al.* [2011]).

Très récemment, il a été démontré que E7 du type HPV 16 peut s'associer avec la protéine Miz-1 (Myc-interacting zinc-finger protein-1) impliquée dans le processus de l'arrêt du cycle

## II.3 Les protéines précoces transformatrices : E5, E6 et E7

Tableau II.1 – Liste de partenaires cellulaires de E7

ACTIVITE	PROTEINE	Références
<b>Régulation transcriptionnelle</b>	c-Jun, JunB, JunD et c-Fos	Antinore <i>et al.</i> [1996]
	E2	Gammoh <i>et al.</i> [2006]
	E2F1	Hwang <i>et al.</i> [2002]
	E2F6	McLaughlin-Drubin <i>et al.</i> [2008]
	FHL2	Campo-Fernandez <i>et al.</i> [2007]
	HDACs, Mi2 $\beta$	Longworth et Laimins [2004a], Brehm <i>et al.</i> [1999]
	HIF-1 $\alpha$	Bodily <i>et al.</i> [2011]
	Kinase d'histone H1	Davies <i>et al.</i> [1993]
	Complexe ISGF3 (p48)	Barnard et McMillan [1999]
	MPP2	Luscher-Firzlaff <i>et al.</i> [1999]
	P300	Bernat <i>et al.</i> [2003]
	pCAF (p/CAF)	Avvakumov <i>et al.</i> [2003], Huang et McCance [2002]
	Rep78	Hermonat <i>et al.</i> [2000]
	Skip	Prathapam <i>et al.</i> [2001]
	TAF-110	Mazzarelli <i>et al.</i> [1995]
	TAFII20, TAFII28, TAFII55, TAFII $\delta$ N135, TBP	Enzenauer <i>et al.</i> [1998]
	TBP	Massimi <i>et al.</i> [1996], Massimi <i>et al.</i> [1997]
	<b>Cycle cellulaire</b>	pRB, p107, p130, cyclin A
CENP-C		Yaginuma <i>et al.</i> [2010]
cyclin A/CDK2, cyclin E/CDK2		Nguyen et Munger [2008]
$\gamma$ -tubulin		Nguyen <i>et al.</i> [2007]
Miz-1		Morandell <i>et al.</i> [2012]
NuMA		Nguyen et Munger [2009]
p130		Zhang <i>et al.</i> [2006]
p21 <sup>cip1</sup>		Funk <i>et al.</i> [1997]
p27 <sup>kip1</sup>		Zerfass-Thome <i>et al.</i> [1996]
p33CDK2, cyclin A		Tommasino <i>et al.</i> [1993]
P600		Huh <i>et al.</i> [2005]
pRB		Wang <i>et al.</i> [2010c]
<b>Signalisation cellulaire</b>	I $\kappa$ B kinase complex	Spitkovsky <i>et al.</i> [2002]
	IRF-1	Park <i>et al.</i> [2000]
	IRF-9	Antonsson <i>et al.</i> [2006]
	PP2A	Pim <i>et al.</i> [2005]
	Smad 1- 4	Habig <i>et al.</i> [2006], Lee <i>et al.</i> [2002]
<b>Apoptose et sénescence cellulaire</b>	IGFBP-3	Santer <i>et al.</i> [2007], Mannhardt <i>et al.</i> [2000]
	PML	Bischof <i>et al.</i> [2005]
	Siva-1	Severino <i>et al.</i> [2007]
<b>Ubiquitination et dégradation</b>	Complexe cullin 2 ubiquitin ligase (S4) ATPase du protéasome 26S	Huh <i>et al.</i> [2007]
	USP1	Berezutskaya et Bagchi [1997]
		Malila <i>et al.</i> [1993]
<b>Métabolisme cellulaire</b>	Acid $\alpha$ glucosidase	Zwerschke <i>et al.</i> [2000]
	M2-PK	Mazurek <i>et al.</i> [2001]
<b>Acides nucléiques</b>	E7-ARN	Chinami <i>et al.</i> [1993]
<b>Immunité</b>	TAP-1	Vambutas <i>et al.</i> [2001]
<b>Modification post-transcriptionnelle</b>	Transglutaminase 2	Jeon <i>et al.</i> [2003]
	Casein kinase II	Barbosa <i>et al.</i> [1990]; Firzlaff <i>et al.</i> [1989]
<b>Chaperons</b>	hTid-1	Schilling <i>et al.</i> [1998]

Tableau récapitulatif des protéines et acides nucléiques capables de lier la protéine E7. Les partenaires cellulaires sont classés selon leurs domaines d'activité biologique. Les références bibliographiques sont citées dans la dernière colonne.

cellulaire via l'induction de l'inhibiteur des cdk, p21<sup>cip1</sup> dépendante de l'exposition aux radiations UV (Morandell *et al.* [2012]). E7 est capable aussi de lier d'autres protéines associées à divers domaines. Citons par exemple dans la sphère des voies de signalisation : le facteur de transcription qui joue un rôle important dans la régulation transcriptionnelle de l'interferon  $\beta$ , IRF1 (Park *et al.* [2000]), les protéines SMAD responsables de la transduction du signal de TGF $\beta$  vers le noyau ( Habig *et al.* [2006]; Lee *et al.* [2002]). Et dans la sphère de la régulation transcriptionnelles : E2F1, HDAC, P300, TBP... (Figure II.1).

### 3.3 La protéine E6

#### 3.3.1 Caractéristiques biochimiques et structurales de E6

E6 est la première protéine de HPV définie comme oncogène, elle occupe une place importante dans le processus d'infection et de détournement de la machinerie cellulaire par les virus HPV (Mantovani et Banks [2001]). E6 est une petite protéine (environ 151 résidus), qui possède deux domaines à doigts de Zinc (E6N et E6C). Les atomes de Zinc sont liés par des motifs à cystéines très conservés (Cys-x-x-Cys) (voir alignement, Figure II.6). Il faut noter ici qu'il a été prouvé que E6 à HRM peut exister sous un autre isoforme tronqué (E6\*) produit par épissage alternatif de l'ARNm (Doorbar *et al.* [1990]).

Selon les prédictions confirmées ultérieurement par les données structurales, le repliement général du E6C est composé de trois hélices  $\alpha$  ( $\alpha 3$ ,  $\alpha 4$ , et  $\alpha 5$ ), quatre brins bêtas ( $\beta 4$ ,  $\beta 5$ ,  $\beta 6$  et  $\beta 7$ ) et une pseudo hélice  $\alpha$ . Le repliement de E6N est composé de deux hélices alpha ( $\alpha 0$  et  $\alpha 1$ ) et trois brins  $\beta$  ( $\beta 1$ ,  $\beta 2$  et  $\beta 3$ ), Les deux domaines sont reliés par une hélice  $\alpha$  relativement longue ( $\alpha 2$ ) (voir structure secondaire, Figure II.6)(Nomine *et al.* [2006]; Zanier *et al.* [2012]). Cette section sera traitée plus amplement dans le chapitre suivant "Introduction au projet de recherche".

## II.3 Les protéines précoces transformatrices : E5, E6 et E7

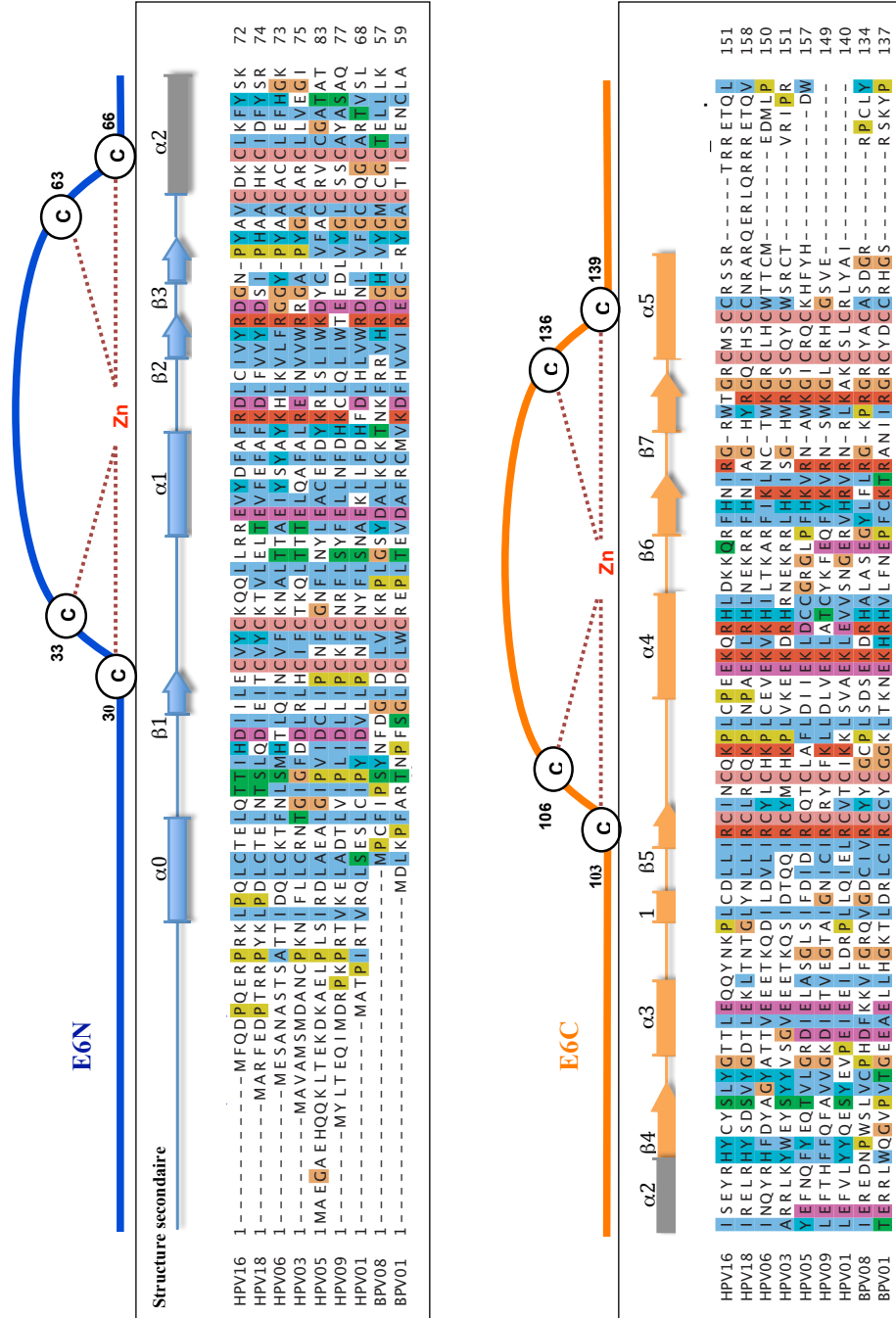


Figure II.6 – Alignements de séquences et structure secondaire de E6

Les séquences alignées des deux domaines de E6 ; E6N (bleu, en haut) et E6C (orange, en bas), appartenant aux 7 HPV (16, 18, 6, 3, 5, 9 et 1) et 2 BPV (8 et 1), avec la structure secondaire de HPV16 E6. Les résidus cystéines liant le zinc sont encerclés (séquences selon HPV16 E6 : 30, 33, 63, 66, 103, 106, 136 et 139).

### 3.3.2 Activité biologique et partenaires cellulaires

Selon l'état actuel de la littérature, E6 interagit avec environ une soixantaine des protéines cellulaires (Tableau II.2), impliquées dans la régulation de transcription, l'apoptose, la signalisation cellulaire ainsi que diverses autres voies cellulaires. Les interactions de E6 avec la plupart de ces protéines ont été identifiées dans des conditions de surexpression *in vivo*.

#### A- Médiateurs d'apoptose :

**A1- Apoptose dépendante de p53 :** La capacité de E6 des HPV à HRM à former un triple complexe avec l'ubiquitine ligase E6AP (E6 associated protein) et p53 (facteur de transcription pro-apoptotique souvent présenté comme le "gardien du génôme"), aboutissant à la dégradation de p53 par la voie d'ubiquitination, constitue une des activités de E6 les plus étudiées (Huibregtse *et al.* [1991]; Scheffner *et al.* [1993, 1990]). Cette interaction de E6 avec p53 (393 résidus), est de nature complexe, multifactorielle et implique plus d'un site d'interaction. Les E6 de HPV à HRM et BRM sont capables toutes les deux de lier un site d'interaction existant sur le domaine C-terminal de p53 via leurs domaines C-terminaux mais cette liaison n'est pas suffisante à elle seule pour induire la dégradation de p53 (Lechner et Laimins [1994]). Une autre interaction nécessaire pour la dégradation pourrait être médiée par le domaine N-terminal de E6 à HRM et le domaine "core" de p53 (Li et Coffino [1996]). Il semble que cette dualité des sites d'interaction E6-p53 existe aussi au niveau de l'interaction E6-E6AP qui nécessite apparemment plus d'une région de contact pour avoir une liaison effective selon certains auteurs (Huibregtse *et al.* [1993]; Nomine *et al.* [2006]; Pim et Banks [2010]). La dégradation de p53 facilite l'immortalisation cellulaire par blocage du processus d'apoptose dépendante de p53, ce qui permet la prolifération continue des cellules, l'accumulation des dommages d'ADN non réparés et la formation des tumeurs malignes (Dalal *et al.* [1996]). En plus de la fonction 'transformatrice' de E6, il a été montré que l'interaction E6-p53 joue un rôle important dans la réplication et le maintien du génome épisomal viral en cas d'une infection virale productive. Cet effet est commun aux deux catégories de HPV muqueux, qu'ils soient à HRM (HPV 16) (Park et Androphy [2002]) ou BRM (HPV 11) (Oh *et al.* [2004]); ce qui souligne l'importance de l'inactivation et pas nécessairement de la dégradation de p53 pendant une infection HPV.

**A2- Apoptose indépendante de p53 :** Des expériences faites sur des souris transgéniques qui n'expriment pas p53 ont montré que HPV16 E6 est capable d'inhiber l'apoptose non dépendante de p53 (Pan et Griep [1995]). En effet, E6 est capable de lier et dégrader la protéine pro-apoptotique Bak. Les interactions de Bak avec les E6 de HPV à HRM (Thomas et Banks

## II.3 Les protéines précoces transformatrices : E5, E6 et E7

Tableau II.2 – Liste de partenaires cellulaires de E6

ACTIVITE	PROTEINE	Références
Régulation transcriptionnelle	ADA3	Zeng <i>et al.</i> [2002]
	ATF3	Wang <i>et al.</i> [2010a]
	BARD1	Yim <i>et al.</i> [2007]
	BRCA1	Zhang <i>et al.</i> [2005]
	CBP/p300	Patel <i>et al.</i> [1999], Zimmermann <i>et al.</i> [2000]
	E6TP1	Gao <i>et al.</i> [1999], Gao <i>et al.</i> [2001], Gao <i>et al.</i> [2002]
	FOXK1/K2	Komorek <i>et al.</i> [2010]
	GADD34/PP1	Kazemi <i>et al.</i> [2004]
	Gps2	Degenhardt et Silverstein [2001b]
	MAML1 (LxxLL)	Brimer <i>et al.</i> [2012]
	Myc	Gross-Mesilaty <i>et al.</i> [1998]
	NFX1-91	Gewin <i>et al.</i> [2004]; Xu <i>et al.</i> [2008]
	PML	Guccione <i>et al.</i> [2004]
	TRIP-Br1	Gupta <i>et al.</i> [2003]
	Tsc2 (LxxLL)	Lu <i>et al.</i> [2004]
	Adhésion et polarité cellulaire	hDLG (PDZ)
hScrib (PDZ)		Nakagawa et Huibregtse [2000]
MAGI-1,2,3 (PDZ)		Glaunsinger <i>et al.</i> [2000], Thomas <i>et al.</i> [2001], Massimi <i>et al.</i> [2004]
Mupp1 (PDZ)		Lee <i>et al.</i> [2000]
PatJ (PDZ)		Storrs et Silverstein [2007]
PTPN3 (PDZ)		Jing <i>et al.</i> [2007]
TIP-1 (PDZ)		Hampson <i>et al.</i> [2004]
TIP-2 (PDZ)		Favre-Bonvin <i>et al.</i> [2005]
Apoptose	Bak	Thomas et Banks [1999]
	FADD	Filippova <i>et al.</i> [2004], Tungteakkhun <i>et al.</i> [2008]
	GRIM-19	Zhou <i>et al.</i> [2011]
	HIPK2	Muschik <i>et al.</i> [2011]
	p53	Werness <i>et al.</i> [1990]
	Procaspace 8	Filippova <i>et al.</i> [2007]
	TNFR1	Filippova <i>et al.</i> [2002]
	Signalisation cellulaire	IRF3 (LxxLL)
NHERF-1		Accardi <i>et al.</i> [2011]
PKA		Kühne <i>et al.</i> [2000]
PKN		Gao <i>et al.</i> [2000]
SMAD3		Mendoza <i>et al.</i> [2006]
TMEM87B		Li <i>et al.</i> [2008]
Tyk2		Li <i>et al.</i> [1999]
Ubiquitination, sumoylation et dégradation	CYLD	An <i>et al.</i> [2008]
	E6AP (LxxLL)	Huibregtse <i>et al.</i> [1993]
	HERC2	Vos <i>et al.</i> [2009]
	PIASy	Bischof <i>et al.</i> [2006]
	UBC9	Heaton <i>et al.</i> [2011]
	USP15	Vos <i>et al.</i> [2009]
Organisation et différenciation épithéliale	ERC-55 (LxxLL)	Chen <i>et al.</i> [1995]
	Fibulin-1	Du <i>et al.</i> [2002]
	Paxillin (LxxLL)	Tong et Howley [1997]; Tong <i>et al.</i> [1997], Vande Pol <i>et al.</i> [1998]
	Zyxin	Degenhardt et Silverstein [2001a]
Epigénétique	CARM1	Hsu <i>et al.</i> [2012]
	PRMT1	Hsu <i>et al.</i> [2012]
	SET7	Hsu <i>et al.</i> [2012]
Stabilité chromosomique	MCM7	Kukimoto <i>et al.</i> [1998]
	MGMT	Srivenugopal et Ali-Osman [2002]
	XRCC1	Iftner <i>et al.</i> [2002]
Trafic nucléaire et vésiculaire	Kap $\alpha 2$ , $\beta 1$ , $\beta 2$	Le Roux et Moroianu [2003]
	CAL (PDZ)	Jeong <i>et al.</i> [2007]
Acides nucléiques	ADN (Jonctions de Holliday)	Nomine <i>et al.</i> [2006]; Ristriani <i>et al.</i> [2001]
	ARNm	Bodaghi <i>et al.</i> [2009]

Tableau récapitulatif des protéines et acides nucléiques capables de lier la protéine E6. Les partenaires cellulaires sont classés selon leurs domaines d'activité biologique. Les références bibliographiques sont citées dans la dernière colonne. Lorsqu'il est connu, le mode d'interaction est indiqué : mode médié par LxxLL (indiqué par LxxLL en rouge) ou mode médié par les PDZ (indiqué par PDZ en rouge).

## Chapitre II. Caractéristiques structurales et fonctionnelles des protéines des HPV

---

[1999]) et de HPV à BRM (Underbrink *et al.* [2008]) sont dépendantes de l'ubiquitine ligase E6AP et provoquent la dégradation de Bak par voie protéasomique. Il a été également montré que E6 du type HPV 16 cause une augmentation de la transactivation du gène de la survivine, membre de la famille des gènes inhibiteurs d'apoptose (IAP). Le produit d'expression de ce gène anti-apoptotique contrôle non seulement l'apoptose mais aussi la division cellulaire (Borbely *et al.* [2006]). En plus de ses implications au niveau des protéines pro-apoptotiques et anti-apoptotiques, E6 est capable aussi d'empêcher la transmission du signal apoptotique via la liaison ou l'interférence avec des éléments des voies d'apoptose comme le récepteur du facteur de nécrose tumorale TNF R1, FADD (Fas-associated death domain) et procaspase 8 (Filippova *et al.* [2007, 2004, 2002]). Finalement il est à noter que des chercheurs chinois ont réalisé récemment des expérimentations qui montrent une fonction importante de la protéine GRIM-19 (sous-unité de NADH-ubiquinone oxidoreductase) dans le contexte d'une infection à HPV à HRM (HPV 18). GRIM-19 se montre capable de perturber le complexe triple E6-E6AP-p53 et de libérer p53 via une association entre le domaine N-terminal de GRIM-19 avec chacune des protéines E6AP et E6 qui protège p53 de la dégradation et provoque l'apoptose des cellules infectées (Zhou *et al.* [2011]). La capacité à empêcher l'apoptose et éviter l'arrêt du cycle cellulaire est plus forte chez E6 de HPV à HRM que chez E6 de HPV à BRM. Cette différence d'efficacité est peut être due au fait que E6 de HPV à HRM inactive plus fortement p53 en la dégradant (Elbel *et al.* [1997]).

**B- Régulation transcriptionnelle :** En plus des médiateurs d'apoptose, E6 peut lier aussi une quinzaine des protéines associées au contrôle de transcription. Outre la dégradation de p53 résultant de l'interaction triple E6-E6AP-p53, E6 est capable d'inhiber l'activité transcriptionnelle de p53 via d'autres mécanismes. En fait, E6 de HPV à HRM et BRM sont capables toutes les deux de lier d'une façon indépendante de E6AP, le complexe p300/CBP (P300/CREB Binding Protein) responsable de l'acétylation de p53. On pense que cette liaison empêche l'acétylation de p53 ce qui affecte l'affinité de p53 envers l'ADN endommagée et supprime sa capacité d'activer ses promoteurs de réponse (Thomas et Chiang [2005]; Zimmermann *et al.* [1999]). Selon des tests de double hybride, E6 à HRM interagit aussi avec Gps2, provoquant sa dégradation (Degenhardt et Silverstein [2001b]). Gps2 lie le complexe p300/CBP et facilite sa fonction à contrôler la régulation transcriptionnelle de p53 (Peng *et al.* [2000]). Donc, la liaison de E6 avec le complexe p300/CBP et avec Gps2 contribuent au blocage d'apoptose et à l'augmentation de prolifération cellulaire. D'autres auteurs évoquent un rôle présumé d'un autre partenaire cellulaire de E6 dans l'annulation de l'acétylation de p53. Il s'agit de la protéine ADA3 (Transcriptional adapter 3), élément du complexe PCAF impliqué dans l'acétylation des

### II.3 Les protéines précoces transformatrices : E5, E6 et E7

---

histones et co-activateur de l'activité transcriptionnelle dépendante de p53. Cette protéine est liée par E6 de HPV à HRM (Kumar *et al.* [2002]). Récemment, il a été montré que E6 du type HPV 23, type cutané responsable de certains cancers, peut interagir avec l'enzyme HIPK2 impliquée dans la régulation d'apoptose et acteur de la phosphorylation de p53, étape importante dans l'expression des gènes pro-apoptotiques dépendants de p53 (Muschik *et al.* [2011]). Dans le même contexte, il a été aussi montré récemment que les E6 de HPV à HRM et BRM interagissent, indépendamment de p53 et E6AP, avec l'acétyltransférase nucléosomique TIP60 (60 kDa Tat-interactive protein) impliquée dans l'acétylation des histones H4, et p53. Cette liaison affecte la capacité de TIP60 à contrôler la régulation positive par p53 des gènes responsables d'apoptose mais, curieusement, n'affecte pas la régulation positive par p53 des gènes responsables de l'arrêt du cycle cellulaire comme p21. Une autre fonction associée à cette liaison E6-TIP60 est l'amélioration de la capacité de TIP60 à réprimer les promoteurs précoces des HPV via l'association de cette dernière (TIP60) avec la protéine Brd4 (Jha *et al.* [2010]). Des analyses de double hybride ont montré que E6 à HRM lie la protéine E6TP1 (Gao *et al.* [2002, 2001, 1999]), qui présente une forte homologie avec les protéines activatrices de GTPase (GAPs) et connue pour avoir une fonction suppressive des tumeurs. D'une façon similaire (double hybride), il a été prouvé que E6 interagit avec la Tubérine, produit du gène suppressif des tumeurs Tsc2, bloquant ainsi son contrôle négatif de la prolifération cellulaire (Lu *et al.* [2004]). De plus, l'expression exogène de E6 active la Télomérase synergiquement avec E7. La capacité de E6 à activer la télomérase est exclusive à E6 de HPV à HRM et indépendante de p53 et ne se fait pas via le motif d'interaction avec les domaine à PDZ (Kiyono *et al.* [1998]; Klingelhutz *et al.* [1996]). En fait, E6 suscite une régulation positive de la transcription de TERT (Telomerase reverse transcriptase), la sous-unité catalytique de la Télomérase, en agissant près de E-Box (Enhancer Box) du promoteur TERT. Il semble aussi que E6 à HRM influence la transcription du TERT en interagissant avec le facteur de transcription c-myc en rendant ce dernier plus actif sur le promoteur de TERT. Cette interaction E6-c-myc requiert E6AP mais ne suscite pas la dégradation de c-myc (Veldman *et al.* [2003]). D'autres chercheurs ont montré que E6 peut lier et dégrader le déacétylase d'histone NFX1-91 qui constitue un répressif de la transcription de TERT. Finalement, E6 peut interagir directement avec la protéine TERT et changer sa localisation dans la cellule (Liu *et al.* [2009]). La capacité d'activation de la télomérase exclusive de E6 à HRM constitue une étape nécessaire pour l'immortalisation cellulaire (Kiyono *et al.* [1998]). Cependant son rôle dans le cycle de vie du virus n'est pas encore clair. Certains auteurs évoquent dans ce contexte que TERT présente une fonction autre que l'élongation de chromosomes durant la réplication, fonction qui consiste en son implication dans la voie de signalisation Wnt dont l'activation est un signe d'environnement favorable pour la réplication virale (Choi



## Chapitre II. Caractéristiques structurales et fonctionnelles des protéines des HPV

---

*et al.* [2008]). E6 de HPV à HRM et BRM lie aussi la protéine PML (Promyelocytic leukemia), impliquée dans la régulation transcriptionnelle et dans l'apoptose. Cette protéine associée à la matrice nucléaire constitue peut-être l'assise d'assemblage des capsides virales comme déjà vu dans la partie de cette introduction consacrée au cycle viral. Selon des études récentes, la protéine ATF3 (Activating transcription factor 3) joue un rôle important dans le mécanisme de défense cellulaire contre les HPV à HRM. Ce rôle est similaire à celui de GRIM-19 décrit un peu plus haut. Au fait, ATF3 perturbe le complexe trimérique de E6-E6AP-p53 sauvant p53 de la dégradation et permettant l'expression des gènes dépendants de cette dernière, ce qui déclenche l'arrêt du cycle cellulaire ou l'apoptose (Wang *et al.* [2010b]). Des tests d'immunoprécipitation, ont permis de prouver que E6 s'associe avec le suppresseur de tumeur BARD1 (BRCA1-associated ring domain protein 1) dans un complexe qui inactive mais ne dégrade pas BARD1 (Yim *et al.* [2007]). E6 lie aussi tout comme E7 la protéine BRCA1 (Breast Cancer 1 Gene) associée avec BARD1 et impliquée dans la régulation répressive de l'activité du récepteur  $\alpha$  à l'œstrogène (ER- $\alpha$ ). Cependant l'interprétation biologique de cette association n'est pas encore claire. La liaison E6-BRCA1 implique deux sites de reconnaissance sur BRCA1 et ne cause pas sa dégradation (Zhang *et al.* [2005]). HPV18 E6 recrute le complexe GADD34/PP1 (Protein phosphatase 1 regulatory subunit 15A / Serine-threonine-protein phosphatase) pour déphosphoryler la sous-unité d'initiation de traduction eIF2 $\alpha$  ( $\alpha$  subunit of eukaryotic translation initiation factor 2) ce qui permet à E6 de contourner le système de blocage de synthèse protéique dans les conditions de stress grâce à la phosphorylation de eIF2 $\alpha$  (Kazemi *et al.* [2004]). Très récemment, il a été montré que E6 des types HPV01, HPV08 et BPV1 interagit avec le coactivateur transcriptionnel MAML1 (Mastermind-like protein 1) via un motif LxxLL localisé dans le domaine C-terminal de MAML1, ce qui réprime la transactivation par MAML1 qui constitue le principal effecteur de la voie de signalisation NOTCH impliquée dans la différenciation et destin cellulaire. Cette interaction permet donc peut-être aux virus de réprimer la voie NOTCH annulant ainsi la différenciation des kératinocytes (Brimer *et al.* [2012]).

Finalement E6 peut interagir avec les facteurs de transcriptions FOXK1/K2 (forkhead transcription factors) (Gupta *et al.* [2003]) impliqués dans le cycle cellulaire.

**C- Adhésion et polarité cellulaire :** E6 de HPV à HRM peut lier et dégrader plusieurs protéines qui contrôlent la polarité cellulaire, l'adhésion intercellulaire et l'adhésion sur la matrice extra-cellulaire, s'agissant de hDLG (homolog of the Drosophila disc large tumor suppressor) (Gardioli *et al.* [1999]), hScrib (homolog of the Drosophila Scribble protein) (Nakagawa et Hui-bregtse [2000]), MuppI (Lee *et al.* [2000]) et PatJ (PALS1-associated TJ protein) (Storrs et Silverstein [2007]). Il a été aussi montré que E6 à HRM provoque la dégradation de PTPN3

### II.3 Les protéines précoces transformatrices : E5, E6 et E7

---

(Tyrosine- protein phosphatase non-receptor type 3) et TIP-2 (Tax interaction protein 2) associées à la signalisation du Facteur de croissance ce qui permet aux cellules infectées de survivre avec un taux minimum de Facteur de croissance (Favre-Bonvin *et al.* [2005]; Jing *et al.* [2007]). Parmi les cibles de E6, les protéines MAGI 1,2,3 (Glaunsinger *et al.* [2000], Thomas *et al.* [2001], Massimi *et al.* [2004]) sont associées aux jonctions serrées. MAGI 1 et 2 contribuent aussi à la régulation du fameux gène PTEN (phosphatase and tensin homolog) dont les mutations sont associées à plusieurs types de cancer (Cairns *et al.* [1997]). Finalement, E6 lie la protéine TIP-1 (Tax1 - binding protein 3), associée aux GTPases. Cette interaction provoque l'augmentation de la motilité des cellules infectées (Hampson *et al.* [2004]). Toutes ces protéines possèdent un domaine PDZ qui peut interagir avec E6 via un motif (x[T/S]x[L/V]) localisé à l'extrémité C-terminale des protéines E6 de HPV à HRM.

**D- Signalisation cellulaire :** Pour assurer la survie de son hôte et continuer à l'exploiter, le virus protège les cellules infectées du système immunitaire via l'interférence de E6 avec certains outils de l'immunité. Des expériences de double hybride ont montré que E6 s'associe avec IRF3 (Interferon regulatory factor 3) empêchant sa fonction de contrôler la progression vers l'apoptose une fois que la cellule infectée est identifiée par le système immunitaire (Ronco *et al.* [1998]). Une des fonctions importantes de E6 est l'activation de la voie de signalisation NF $\kappa$ B. Cette voie impliquée dans l'inhibition d'apoptose et l'adaptation en cas d'hypoxie est activée dans divers types de cancers épithéliales (Prasad *et al.* [2010]). Bien que le mécanisme d'activation ne soit pas encore clair, certains chercheurs évoquent le rôle des protéines à domaine PDZ (James *et al.* [2006]) ou celui de la deubiquitinase CYLD (Ubiquitin carboxyl-terminal hydrolase CYLD) (An *et al.* [2008]). De plus, E6 interagit aussi avec d'autres médiateurs de signalisation cellulaires, à savoir NHERF-1 (Accardi *et al.* [2011]), PKA (Kühne *et al.* [2000]), PKN (Gao *et al.* [2000]) SMAD3 (Mendoza *et al.* [2006]), TMEM87B (Transmembrane protein 87B) (Li *et al.* [2008]) et Tyk2 (Li *et al.* [1999]).

**E- Ubiquitination, sumoylation et dégradation :** Plusieurs protéines reconnues par E6 sont dégradées par la voie de polyubiquitination via le recrutement de la protéine E3 ubiquitine-ligase, E6AP (Huibregtse *et al.* [1993]). La dégradation de la protéine deubiquitinase CYLD (Ubiquitin carboxyl-terminal hydrolase CYLD) par E6 à HRM active la voie NF $\kappa$ B en cas d'hypoxie comme déjà mentionné (An *et al.* [2008]). en plus de E6AP et CYLD, E6 est capable de lier aussi HERC2 (HECT domain protein) (Vos *et al.* [2009]) et la deubiquitinase (USP15) (Vos *et al.* [2009]). La protéine E6 lie aussi la protéine PIASy (E3 SUMO ligase) qui stimule la sumoylation et l'activité transcriptionnelle de p53 et est responsable aussi d'une augmentation

## Chapitre II. Caractéristiques structurales et fonctionnelles des protéines des HPV

---

de la répression des des gènes dépendants de Rb en agissant sur les promoteurs de réponse de E2F. PIASy peut aussi stimuler l'apoptose et la senescence. L'interaction de PIASy avec E6 bloque ces activités (Bischof *et al.* [2006]). Finalement, E6 interagit avec UBC9 (Ubiquitin carrier protein 9) (Heaton *et al.* [2011]).

**F- Organisation et différenciation épithéliale :** A l'état normal, l'organisme possède des mécanismes aux niveaux tissulaires et cellulaires qui lui permettent de contrôler la prolifération et la différenciation cellulaires. Les cellules adhèrent à la matrice extra-cellulaire pour empêcher la dysplasie et ne déclenchent la prolifération qu'après la réception des signaux de prolifération. Elles l'arrêtent après l'achèvement du signal ou après l'atteinte d'une taille appropriée du tissu. E6 de HPV à HRM et BRM perturbe ce système de contrôle et d'organisation en interagissant avec plusieurs protéines impliquées. Parmi ces cibles se trouve la Paxilline, protéine associée à l'adhésion focale impliquée dans l'organisation des fibres d'actine et la maintenance de cytosquelette. La capture de la Paxilline par les E6 de certains Papillomavirus perturbe la formation des fibres d'actine et diminue l'intégrité épithéliale ce qui facilite la transmission virale (Tong et Howley [1997]; Tong *et al.* [1997]; Vande Pol *et al.* [1998]). E6 lie aussi la protéine Zyxine qui assure la connexion entre la cytosquelette et la matrice extra-cellulaire, la transmission du signal jusqu'au noyau, et coopère avec la Paxilline dans l'organisation des fibres d'actine (Deegenhardt et Silverstein [2001a]). De plus E6 déstabilise la matrice extra-cellulaire et empêche la différenciation terminale en interagissant respectivement avec la Fibuline 1(FBLN1) et ERC-55 (E6BP) (Calcium binding protein) (Chen *et al.* [1995]; Du *et al.* [2002]).

**G- Epigénétique :** Très récemment il a été rapporté aussi que E6 interagit avec trois méthyltrans- férases d'histones (CARM1, PRMT1, SET7). CARM1et PRMT1 catalysent la méthylation d'histones au niveau des promoteurs de réponse de p53 provoquant l'annulation de l'interaction entre p53 et la chromatine, indépendamment de la dégradation de p53 induite par E6. SET7 quant à elle, catalyse la méthylation du p53 à la lysine-372(p53K372 mono-methylation), ce qui protège p53 de la dégradation induite par E6 (Hsu *et al.* [2012]).

**H- Stabilité chromosomique** E6 lie la protéine MCM7 (minichromosome maintenance seven protein), nécessaire pour l'initiation de la réplication des génomes eucaryotes. E6 à HRM dégrade MCM7 et inhibe sa reconnaissance de la chromatine, ce qui cause l'accumulation des dommages et anomalies chromosomiques qui contribuent au processus de carcinogénèse (Kukimoto *et al.* [1998]). A l'état normal, les anomalies chromosomiques sont réparées par les systèmes XRCC1 et MGMT (O-6-methylguanine-DNA methyltransferase). Or, l'interaction de

### II.3 Les protéines précoces transformatrices : E5, E6 et E7

---

E6 avec ces deux protéines les déstabilise et empêche la réparation d'ADN (Iftner *et al.* [2002]; Srivenugopal et Ali-Osman [2002]). Ces activités de E6 permettent une réplication continue du génome viral, et pour contourner la limitation éventuelle causée par raccourcissement des chromosomes, E6 de HPV à HRM augmente la production de l'enzyme de Télomérase comme nous l'avons déjà mentionné.

**I- Trafic nucléaire et vésiculaire :** Il semble que la E6 de HPV à HRM perturbe le système du trafic cellulaire au niveau nucléaire et vésiculaire. Deux études consacrées à ce sujet ont été publiées en 2003 et 2007. La première étude montre une interaction entre E6 et les karyopherin (Kap)  $\alpha 2$ ,  $\beta 1$  et  $\beta 2$  responsables du transport des macromolécules du noyau vers le cytoplasme ou l'inverse (Le Roux et Moroianu [2003]). La deuxième étude montre que les E6 des types HPV16 et 18 lient et dégradent la protéine CAL porteuse du domaine PDZ (Golgi-associated PDZ), impliquée dans le trafic vésiculaire mais aussi dans la dégradation des protéines. Cette interaction se fait dans le cadre d'un triple complexe entre CAL, E6 et E6AP, qui constituent un ligand naturel de CAL. L'association E6-CAL est formée par la liaison du domaine PDZ de CAL et le motif de reconnaissance des PDZ au C-terminus de E6 (Jeong *et al.* [2007]). Ces deux activités de E6 expliquent peut-être la diversité des localisations des protéines reconnues par E6 (Facteurs de transcription nucléiques, protéines membranaires et cytoplasmiques).

**J- Acides nucléiques :** En plus des interactions protéiques décrites précédemment, notre laboratoire a montré que E6 peut lier *in vitro* des motifs structuraux de l'ADN. Cette reconnaissance exclusive des E6 de HPV à HRM se fait au niveau des jonctions de hollyday impliquées dans la recombinaison génétique (Nomine *et al.* [2006]; Ristriani *et al.* [2001]). E6 du type HPV16 peut aussi lier son propre transcrit pré-ARNm, ce qui laisse penser à une éventuelle régulation post-transcriptionnelle (Bodaghi *et al.* [2009]).

**Les motifs de reconnaissance de E6 :** Selon l'état actuel de la littérature, on peut classer les modes d'interaction de E6 en 4 types :

- **Interaction directe avec E6 via un motif LxxLL chargé négativement :** les deux domaines de E6 repliés forment une poche qui reconnaît un motif présentant le consensus LxxLL combiné avec des résidus acides trouvés chez de nombreuses protéines ciblées par E6 (E6AP, E6BP, IRF3, Paxilline, Tsc2. . .)(Baleja *et al.* [2006]) ;
- **Interaction avec E6 conditionnée par la formation préalable du complexe E6/E6AP :** Bak, c-myc, E6TP1, hMCM7 et P53 ;

## Chapitre II. Caractéristiques structurales et fonctionnelles des protéines des HPV

---

- **Domaines PDZ** : les domaines "PDZ" qui reconnaissent un motif consensus X[T/S]X[L/V] présent dans l'extrémité C-terminale non repliée de E6 à HRM (MAGI, DLG, Scrib, MuppI, PTPH1, TIP1, TIP2, PatJ...);
- **Mode d'interaction inconnu** : FADD, Gps2, hADA3, Procaspase 8... etc.

---

---

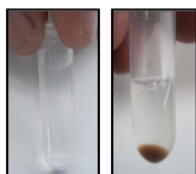
## Chapitre III

---

### Introduction au projet de recherche

Surproduite sous forme recombinante, la protéine E6 a toujours présenté des difficultés de purification. Malgré sa taille petite (150 résidus) et malgré 25 ans de recherches effectuées sur cette protéine dans plusieurs laboratoires au monde, E6 a très longtemps résisté aux expériences visant à la résolution de sa structure à cause de la nature insoluble et très agrégée des échantillons exprimés dans tous les systèmes d'expression testés ([Androphy \*et al.\* \[1987a\]](#); [Grossman \*et Laimins\* \[1989\]](#); [Imai \*et al.\* \[1989\]](#)) (Figure III.1).

Depuis 1995, notre équipe a beaucoup travaillé sur ce problème en utilisant plusieurs approches (optimisation de protocoles de production, mutagenèse dirigée, co-purification avec ligands . . . ) afin de résoudre sa structure et d'étudier ses interactions avec divers ligands. En 2001, l'équipe obtenait les premiers échantillons de E6 solubles et repliés. Deux ans plus tard notre équipe démontrait que E6 était composée de deux domaines à zinc (E6N et E6C). En 2006 l'équipe résolvait par RMN la structure du domaine C-terminal (E6C) de E6 du type HPV16 ([Nomine \*et al.\* \[2006\]](#)). Avant le début de cette thèse notre équipe avait déjà réussi la production d'un mutant de HPV16 E6 biologiquement actif et présentant une solubilité améliorée par rapport à la version sauvage. Dans ce mutant appelé E6 6C/6S, on avait remplacé tous les résidus cystéines non conservés (non impliqués dans la liaison avec le zinc) par des ré-



**Figure III.1 – Tendance à l'agrégation de E6 sauvage**

Aspect agrégatif de la version sauvage de la protéine HPV16 E6 (à droite) en comparaison avec celui du mutant HPV16 E6 F47R4C4S (à gauche), observée après ultracentrifugation à haute vitesse et longue durée de la protéine MBP-E6 purifiée.

### Chapitre III. Introduction au projet de recherche

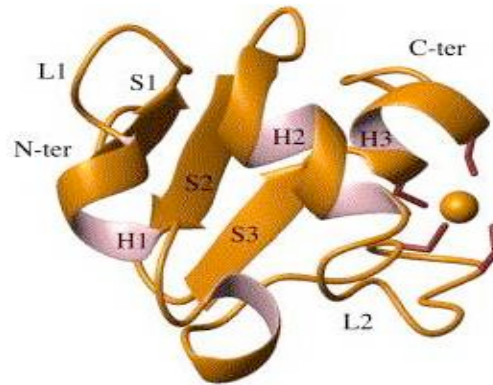
---

sidus sérines (C16S, C51S, C80S, C97S, C111S et C140S) dans le but d'empêcher l'agrégation provoquée par les ponts disulfure intermoléculaires formés durant la production et le stockage de la protéine. Ce mutant était bien purifiable sous forme soluble et native, mais lors de l'étape de concentration nécessaire pour les analyses structurales (RMN, Cristallographie... etc.), il ne dépassait pas un seuil de 20  $\mu$ M. Au delà de ce seuil, il initiait un processus d'agrégation lente aboutissant à une précipitation sous forme de fibres (Nomine *et al.* [2003, 2001]; Zanier *et al.* [2010]).

Après avoir muté le maximum des cystéines, notre équipe a alors effectué des études supplémentaires pour améliorer et raffiner ce mutant, ce qui a permis la "dé-mutation" des deux cystéines précédemment mutées au coeur du domaine E6N (C16 et C51). Une amélioration de la solubilité a été alors observée pour le mutant E6 4C/4S par rapport au mutant initial E6 6C/6S mais ce progrès restait insuffisant pour aborder l'étude structurale. Dans un stade plus avancé, il a été remarqué que la phénylalanine 47 participait à la formation d'un dimère du domaine E6N qui enclenchait un processus d'oligomérisation (Zanier *et al.* [2010]). Ceci a abouti à la conception du mutant E6 F47R 4C4S, dont la solubilité décuplée a enfin permis d'aborder l'étude structurale. Grâce à ces études, notre équipe a réussi à produire séparément les deux domaines de HPV16 E6 et à résoudre leurs structures en solution en utilisant les analyses RMN (Résonance Magnétique Nucléaire). La première analyse structurale de E6 a été la structure RMN, obtenue par notre équipe, du domaine C-terminal (E6C) (Nomine *et al.* [2006]) (Figure III.2).

Cependant la protéine E6 complète (full-length) restait largement insoluble. C'est dans ce cadre que j'ai commencé mes travaux de thèse en octobre 2008, avec pour but de continuer la quête de données structurales sur E6 tout en acquérant de nouvelles informations sur ses modes d'interaction avec ses cibles cellulaires. Très récemment j'ai eu la chance, avec mes collègues, de couronner les travaux de l'équipe par l'obtention de la structure cristallographique de E6 de la souche HPV16 (en complexe avec un motif LxxLL issu de l'ubiquitine-ligase E6AP) attendue depuis longtemps par la communauté scientifique. Il faut aussi noter qu'après un travail fait indépendamment de mon projet de thèse, notre équipe vient également de publier cette année une structure en solution du domaine E6N (mutant E6N F47R) et une structure corrigée du domaine E6C (mutant E6C 4C/4S) (Zanier *et al.* [2012]) (Figure III.3).

Cette thèse est composée de deux grandes parties : la première sera consacrée à l'étude structurale de deux protéines E6 issues de deux Papillomavirus, l'un humain à haut risque génital (HPV16 E6) et l'autre bovin (BPV1 E6). La deuxième partie traitera d'une analyse interactionnelle effectuée sur la base d'une large étude expérimentale testant l'interaction entre les protéines E6 de 6 variants de HPV plus le type bovin (BPV1 E6) avec une banque de 93 pep-

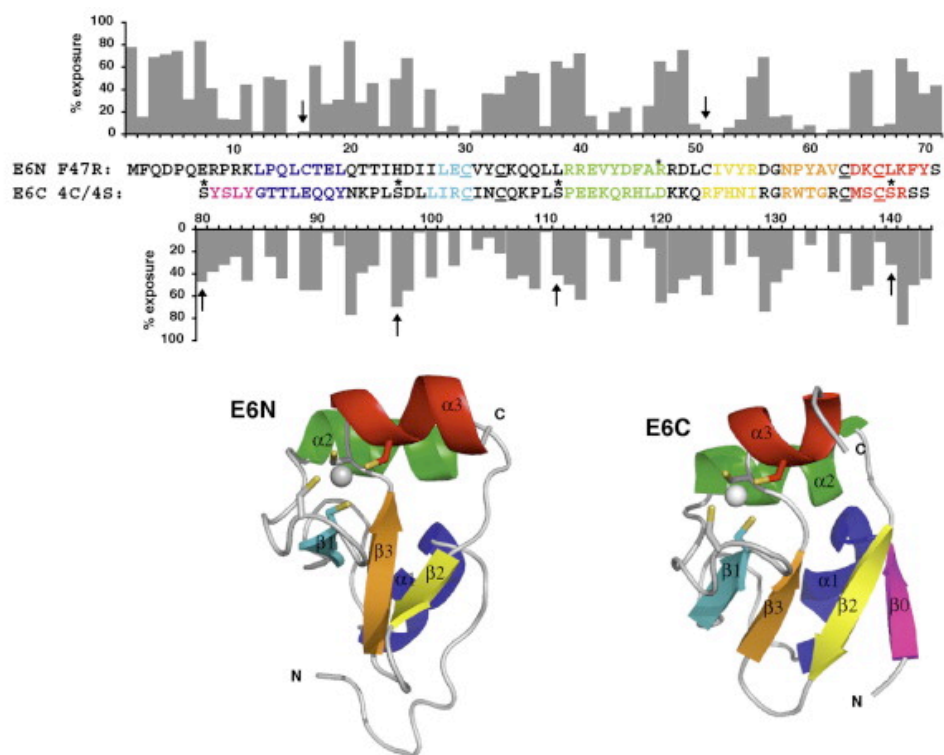


**Figure III.2 – Structure en solution du domaine E6C**

Structure 3D du domaine C-terminal (résidus : 87-141) de HPV16 E6, avec 3 brins  $\beta$  (S1, S2 et S3), 2 hélices  $\alpha$  (H1 et H2), une pseudo-hélice  $\alpha$  (H3), et une molécule du zinc attachée à la boucle L2 qui s'étend entre H2 et H3. Certains résidus de H3 participent à la liaison avec le zinc (d'après [Nomine \*et al.\* \[2006\]](#)).

tides dont 46 mutants de E6AP et 47 peptides extraits à partir de 30 protéines connues comme ligands de E6. Outre ces deux grands axes, j'évoquerai aussi dans cette thèse un projet annexe portant sur l'étude structurale de l'interaction entre le facteur de transcription papillomaviral E2 (HPV 16) et des peptides phosphorylés issus de la protéine cellulaire Brd4.





**Figure III.3 – Structures en solution des deux domaines de HPV16 E6**

En bas : Présentation des structures en solution des 2 domaines de E6 (E6N [à gauche] et E6C [à droite]). En haut : Pourcentage d'exposition au solvant des deux structures présenté sur un alignement des deux domaines. L'axe Y du E6C est inversé pour la clarté. Les résidus sont colorés selon les structures secondaires. Les résidus cystéines liant le zinc sont soulignés tandis les résidus cystéines non conservés sont indiqués par flèches. Les astérisques indiquent les mutations (F47R et 4C4S) (d'après Zanier *et al.* [2012]).

---

## RÉSULTATS ET DISCUSSION

## RESULTATS ET DISCUSSION

---

---

---

# Chapitre I

---

## Etude biochimique et structurale de la protéine E6

### 1 Stratégies d'expression et de purification de la protéine E6 issue des Papillomavirus bovins (BPV1 E6)

Article :

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

Nous allons présenter dans cet article un ensemble de procédures et techniques optimisées que nous avons utilisées pour produire la protéine E6 du type bovin 1 (BPV1 E6). Dans des études précédentes, il a été montré que BPV1 E6 interagit avec plusieurs peptides qui portent le motif consensus (LxxLL) (Bohl *et al.* [2000]; Vande Pol *et al.* [1998]). En utilisant des tests de Biacore (SPR), nous avons remarqué que BPV1 E6 interagit avec le peptide portant le motif (LxxLL) extrait du N-terminal de la protéine Paxilline, avec une constante de dissociation ( $K_d$ ) de 35,1 nM. Cette affinité est significativement plus élevée que les autres interactions E6/peptides déjà obtenues ( $K_d$  de l'ordre du micromolaire (Zanier *et al.* [2005a])). La protéine BPV1 E6 exprimée chez la bactérie *E. coli.* est peu soluble à l'état libre. Cependant, l'ajout du peptide LxxLL de la Paxilline pendant la biosynthèse de BPV1 E6 a amélioré considérablement la solubilité de cette protéine. La stratégie que nous avons utilisée pour produire le complexe BPV1 E6/Peptide LxxLL (Paxilline) a consisté à utiliser un vecteur d'expression protéique (plasmide de *E. coli* BL21 DE3) porteur d'un gène encodant une triple fusion contenant respectivement du N-terminal au C-terminal : la protéine de fusion solubilisante MBP (Maltose Binding Protein), le peptide LxxLL extrait de la protéine Paxilline (MDDL DALLADKE) et la protéine E6. En

## Chapitre I. Etude biochimique et structurale de la protéine E6

---

plus des ces trois éléments nous avons placé un site de reconnaissance de la protéase TEV selon les deux schémas suivants :

- **Complexe (MBP-tevs-LxxLL-E6) :** dans lequel le site de reconnaissance de TEV (tevs) est placé entre MBP et LxxLL permettant la création d'une fusion covalente entre la protéine E6 et le peptide LxxLL.

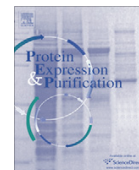
- **Complexe (MBP-LxxLL-tevs-E6) :** où le site de reconnaissance de TEV est placé entre LxxLL et E6 générant une fusion non covalente. Ces deux constructions ont permis la production du complexe BPV1 E6/LxxLL sous une forme soluble atteignant des concentrations très élevées.

Les analyses par RMN nous ont montré que la protéine E6 a conservé le même repliement dans les deux types de constructions. L'utilisation d'une stratégie similaire, à savoir l'expression de E6 en complexe avec son ligand peptidique préférentiel, a permis plus tard à notre équipe de résoudre la structure cristallographique de cette même protéine BPV1 E6 (voir annexe 1). Nous avons aussi résolu, de la même façon, la structure de E6 humain du type HPV 16, que nous traiterons dans la section suivante de ce chapitre. Cette approche originale que nous avons développée peut être utilisée dans des cas similaires où la protéine étudiée n'est stable et/ou soluble qu'en présence de son peptide préférentiel.



Contents lists available at ScienceDirect

## Protein Expression and Purification

journal homepage: [www.elsevier.com/locate/yprep](http://www.elsevier.com/locate/yprep)

## Strategies for bacterial expression of protein–peptide complexes: Application to solubilization of papillomavirus E6

Abdellahi Ould M'hamed Ould Sidi<sup>a</sup>, Khaled Ould Babah<sup>a</sup>, Nicole Brimer<sup>b</sup>, Yves Nominé<sup>a</sup>, Christophe Romier<sup>c</sup>, Bruno Kieffer<sup>c</sup>, Scott Vande Pol<sup>b</sup>, Gilles Travé<sup>a,\*</sup>, Katia Zanier<sup>a,\*</sup>

<sup>a</sup> Ecole Supérieure de Biotechnologie de Strasbourg, UMR 7242, Boulevard Sébastien Brant, BP 10413, 67412 Illkirch, France

<sup>b</sup> Department of Pathology, University of Virginia, Charlottesville, VA 22908, USA

<sup>c</sup> Institut de Génétique et de Biologie Moléculaire et Cellulaire, UMR 7104, INSERM U 964, Université de Strasbourg, 1 rue Laurent Fries, BP 163, 67404 Illkirch, France

### ARTICLE INFO

#### Article history:

Received 28 March 2011  
and in revised form 21 June 2011  
Available online 14 July 2011

#### Keywords:

BPV  
E6  
Paxillin  
LXXLL motif  
Protein aggregation  
Protein–peptide interactions

### ABSTRACT

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

© 2011 Elsevier Inc. All rights reserved.

### Introduction

Papillomaviruses (PVs) are small DNA viruses that infect epithelial tissues and produce hyperplasias within mammals, birds, and reptiles. Almost 200 PV types have been identified [1]. PVs display distinct tropisms for different body sites (skin, mouth, and genitalia) as well as distinct levels of pathogenicity. Although papillomaviruses initially produce benign hyperplasias, in a subset of virus types termed “high-risk”, these lesions can progress to cancer. In humans, cervical cancers are caused by infection with high-risk genital human papillomaviruses (HPVs)<sup>1</sup> [2] of which HPV16 is the most common “high-risk” genital type [3]. Bovine papillomavirus type 1 (BPV1) is one of the most extensively studied animal papillomavirus, serving as a model system for papillomavirus transcription, transformation, and replication [4].

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

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

\* Corresponding authors. Fax: +33 3 68854718.

E-mail addresses: [gilles.trave@unistra.fr](mailto:gilles.trave@unistra.fr) (G. Travé), [zanier@unistra.fr](mailto:zanier@unistra.fr) (K. Zanier).

<sup>1</sup> Abbreviations used: MBP, maltose binding protein; HPV, human papillomavirus; BPV, bovine papillomavirus; TEV, tobacco etch virus; SPR, surface plasmon resonance; RU, response units; NMR, nuclear magnetic resonance spectroscopy.

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

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

## Materials and methods

### MBP-E6 fusion constructs

Three constructs were generated where the DNA encoding for full-length BPV1 E6 (137, Genbank CAB 4650) was inserted in the pETM41 vector (kindly provided by G. Stier, EMBL, Heidelberg). In the first construct, BPV1 E6 was cloned at the C-terminus of 6His-tagged maltose binding protein (MBP) via a linker sensitive to TEV protease as previously described [26]. The second construct, referred to as MBP-tevs-LXXLL-E6, consists of the 10 amino acids of the N-terminal E6-binding LXXLL motif of paxillin (residues 1–10) fused to the N-terminus of E6 via a short 7 residue linker (detailed sequence: [6HisMBP]-[MSENYFQGMDDLDALLADKESGSA]-[E6]). In the third construct, named MBP-LXXLL-tevs-E6 we fused directly the LXXLL motif to the C-terminus of 6his-MBP and introduced the TEV sensitive linker at the junction between the LXXLL motif and the E6 sequence (detailed sequence: [6HisMBP]-[MDDLDALLADGSGSNENLYFQGS]-[E6]). All clones were verified by DNA sequencing.

### GST-peptide constructs

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

### Expression and purification procedures

BPV1 E6 fusion constructs were expressed overnight at 15 °C in *E. coli* BL21 (DE3) cells grown in either LB or M9 medium supplemented with <sup>15</sup>N labeled ammonium sulfate salts. All purification buffers were extensively degassed and equilibrated with argon. Expression cultures were harvested by centrifugation and resuspended in buffer A (Tris 50 mM pH 6.8, NaCl 400 mM, and DTT 2 mM) supplemented with 5% glycerol, 2.5 µg/ml DNase I, 2.5 µg/ml RNase I, 1 µg/ml lysozyme, and anti-protease cocktail (Roche Applied Science) at the concentration recommended by the manufacturer. Cells were broken by sonication on ice and then centrifuged at 38,000 rpm for 1 h at 4 °C. The supernatant was loaded on an affinity amylose resin column (New England Biolabs)

equilibrated in buffer A. After extensive washing, the protein fusion was eluted using buffer A containing 10 mM maltose. Hence, the affinity purified sample was ultracentrifuged overnight at 110,000g and 4 °C in a SW41 rotor (Beckman). Subsequently, the MBP moiety was removed by TEV protease digestion carried out at 8 °C for 12 up to 36 h. In the case of the 6His-MBP-LXXLL-tevs-E6 TEV protease digestion was performed in the presence of a fourfold stoichiometric excess of synthetic peptide derived from paxillin (see “Results” section for further details). The digested product was concentrated and loaded on a HiLoad 16/60 Superdex 75 gel filtration column (Amersham Biosciences) previously equilibrated in buffer B (20 mM sodium phosphate buffer (pH 6.8), NaCl 200 mM, and DTT 2 mM). Protein concentration was estimated by absorption at 280 nm.

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

### Surface plasmon resonance experiments

Surface plasmon resonance (SPR) experiments were performed at 25 °C on a Biacore 2000 instrument (Biacore) equilibrated in buffer B. The experimental set-up followed was described in a previous publication [28]. Briefly, peptides containing the LXXLL motives of paxillin and E6AP and a negative control peptide were expressed fused to GST (see above) and captured onto the Biacore's surface via a polyclonal anti-GST antibody. The analyte protein was monomeric BPV1 E6 fused to MBP (MBP-E6) isolated by gel-filtration 24 h prior to injection and concentrated to a stock concentration below the aggregation threshold (i.e. 20 µM). We previously demonstrated that the MBP tag does not significantly alter binding properties of E6 proteins to their target peptide while permitting the preservation of the E6 as a soluble monomer and an increase of the SPR signal [29]. The analyte MBP-E6 protein was therefore injected and the interaction with the captured peptides monitored. Immediately after synthetic peptide containing the LXXLL motif from either E6AP or paxillin (depending on whether GST-E6AP or GST-paxillin was immobilized on the surface) was injected at a concentration of 20 µM during the analyte's dissociation phase, thus avoiding a regeneration step. The negative control peptide sensorgram was used to correct for non-specific binding whereas buffer effects were corrected by subtracting buffer injection sensorgrams. All data were interpreted using the BiaEvaluation 3.2 software (Biacore). Kinetic data were interpreted by global fit using a simple Langmuir interaction model. The kinetic parameters reported were obtained from fits with  $\chi^2$  values smaller than 4% of the globally fitted R<sub>max</sub> values.

### Synthetic peptide

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

### NMR spectroscopy

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

### Mass spectrometry

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

### Results

#### Domain architecture and stability of BPV1 E6

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

We therefore applied to BPV1 E6 the same expression and purification protocols developed for the other HPV E6 proteins [26]. The wild-type sequence of full-length BPV1 E6 was fused to the C-terminus of maltose binding protein (MBP) via a linker sensitive to TEV protease, over-expressed and affinity purified on amylose resin. Overnight ultracentrifugation to clear soluble MBP-E6

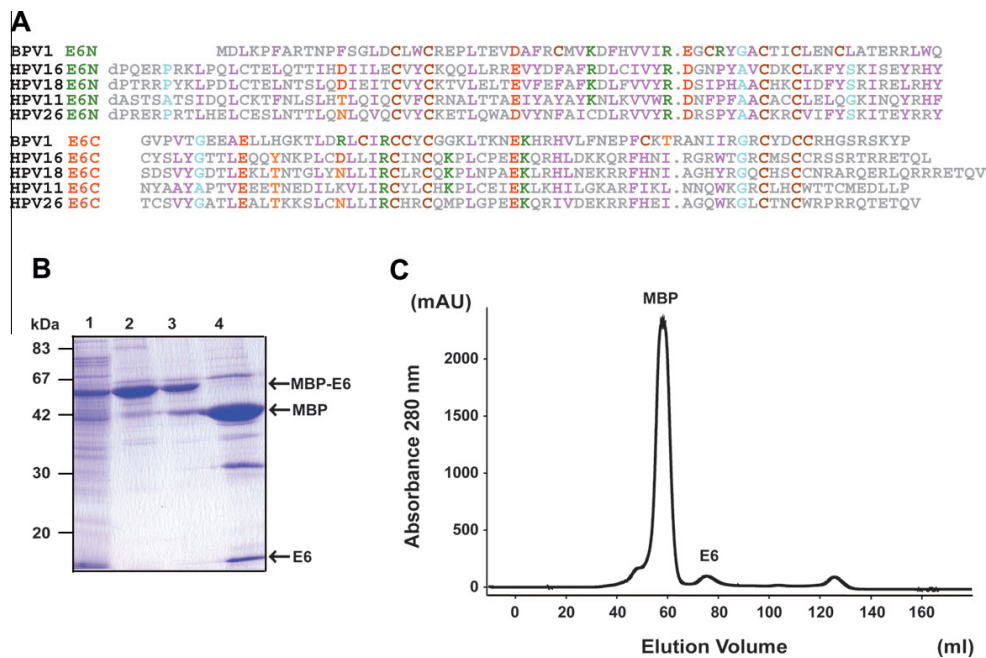
aggregates led to a significant loss of material in the supernatant (Fig. 1B, lanes 2 and 3). The cleared sample was then subject to TEV protease cleavage. The identity of the BPV1 E6 band generated after TEV digestion and migrating at 16 kDa (Fig. 1B, lane 4) was confirmed by MALDI-TOF mass spectrometry analysis (Table S1, Supplementary data). The digested sample was subsequently loaded on a gel-filtration column (Fig. 1C). We could only recover low amounts of protein at the elution volume expected for monomeric BPV1 E6 separated from the MBP tag. In addition gel filtrated E6 could not even stand further concentration due to precipitation onto the concentrator's membrane.

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

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

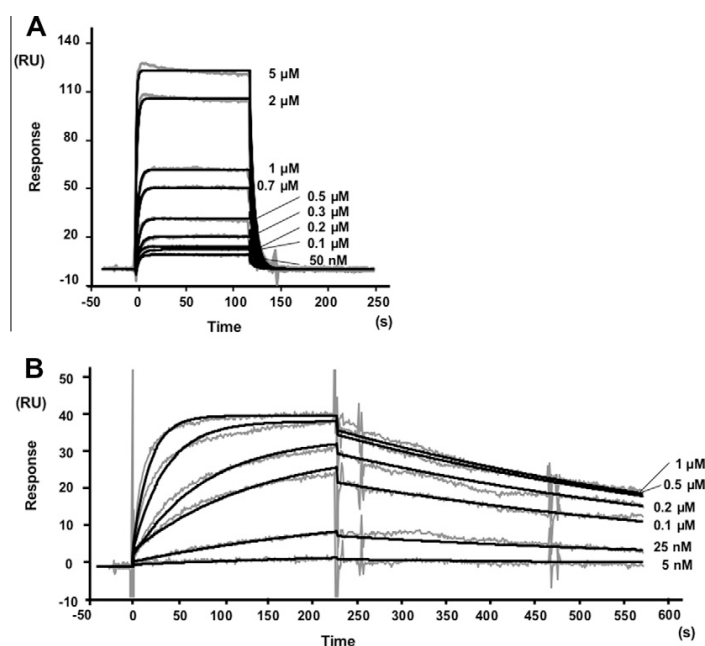
The BPV1 E6 protein has been shown to bind to several peptides bearing the LXXLL consensus motif [17,18]. We investigated the binding of BPV1 E6 to two distinct peptides containing respectively the LXXLL motif of E6AP and the N-terminal LXXLL motif of paxillin by SPR (Biacore). Freshly purified monomeric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

The sensorgrams shown in Fig. 2 denote a remarkably slow dissociation phase for the E6/paxillin interaction as compared to E6/E6AP. Kinetic analysis using a simple bimolecular Langmuir



**Fig. 1.** (A) Alignment of E6 sequences (upper half E6-N and lower half E6-C) from the BPV1 and several HPV species. Positions exhibiting conserved physicochemical characteristics are colored as follows: magenta, hydrophobic (W, F, Y, L, I, V, and M); green, basic (K, R, and H); red, acidic (E and D); orange, polar (Q, N, T, and S); brown, cysteine (C); cyan, small (G, A, and P). (B) 15% SDS-PAGE gel showing the different purifications steps of BPV1 E6. Affinity-purified MBP-E6 was cleared from soluble aggregates by ultracentrifugation and digested using TEV protease. Subsequently, the sample was concentrated and injected on a HighLoad Superdex75 16/60 column. Lane 1: total cell lysate; lane 2: eluate from amylose resin before overnight ultracentrifugation; lane 3: eluate from amylose resin after overnight ultracentrifugation; lane 4: TEV protease digest. (C) Gel filtration after separation from the MBP by TEV digestion. Peaks corresponding to the MBP and E6 proteins are indicated. Note that very small amounts of monomeric E6 are recovered after gel filtration. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)





**Fig. 2.** Biacore datasets used for kinetic analysis of the interactions of MBP-BPV1 E6 with peptides containing (A) the LXXLL motif of E6AP and (B) the N-terminal LXXLL motif of paxillin. Gray lines represent SPR sensorgrams obtained by injecting analyte MBP-E6 on surfaces capturing 60 (A) and 30 (B) RUs of GST-peptide ligand. The concentration of analyte is indicated on the right hand side of each sensorgram. Black lines are the global fits of the sensorgrams to a simple bimolecular Langmuir model. The ratio of the globally fit  $R_{max}$  ( $R_{max}^{global}$ ) over the theoretical  $R_{max}$  ( $R_{max}^{theoretical}$ —estimated from the amount of fusion peptide captured—)  $R_{max}^{global}/R_{max}^{theoretical}$  was 0.9 and 0.6 for the GST-paxillin and GST-E6AP ligands. Kinetic parameters are reported in Table 1.

interaction model yielded  $k_{off}$  rate values of 0.2 and  $1.7 \times 10^{-3} \text{ M}^{-1} \text{ s}^{-1}$  for the E6AP and paxillin peptides, respectively. The equilibrium ( $K_D$ ), values estimated from the  $k_{off}/k_{on}$  ratios are 1.3  $\mu\text{M}$  and 35.1 nM (Table 1). Therefore, whereas the affinity of the BPV1 E6/E6AP interaction approximates that of HPV16 E6/E6AP (4  $\mu\text{M}$ ) [29], the affinity of BPV1 E6 for the paxillin peptide is significantly stronger.

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

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

Initially we have tested a number of co-expression and co-purification protocols. Co-purification of GST-paxillin peptide and MBP-BPV1 E6 fusion expressed in separate bacterial cultures was attempted on either amylose (MBP-binding) or glutathione (GST-binding) resins. This approach yielded very little amounts of complex, most probably because MBP-E6 aggregates pre-formed upon bacterial expression were not proficient for binding to the peptide (data not shown). A related approach, consisting of adding an excess of synthetic paxillin peptide to MBP-E6 expressions during cell

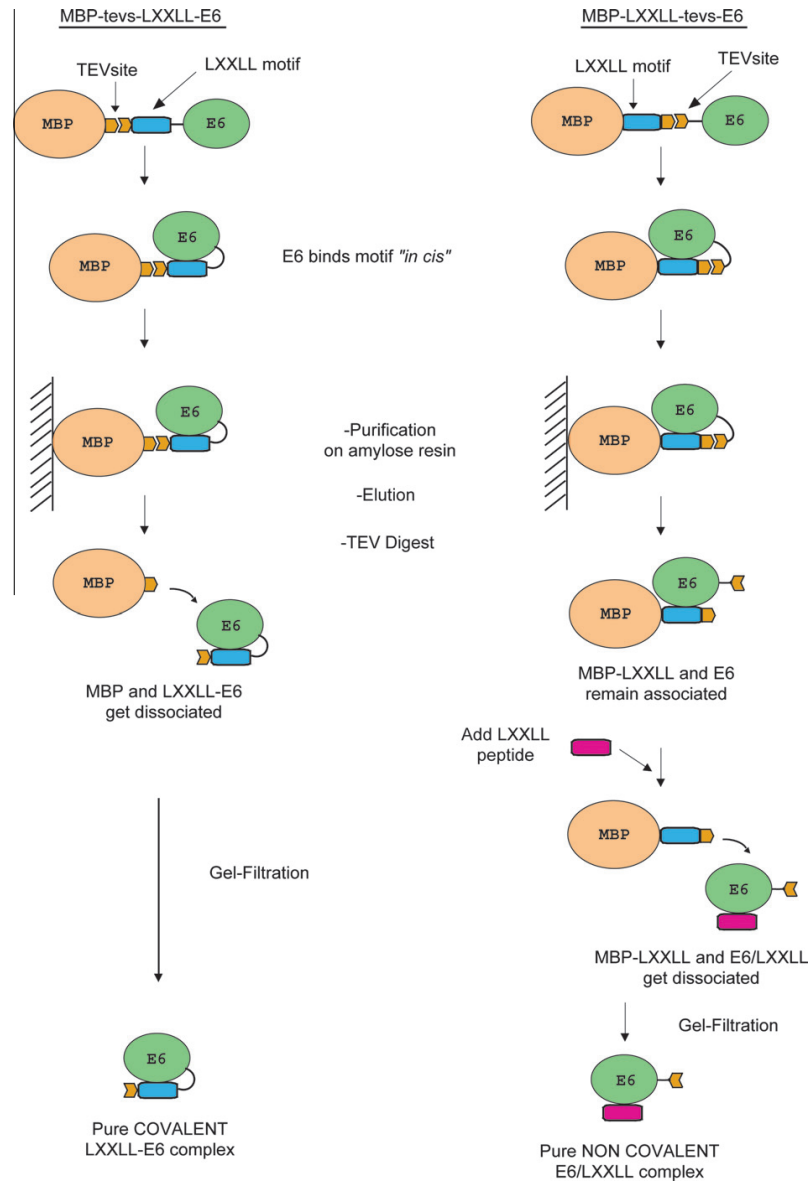
lysis was also unsuccessful (data not shown). Finally, co-expression within the same bacterial cell of GST-paxillin peptide and untagged E6, both constructs being expressed by plasmids bearing different replication origins [30], led to co-purification of BPV1 E6 and GST-paxillin peptide, albeit at levels too modest for biophysical and structural studies (data not shown).

We therefore tried to stabilize BPV1 E6 by expressing the paxillin LXXLL motif and the E6 protein “*in cis*” on the same fusion construct. Two alternative constructs were generated. The first fusion protein (MBP-tevs-LXXLL-E6) was constructed as follows: [6HisMBP]-[TEV site]-[paxillin peptide]-[E6] (Fig. 3, left panel). The MBP-tevs-LXXLL-E6 fusion protein was expressed and purified exactly as described above for MBP-E6. The affinity purification on amylose resin raised amounts of soluble MBP-tevs-LXXLL-E6 fusion equivalent to the amounts of soluble MBP-E6 previously obtained. However, the ultracentrifugation step indicated that the MBP-tevs-LXXLL-E6 preparations contained very little amounts of soluble aggregates as compared to MBP-E6 preparations (Fig. 4A, lanes 2 and 3). TEV digestion of the MBP-tevs-LXXLL-E6 fusion was unusually slow (up to 36 h), suggesting that the conformation of the LXXLL-E6 construct hindered accessibility to the proteolytic site. The identity of the LXXLL-E6 band (Fig. 4A, lane 3) was confirmed by MALDI-TOF analysis (Table S2, Supplementary data). By contrast to what was previously observed for the MBP-E6 construct, the LXXLL-E6 moiety released by TEV proteolysis did not undergo precipitation neither during the proteolysis reaction nor during the subsequent concentration step. Gel filtration raised large amounts of monomeric LXXLL-E6 (Fig. 4B). In addition LXXLL-E6 samples could be concentrated up to 300  $\mu\text{M}$  without undergoing any detectable aggregation.

To extend this approach, a second fusion protein (MBP-LXXLL-tevs-E6) was constructed as follows: [6HisMBP]-[paxillin peptide]-[TEV site]-[E6] (Fig. 3 right panel). This fusion protein was highly

**Table 1**  
Kinetic parameters derived from Biacore sensorgrams shown in Fig. 2.

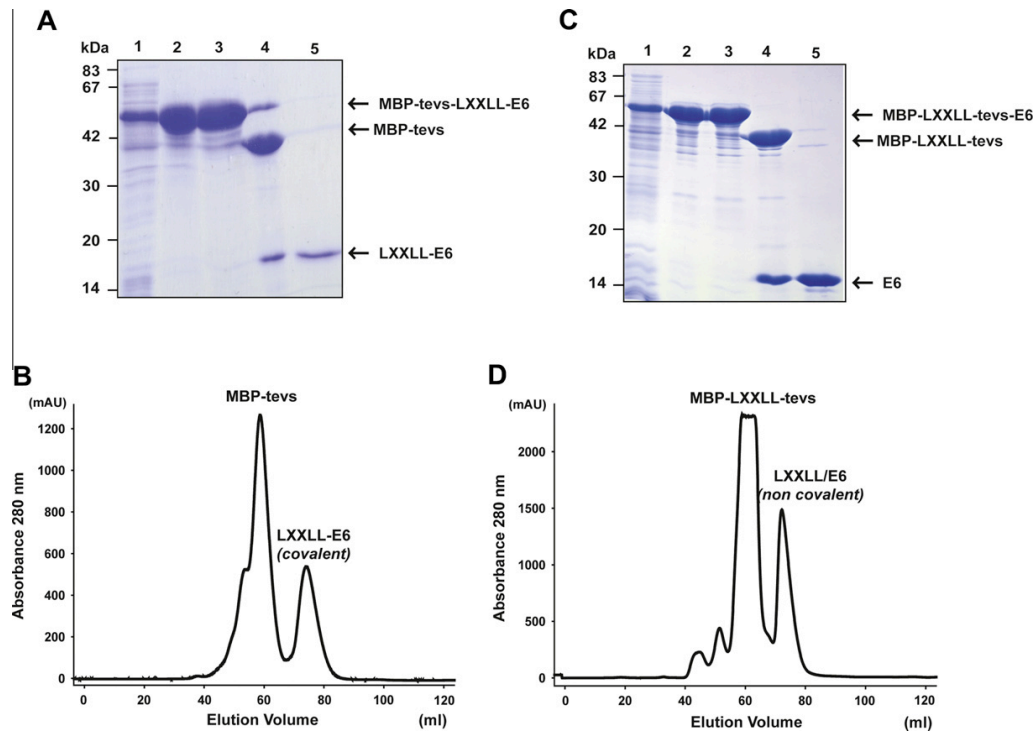
	$k_{on}$ ( $\text{M}^{-1} \text{s}^{-1}$ )	$k_{off}$ ( $\text{s}^{-1}$ )	$k_{off}/k_{on}$ (M)
BPV1 E6/E6AP	$1.5 \times 10^5$	0.2	$1.3 \times 10^{-6}$
BPV1 E6/paxillin	$5.0 \times 10^4$	$1.7 \times 10^{-3}$	$3.5 \times 10^{-8}$



**Fig. 3.** Purification scheme for the MBP-tevs-LXXLL-E6 construct (left) and the MBP-LXXLL-tevs-E6 construct (right). Both constructs stabilize the BPV1 E6 protein during expression by allowing it to form a complex *in cis* with its cognate LXXLL motif peptide. TEV digestion is carried out for both constructs after the first affinity purification step. In the case of the MBP-tevs-LXXLL-E6 construct, this leads to the isolation of pure covalent LXXLL-E6 complex after gel filtration. On the other hand, TEV proteolysis of the MBP-LXXLL-tevs-E6 construct is performed in the presence of a fourfold excess of LXXLL synthetic peptide, which binds to the E6 moiety, allowing the displacement of the MBP-LXXLL fusion and leading to isolation of the non-covalent E6/LXXLL complex after gel filtration.

expressed as a soluble form. Like for the MBP-tevs-LXXLL-E6 construct, affinity purified MBP-LXXLL-tevs-E6 contained only minor amounts of aggregated material when submitted to overnight ultracentrifugation (Fig. 4C, lanes 2 and 3). The MBP-LXXLL-tevs-E6 preparation was further processed as the MBP-tevs-LXXLL-E6 construct, with the following modification: TEV digestion was performed in the presence of a fourfold stoichiometric excess of synthetic LXXLL paxillin peptide, to allow the released E6 moieties to be displaced from the MBP-fused LXXLL peptides to the small synthetic peptides. The E6/peptide complex was then separated from MBP-paxillin peptide and TEV by gel-filtration, eluting at

the volume expected for a monomeric complex (Fig. 4D). The gel-filtrated sample could be concentrated up to 300  $\mu$ M just like the covalent LXXLL-E6 fusion described before. To further verify the integrity of the E6 and LXXLL peptide in the non-covalent complex, we performed electrospray mass spectrometry analysis in both denaturing and native conditions. In denaturing conditions we identified two species of 1724.8 and 16051.4 Da, corresponding, respectively, to the LXXLL peptide (predicted molecular weight 1725.8 Da) and BPV1 E6 (predicted molecular weight 16051.6 Da). When the analysis was performed in native conditions we obtained a single species of 17904.3 Da, which corresponds to the



**Fig. 4.** Purification of the construct of MBP-tevs-LXXLL-E6 (A and B) and MBP-LXXLL-tevs-E6 (C and D) constructs. (A and C) 15% SDS-PAGE gels for the two purifications. Lane 1: total cell lysate; lane 2: eluate from amylose resin before overnight ultracentrifugation; lane 3: eluate from amylose resin before overnight ultracentrifugation; lane 4: TEV protease digest; lane 5: pooled gel-filtration fractions at the elution volume expected for the E6/paxillin complex. (B and D) Gel filtration profiles of LXXLL/E6 and E6/LXXLL samples after separation from the MBP. Peaks corresponding to MBP-tevs or MBP-LXXLL-tevs and to E6/paxillin complexes are indicated. See also legend of Fig. 1.

sum of the molecular weights of the LXXLL peptide, BPV1 E6 and the two zinc atoms (predicted molecular weight of the complex: 17911.2 Da) (Table 2).

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

#### NMR analysis of BPV1 E6/paxillin complexes

The expression of both the MBP-tevs-LXXLL-E6 and MBP-LXXLL-tevs-E6 fusion constructs were reproduced in appropriate minimal M9 media, giving rise to  $^{15}\text{N}$ -labeled samples that were subsequently purified according to the protocols described above and concentrated up to 300  $\mu\text{M}$ . The  $^1\text{H}$ ,  $^{15}\text{N}$ -HSQC spectrum of the covalent LXXLL-E6 fusion complex (Fig. 5A) displays the chemical shift dispersion typical for a folded protein. Approximately 147 amide peaks can be counted in the HSQC spectrum. This number is

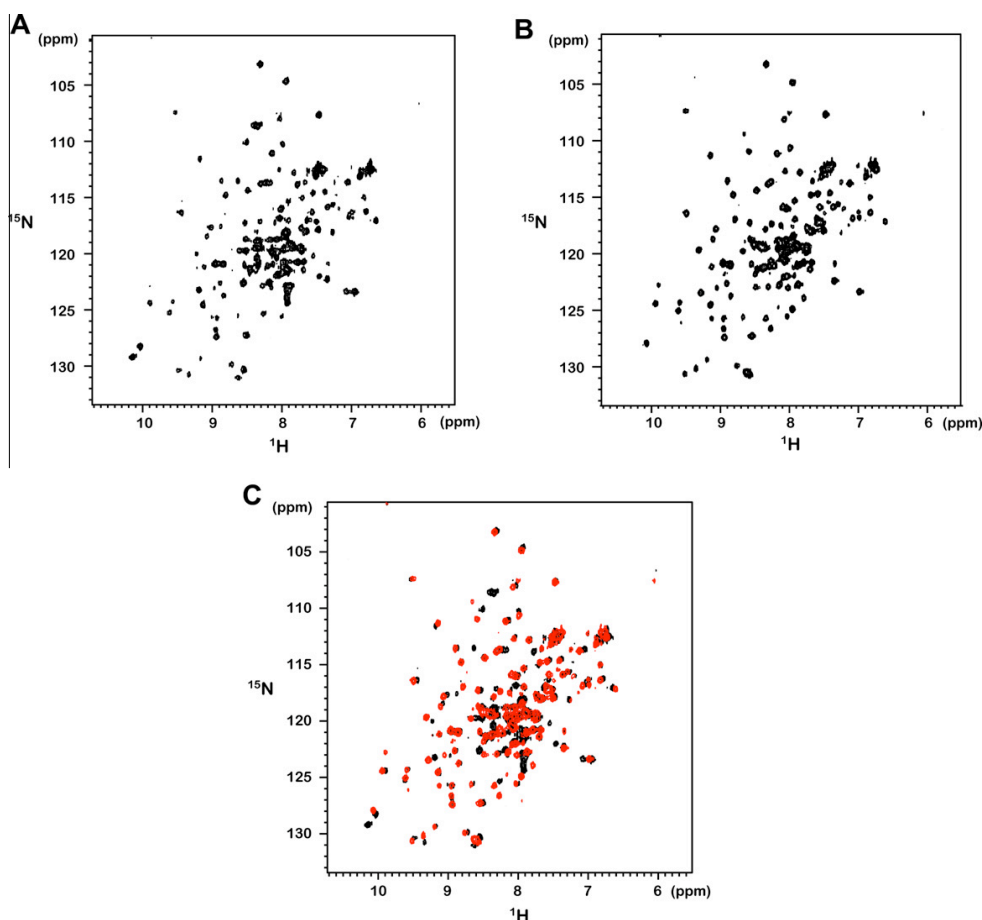
close to that predicted from the sequence of the LXXLL-E6 construct (consisting of 155 amino acids of which six are proline residues). The  $^1\text{H}$ ,  $^{15}\text{N}$ -HSQC spectrum of the non-covalent complex of BPV1 E6 with non-labeled synthetic LXXLL peptide (E6/LXXLL) is also typical for a folded protein (Fig. 5B). The number of amide peaks counted in the spectrum is 130, which is close to the number of amide resonances in the  $^{15}\text{N}$  labeled BPV1 E6 counterpart (140 residues of which six are prolines) of the complex. Interestingly, signal line-width in the latter spectrum appears more homogeneous when compared to that of the first spectrum. This suggests that the covalent peptide fusion to the N-terminus of BPV1 E6 might introduce some constraints to complex dynamics and stability. Superimposition of the two spectra (Fig. 5C) indicates that the majority of amide signals in the better dispersed regions of the spectrum have very similar chemical shift values, suggesting that the conformations of E6 in the covalent and non covalent complexes are very similar.

#### Discussion

Papillomavirus E6 proteins are notoriously difficult to express and purify. Unfused E6 proteins form insoluble aggregates upon bacterial overexpression. Usual purification tags such as the 6His-tag or the Glutathione-S-Transferase (GST) generate no or little solubilization of E6. By contrast, fusion to the maltose binding protein (MBP) is particularly efficient in solubilizing E6 proteins. However, a large fraction of the soluble MBP-E6 proteins may exist as soluble aggregates. We have previously developed various assays to detect these soluble aggregates and used ultracentrifugation approaches to eliminate them [23,26,31]. However,

**Table 2**  
Molecular weights of the non-covalent BPV1 E6/paxillin complex and its components as derived from both denaturing and native mass spectrometry measurements.

	BPV1 E6 (-Zn <sup>2+</sup> ) (Da)	Paxillin (Da)	BPV1 E6 (+Zn <sup>2+</sup> )/ LXXLL (Da)
Predicted	16051.6	1725.8	17903.2
Denaturing mass-spec.	16051.4	1724.8	-
Native mass-spec	-	-	17904.3



**Fig. 5.**  $^1\text{H}$ - $^{15}\text{N}$  HSQC spectra of (A)  $^{15}\text{N}$  labeled LXXLL-E6 fusion (covalent) and of (B)  $^{15}\text{N}$  labeled E6 bound to unlabeled LXXLL peptide (non-covalent). Spectra were acquired with 32 scans at 296 K. (C) Overlay of spectra shown in A (black) and in B (red). (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

monomeric E6 proteins obtained by these procedures generally remain prone to self-association at high concentration, hindering their biophysical and structural analysis [27].

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

To turn around this problem, we have searched for means of improving the solubility of BPV1 E6. Previous studies had shown that a construct of BPV1 E6 fused to the C-terminus of the same paxillin-derived LXXLL peptide used in the present study was inactivated *in vivo* for biological functions requiring interactions with partners possessing a LXXLL E6-binding motif. This suggested that the fused peptide was efficiently blocking the LXXLL-binding site of BPV1 E6 [18]. SPR measurements showed that BPV1 E6 binds to the paxillin LXXLL peptide with high affinity and slow dissociation rate kinetics. Given its stability, we reasoned that the protein/peptide complex may persist through the purification steps and ensure

the stabilization of the E6 protein at high concentrations. Indeed the conserved leucine residues of the LXXLL motif are likely to interact with a hydrophobic patch exposed on the surface of E6, which may promote aggregation of the unbound protein. We tested several alternative strategies. Approaches based on co-expression and co-purification of BPV1 E6 and the cognate peptide were unsuccessful or yielded insufficient amounts of complexes. By contrast, approaches based on “*in cis*” expression of MBP, peptide and E6, in the form of triple fusion, turned out to be remarkably successful. Two alternative strategies were used to express and purify samples of BPV1 E6/paxillin complexes. One strategy, based on the MBP-tevs-LXXLL-E6 construct, led to a covalent (intra-molecular) LXXLL-E6 fusion construct, which proved soluble at high concentrations and amenable to biophysical analysis by NMR. The second strategy, based on the MBP-LXXLL-tevs-E6 construct, allowed us to obtain samples of a non-covalent (inter-molecular) E6/LXXLL complex, which was equally soluble at high concentrations and also amenable to NMR studies. This latter approach has the advantage of allowing for subunit specific labeling, thus reducing the complexity of NMR spectra. The success of the triple fusion approach indicates that both peptidic ligand and MBP tag are required during biosynthesis for the solubilization of BPV1 E6.

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

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

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

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

## Acknowledgments

We thank F. Ruffenach (CEBGS, IGBMC) for performing MS analysis and P. Eberling (IGBMC) for peptide synthesis. 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 National Institute of Health (NIH). A.O.M.O.S. was supported by a grant of the ARC.

## Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at doi:10.1016/j.pep.2011.06.013.

## References

- [1] E.M. de Villiers, C. Fauquet, T.R. Broker, H.U. Bernard, H. zur Hausen, Classification of papillomaviruses, *Virology* 324 (2004) 17–27.
- [2] H. zur Hausen, Papillomaviruses in human cancers, *Proc. Assoc. Am. Phys.* 111 (1999) 581–587.
- [3] F.X. Bosch, M.M. Manos, N. Munoz, M. Sherman, A.M. Jansen, J. Peto, M.H. Schiffman, V. Moreno, R. Kurman, K.V. Shah, Prevalence of human papillomavirus in cervical cancer: a worldwide perspective. International biological study on cervical cancer (IBSCC) Study Group, *J. Natl. Cancer Inst.* 87 (1995) 796–802.
- [4] L. Nasir, M.S. Campo, Bovine papillomaviruses: their role in the aetiology of cutaneous tumours of bovinds and equids, *Vet. Dermatol.* 19 (2008) 243–254.
- [5] M. Scheffner, J.M. Huibregtse, R.D. Vierstra, P.M. Howley, The HPV-16 E6 and E6-AP complex functions as a ubiquitin-protein ligase in the ubiquitination of p53, *Cell* 75 (1993) 495–505.
- [6] B.A. Werness, A.J. Levine, P.M. Howley, Association of human papillomavirus types 16 and 18 E6 proteins with p53, *Science* 248 (1990) 76–79.
- [7] O. Chakrabarti, S. Krishna, Molecular interactions of 'high risk' human papillomaviruses E6 and E7 oncoproteins: implications for tumour progression, *J. Biosci.* 28 (2003) 337–348.
- [8] M. Thomas, J. Dasgupta, Y. Zhang, X. Chen, L. Banks, Analysis of specificity determinants in the interactions of different HPV E6 proteins with their PDZ domain-containing substrates, *Virology* 376 (2008) 371–378.
- [9] J.M. Huibregtse, M. Scheffner, P.M. Howley, Localization of the E6-AP regions that direct human papillomavirus E6 binding, association with p53, and ubiquitination of associated proteins, *Mol. Cell. Biol.* 13 (1993) 4918–4927.
- [10] J.J. Chen, Y. Hong, E. Rustamzadeh, J.D. Baleja, E.J. Androphy, Identification of an alpha helical motif sufficient for association with papillomavirus E6, *J. Biol. Chem.* 273 (1998) 13537–13544.
- [11] L.V. Ronco, A.Y. Karpova, M. Vidal, P.M. Howley, Human papillomavirus 16 E6 oncoprotein binds to interferon regulatory factor-3 and inhibits its transcriptional activity, *Genes Dev.* 12 (1998) 2061–2072.
- [12] D. Patel, S.M. Huang, L.A. Baglia, D.J. McCance, The E6 protein of human papillomavirus type 16 binds to and inhibits co-activation by CBP and p300, *EMBO J.* 18 (1999) 5061–5072.
- [13] H. Zimmermann, R. Degenkolbe, H.U. Bernard, M.J. O'Connor, The human papillomavirus type 16 E6 oncoprotein can down-regulate p53 activity by targeting the transcriptional coactivator CBP/p300, *J. Virol.* 73 (1999) 6209–6219.
- [14] R.C. Elston, S. Naphine, J. Doorbar, The identification of a conserved binding motif within human papillomavirus type 16 E6 binding peptides, E6AP and E6BP, *J. Gen. Virol.* 79 (1998) 371–374.
- [15] K. Das, J. Bohl, S.B. Vande Pol, 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.* 74 (2000) 812–816.
- [16] X. Tong, W. Boll, T. Kirchhausen, P.M. Howley, Interaction of the bovine papillomavirus E6 protein with the clathrin adaptor complex AP-1, *J. Virol.* 72 (1998) 476–482.
- [17] S.B. Vande Pol, M.C. Brown, C.E. Turner, Association of Bovine Papillomavirus Type 1 E6 oncoprotein with the focal adhesion protein paxillin through a conserved protein interaction motif, *Oncogene* 16 (1998) 43–52.
- [18] J. Bohl, K. Das, B. Dasgupta, S.B. Vande Pol, Competitive binding to a charged leucine motif represses transformation by a papillomavirus E6 oncoprotein, *Virology* 271 (2000) 163–170.
- [19] F. Lipari, G.A. McGibbon, E. Wardrop, M.G. Cordingley, Purification and biophysical characterization of a minimal functional domain and of an N-terminal  $Zn^{2+}$ -binding fragment from the human papillomavirus type 16 E6 protein, *Biochemistry* 40 (2001) 1196–1204.
- [20] Y. Nominé, S. Charbonnier, T. Ristriani, G. Stier, M. Masson, N. Cavusoglu, A. Van Dorselaer, E. Weiss, B. Kieffer, G. Travé, Domain substructure of HPV E6 protein: biophysical characterization of E6 C-terminal DNA-binding domain, *Biochemistry* 42 (2003) 4909–4917.
- [21] Y. Nominé, M. Masson, S. Charbonnier, K. Zanier, T. Ristriani, F. Deryckere, A.P. Sibler, D. Desplancq, R.A. Atkinson, E. Weiss, G. Orfanoudakis, B. Kieffer, G. Trave, Structural and functional analysis of E6 oncoprotein: insights in the molecular pathways of human papillomavirus-mediated pathogenesis, *Mol. Cell* 21 (2006) 665–678.
- [22] P.R. Daniels, C.M. Sanders, P. Coulson, N.J. Maitland, Molecular analysis of the interaction between HPV type 16 E6 and human E6-associated protein, *FEBS Lett.* 416 (1997) 6–10.
- [23] Y. Nominé, T. Ristriani, C. Laurent, J.F. Lefèvre, E. Weiss, G. Travé, A strategy for optimizing the monodispersity of fusion proteins: application to purification of recombinant HPV E6 oncoprotein, *Protein Eng.* 14 (2001) 297–305.
- [24] R. Degenkolbe, P. Gilligan, S. Gupta, H.U. Bernard, Chelating agents stabilize the monomeric state of the zinc binding huma, papillomavirus 16 E6 oncoprotein, *Biochemistry* 42 (2003) 3868–3873.
- [25] Y. Liu, J.J. Cherry, J.V. Dineen, E.J. Androphy, J.D. Baleja, Determinants of stability for the E6 protein of papillomavirus type 16, *J. Mol. Biol.* 386 (2009) 1123–1137.
- [26] K. Zanier, Y. Nominé, S. Charbonnier, C. Ruhlmann, P. Schultz, J. Schweizer, G. Trave, 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 (2007) 59–70.
- [27] K. Zanier, C. Ruhlmann, F. Melin, M. Masson, A. Ould M'hamed Ould Sidi, X. Bernard, B. Fischer, L. Brino, T. Ristriani, V. Rybin, M. Baltzinger, S. Vande Pol, P. Hellwig, P. Schultz, G. Trave, E6 proteins from diverse papillomaviruses self-associate both in vitro and in vivo, *J. Mol. Biol.* 396 (2010) 90–104.
- [28] S. Fournane, S. Charbonnier, A. Chapelle, B. Kieffer, G. Orfanoudakis, G. Trave, M. Masson, Y. Nominé, Surface plasmon resonance analysis of the binding of high-risk mucosal HPV E6 oncoproteins to the PDZ1 domain of the tight junction protein MAGI-1, *J. Mol. Recognit.* 24 (2010) 511–523.
- [29] K. Zanier, S. Charbonnier, M. Baltzinger, Y. Nominé, D. Altschuh, G. Trave, Kinetic analysis of the interactions of Human Papillomavirus E6 oncoproteins with the ubiquitin ligase E6AP using surface plasmon resonance, *J. Mol. Biol.* 349 (2005) 401–412.

- [30] C. Romier, M. Ben Jelloul, S. Albeck, G. Buchwald, D. Busso, P.H. Celie, E. Christodoulou, V. De Marco, S. van Gerwen, P. Knipscheer, J.H. Lebbink, V. Notenboom, A. Poterszman, N. Rochel, S.X. Cohen, T. Unger, J.L. Sussman, D. Moras, T.K. Sixma, A. Perrakis, Co-expression of protein complexes in prokaryotic and eukaryotic hosts: experimental procedures, database tracking and case studies, *Acta Crystallogr. D Biol. Crystallogr.* 62 (2006) 1232–1242.
- [31] Y. Nominé, T. Ristriani, C. Laurent, J.F. Lefèvre, E. Weiss, G. Travé, Formation of soluble inclusion bodies by HPV E6 oncoprotein fused to Maltose-binding protein, *Protein Expr. Purif.* 23 (2001) 22–32.

## 2 Structure cristallographique de la protéine E6 des Papillomavirus humains (HPV16) en complexe avec le peptide E6APp

### 2.1 Problématique

Comme nous l'avons décrit dans le chapitre introductif, notre équipe a beaucoup travaillé sur la production de E6 en vue d'études structurales. Bien que les structures de ses deux domaines à zinc (E6N et E6C) aient été résolues dans notre laboratoire, celle de la protéine entière restait encore non élucidée à cause de l'auto-oligomérisation de cette dernière.

Ici, je présente la procédure technique qui m'a permis d'obtenir la structure cristallographique de la protéine E6 entière issue du type HPV 16 en complexe avec son peptide préféré portant le motif LxxLL (E6APp). Cette démarche comprend deux étapes majeures :

- Une étape de mutagenèse de E6 qui vise à obtenir un bon candidat à la cristallographie c'est-à-dire une protéine à la fois bien soluble et portant le minimum de mutations ;
- La deuxième étape consistait à l'optimisation de la technique d'expression et de purification de cette protéine par fusion avec un ligand stabilisant (E6APp) et une protéine solubilisante (MBP).

## I.2 Structure cristallographique de la protéine E6 des Papillomavirus humains (HPV16) en complexe avec le peptide E6APp

---

### 2.2 Résultats

#### 2.2.1 Recherche du meilleur mutant

Pour obtenir un bon candidat E6 pour la cristallographie, notre travail de mutagenèse s'est basé sur la procédure ayant produit le mutant E6 F47R 4C4S qui nous semble un bon exemple car ce peptide a été obtenu avec une solubilité supérieure à 350  $\mu\text{M}$ . Comme déjà mentionné (introduction générale), ce mutant résulte du remplacement de la phénylalanine 47 par une arginine dans le domaine N-terminal (E6N), ce qui empêche la dimérisation de ce domaine, et la mutation des quatre cystéines non conservées (E6C) par des sérines aux positions respectives 80, 97, 111 et 140 évitant ainsi l'agrégation de la protéine induite par les ponts di-sulfures entre les groupements sulfhydryles (SH).

En pratique, nous avons d'abord conçu cinq mutants :

- le mutant E6 F47R qui porte une seule mutation, celle du domaine E6N nécessaire pour empêcher l'auto-oligomérisation de E6.









- quatre mutants E6 F47R C80S, E6 F47R C90S, E6 F47R C111S et E6 F47R C140S qui portent chacun deux mutations : la mutation F47R plus une des quatre mutations cystéine-sérine que porte E6 F47R 4C4S sur son domaine E6C.

Ces cinq mutants ainsi la protéine E6 non sauvage et le mutant E6 F47R 4C4S ont ensuite été exprimés et purifiés. Une analyse comparative de la qualité des mutants a été effectuée en se basant sur deux paramètres :

- Le premier paramètre consiste à mesurer la concentration maximale atteinte par chaque mutant avant que celui-ci ne commence son processus de précipitation. Pour rendre ce paramètre plus précis pour la comparaison, nous avons fixé le temps de concentration à 2 heures. Aucun de nos cinq mutants n'a pu atteindre la qualité du E6 F47R 4C4S c'est-à-dire une concentration de précipitation d'environ 350  $\mu\text{M}$ . La concentration la plus proche de cette valeur a été obtenue avec le mutant (E6 F47R C111S) et ne dépasse pas 172,5  $\mu\text{M}$ . La concentration la plus basse, hormis celle du E6 6C6S C (20  $\mu\text{M}$ ) (introduction générale), a été obtenue avec le mutant E6 F47R (75,5  $\mu\text{M}$ ), qui ne porte aucune mutation cystéine sur le domaine C-ter de E6 (Figure I.1). Il est à noter, que dans un travail différent, réalisé dans notre équipe, une solubilité de 400  $\mu\text{M}$  a été atteinte avec le mutant Y43E F47R 4C4S (Zanier *et al.* [2012]).
- Le deuxième paramètre dans cette comparaison est la qualité des spectres générés par la spectroscopie de résonance magnétique nucléaire (RMN) de corrélation à 2D, notamment l'homogénéité et l'intensité des pics (Figure I.2). Ce critère a confirmé les résultats de premier paramètre et a généré à peu près le même ordre de classification des mutants selon



## Chapitre I. Etude biochimique et structurale de la protéine E6

Constructions		C. Max ( $\mu\text{M}$ )
E6wt		
E6 6C/6S		20
E6F47R		75.5
E6 F47R C140S		85
E6 F47R C97S		100
E6 F47R C80S		127.5
E6 F47R C111S		172.5
E6 F47R 4C/4S		350

**Figure I.1 – Comparaison de la solubilité des mutants de E6**

Dans cette figure : Les principaux mutants de E6 utilisés dans ce projet avec leurs concentrations maximales (entre 20 et 350  $\mu\text{M}$ ). Les mutations sont indiquées par des astérisques rouges, E6N et E6C par des rectangles bleu et orange respectivement.

## I.2 Structure cristallographique de la protéine E6 des Papillomavirus humains (HPV16) en complexe avec le peptide E6APp

---

leurs qualités respectives. Ainsi comme on peut voir dans la Figure I.1, le mutant E6 F47R 4C4S possède le meilleur spectre tandis que E6 F47R et E6 F47R C140S sont toujours les plus bas en termes de qualité. Cependant, les trois mutants qui portent en plus de F47R les mutations C80S, C97S ou C140S ont des spectres RMN assez bons (l'identification des différents pics des spectres HSQC est possible), malgré la concentration maximale de précipitation relativement faible de ces mutants. Sur la base de ces données, nous avons choisi le mutant E6 F47R 4C4S  $\Delta$ 8CT sur lequel une dernière modification a été effectuée avant de passer à l'étape de cristallographie : troncation des 8 derniers résidus du E6 qui constituent une région de désordre qui porte le motif de reconnaissance des protéines à PDZ. Cette modification avait pour but de favoriser la cristallisation du mutant.

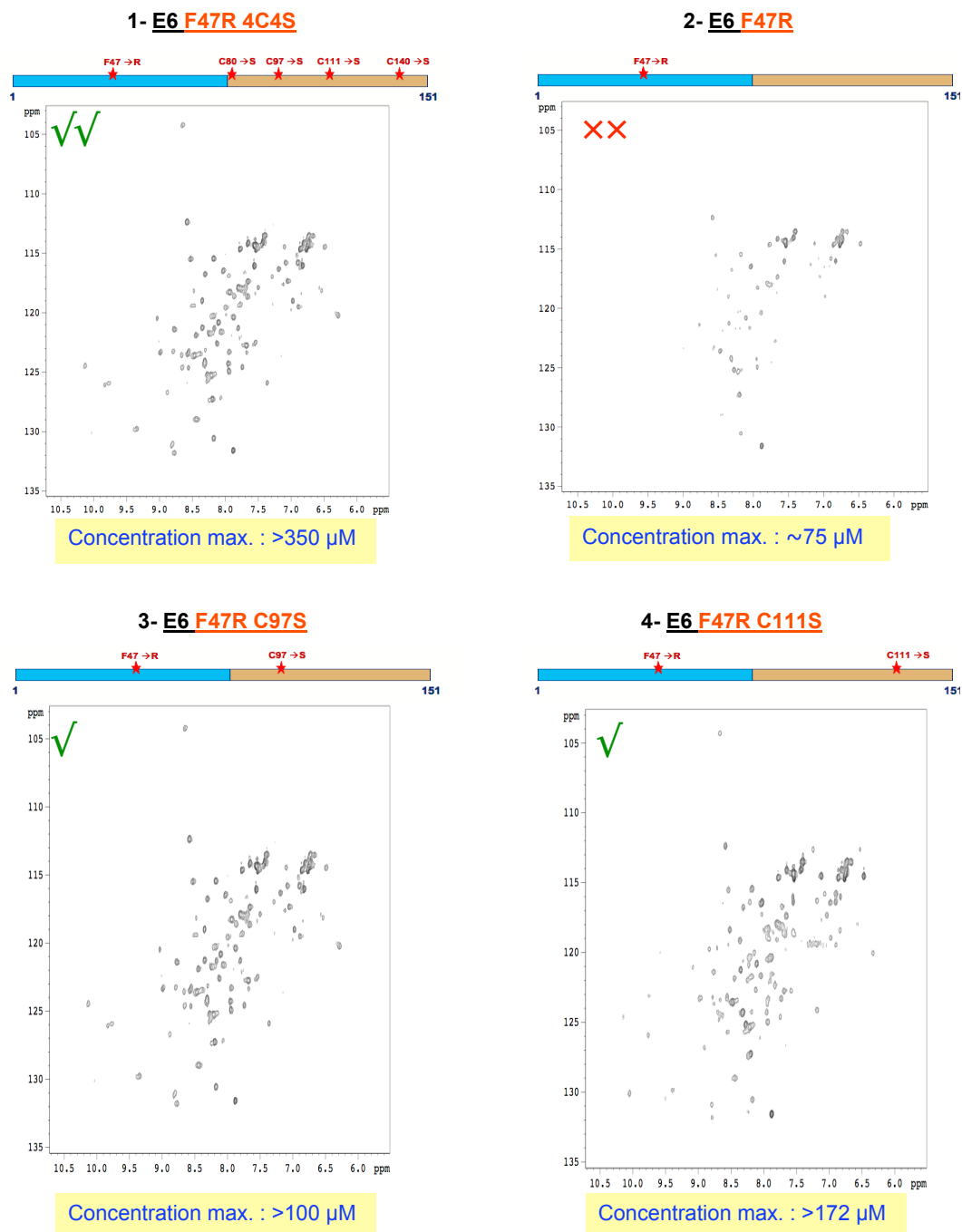
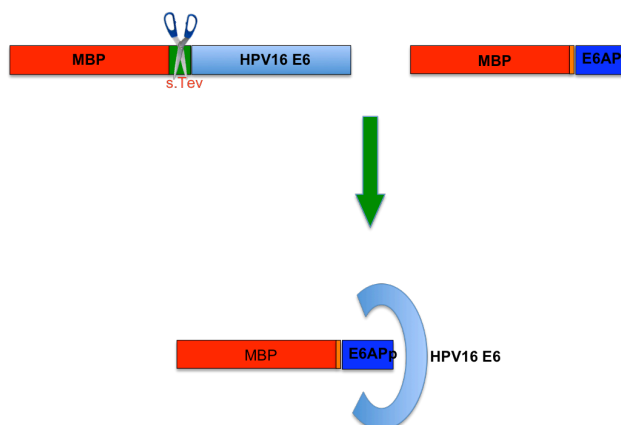


Figure I.2 – Spectres HSQC des principaux mutants E6

Les spectres HSQC (Heteronuclear Simple Quantum Correlation) à deux dimensions ( $H^1N^{15}$ ) correspondants aux différents mutants de E6, Ces spectres montrent la qualité faible du mutant E6F47R (X rouge), la qualité moyenne des mutants E6F47RC97S et E6F47RC111S (1V verte) et la bonne qualité de E6F47R4C4S (2V vertes).

## I.2 Structure cristallographique de la protéine E6 des Papillomavirus humains (HPV16) en complexe avec le peptide E6APp

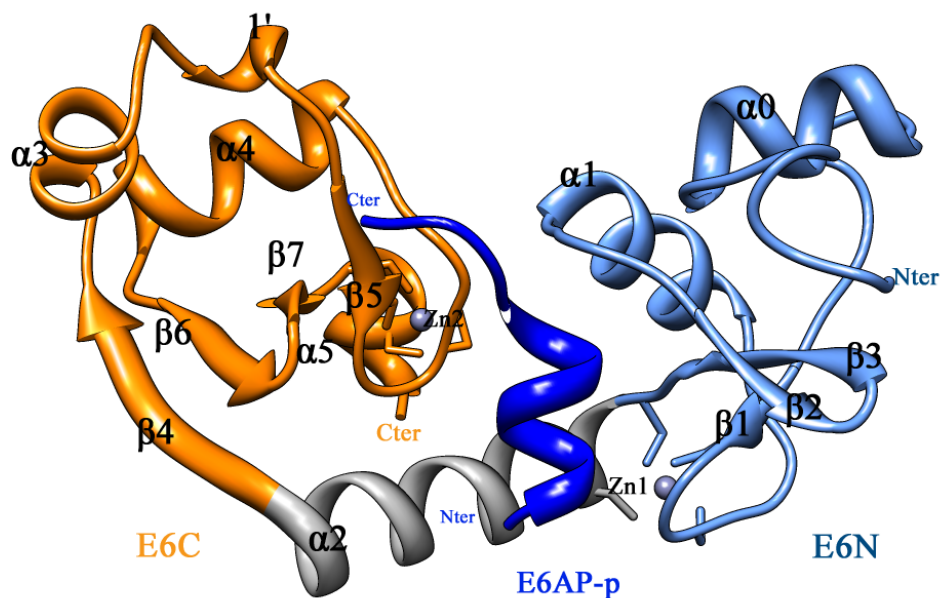


**Figure I.3 – Stratégie de production d’une protéine E6 soluble**

Schéma représentatif de la technique utilisée pour produire une protéine E6 soluble et stable, grâce à la présence d’un volume équimolaire du complexe MBP-E6APp au moment de la concentration de la protéine E6 purifiée et pré-clivée de MBP par l’enzyme de Tev.

### 2.2.2 Stratégie d’étude de la structure de HPV16 E6

Pour résoudre la structure de la protéine HPV16 E6 nous avons appliqué la stratégie de stabilisation de la protéine E6 décrite dans l’article présenté ci-dessus. Cette stratégie a aussi permis à notre équipe de résoudre la structure de E6 du Papillomavirus bovin (BPV1 E6) (manuscrit en cours de rédaction, voir annexe 1). En pratique, le mutant E6 F47R 4C4S  $\Delta$ 8CT a été fusionné - à travers son extrémité N-terminale - avec la protéine MBP (Maltose-Binding Protein) via un court fragment (linker) comportant le site de clivage de l’enzyme TEV. Après expression et purification, la liaison MBP-E6 a été clivée en utilisant l’enzyme TEV. Le produit E6 a été ensuite concentré en présence d’un volume équimolaire du complexe MBP-E6APp dans lequel le peptide E6AP ( séquence : E<sub>1</sub>L<sub>2</sub>T<sub>3</sub>L<sub>4</sub>Q<sub>5</sub>E<sub>6</sub>L<sub>7</sub>L<sub>8</sub>G<sub>9</sub>E<sub>10</sub>E<sub>11</sub>R<sub>12</sub>), est fusionné en C-terminale de la protéine MBP pour améliorer la solubilisation du complexe et favoriser sa cristallisation (Figure I.3). Le complexe MBP-E6APp/E6 a cristallisé selon le groupe d’espace P2<sub>1</sub>2<sub>1</sub>2<sub>1</sub> permettant la diffraction des radiations du synchrotron avec une résolution de 2.6Å. La structure du complexe E6- E6APp a été résolue par remplacement moléculaire en utilisant la structure connue de MBP comme support (voir annexe 2, Tableau 1 et Figure 1). La première méthionine de E6 n’est pas visible dans la carte de densité électronique, probablement du fait de la mobilité du fragment N-terminal non replié de E6.

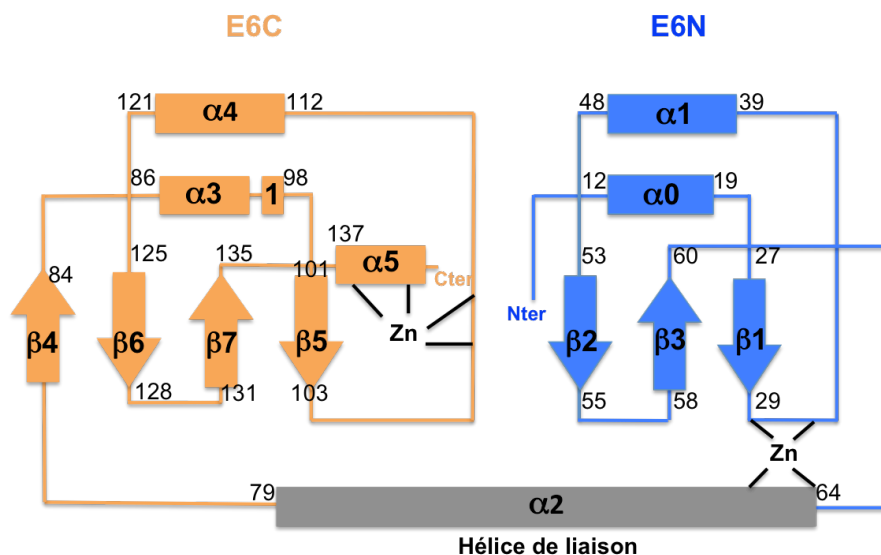


**Figure I.4 – Structure cristallographique de E6 en complexe avec E6APp**  
Représentation "cartoon" de la structure 3D du complexe HPV16 E6-E6APp. Chaque élément structural est coloré différemment : Les deux domaines E6N et E6C sont colorés en bleu et orange respectivement, l'hélice inter-domaines ( $\alpha 2$ ) est grise, le peptide E6AP est coloré en bleu foncé et les molécules de zinc sont représentés par des sphères de couleur violet. 1' est une courte hélice 3/10.

### 2.2.3 Description structurale du cristal HPV16E6/E6APp

La structure du complexe HPV16 E6-E6APp montre que la partie E6 est composée de deux domaines à zinc : le domaine N-terminal (E6N, résidus : 2-72) et le domaine C-terminal (E6C, résidus : 73-143), liés par une hélice alpha relativement longue ( $[\alpha 2]$ , résidus : 64-79). Le repliement du E6C est composé de trois hélices alpha ( $\alpha 3$ ,  $\alpha 4$ , et  $\alpha 5$ ) et quatre régions bêta ( $\beta 4$ ,  $\beta 5$ ,  $\beta 6$  et  $\beta 7$ ) avec une hélice 3/10 (1') tandis que celui du E6N montre deux hélices alpha ( $\alpha 0$  et  $\alpha 1$ ) et trois régions beta ( $\beta 1$ ,  $\beta 2$  et  $\beta 3$ ). Le peptide E6APp, replié en forme d'hélice alpha entre les résidus L2 et L8, est logé dans la poche formée par les deux domaines (E6N et E6C) et l'hélice inter-domaines  $\alpha 2$  (Figure I.4). Les trois leucines ( $L_4$ ,  $L_7$  et  $L_8$ ) du motif LxxLL sont liées à des résidus hydrophobes provenant principalement du domaine E6N et de l'hélice de liaison  $\alpha 2$ . L'interaction du peptide avec E6C est assurée par des liaisons de charge entre le peptide négativement chargé et certains résidus basiques de E6C. Ces liaisons de charge sont

## I.2 Structure cristallographique de la protéine E6 des Papillomavirus humains (HPV16) en complexe avec le peptide E6APp



**Figure I.5 – Topologie structurale de E6**

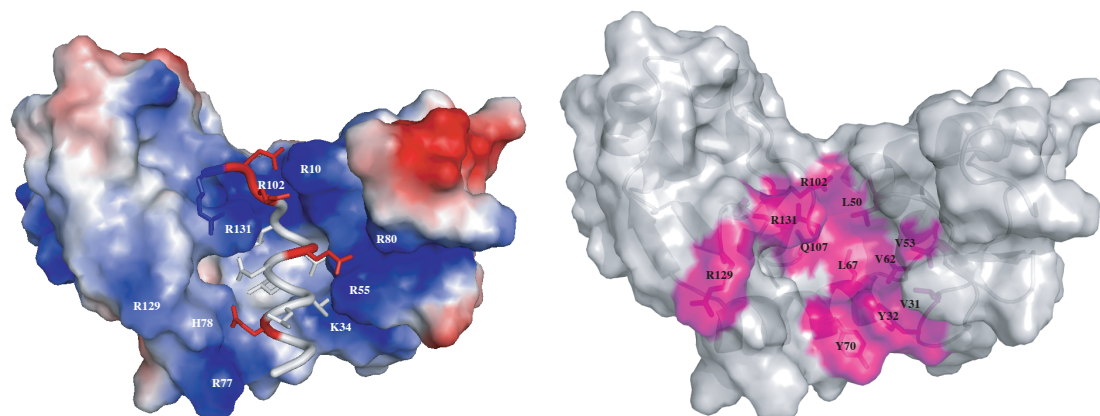
Dans cette figure : Diagramme représentatif de la topologie de la protéine HPV16 E6.

aussi observées le peptide et E6N (Figure I.4, I.5). Nous étudierons plus largement l'interaction entre E6 et les motifs LxxLL dans le chapitre suivant.

## 2.3 Discussion

### 2.3.1 Repliements E6N et E6C

Les repliements de E6N et E6C dans la protéine entière sont très similaires à ceux observés dans les structures en solution (par RMN) des domaines isolés publiés par notre équipe (Nomine *et al.* [2006]; Zanier *et al.* [2012]), à l'exception de certaines fines différences de position des éléments de structure secondaire. En regardant la topologie de cette structure (Figure I.5), on remarque que les deux domaines E6N et E6C partagent en gros les mêmes traits structuraux sauf au niveau de leurs régions N-terminales respectives. Ainsi, la région N-terminale de E6N est une boucle flexible (résidus 2-10), tandis que celle du E6C montre un repliement de forme bêta ( $\beta 4$ ) au niveau des résidus 80-84. Malgré la similarité structurale entre les deux repliements de E6N et E6C, leurs propriétés de charge de surface sont assez dissemblables : la surface de E6N combine des régions acides et des régions basiques alors que celle de E6C est principalement chargée positivement. De plus, concernant l'hydrophobicité de surface, E6N contient deux zones principales d'hydrophobicité. La première, formée par les résidus du brin  $\beta 1$ , de l'hélice  $\alpha 1$  et



**Figure I.6 – Surface électrique et hydrophobe de E6**

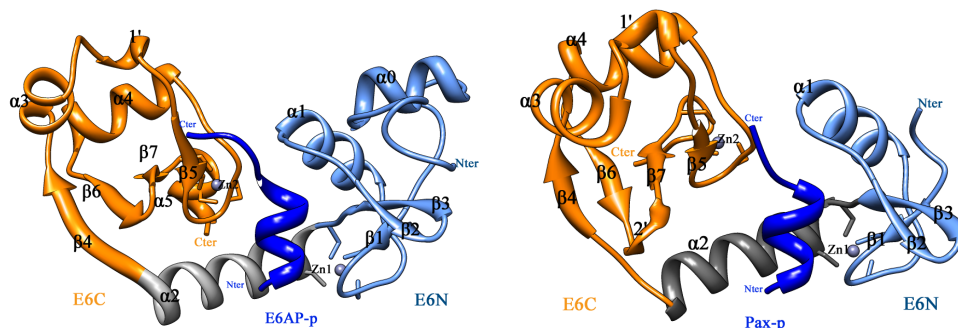
A gauche : Distribution surfacique de charge de HPV16 E6 ; la coloration bleu indique la charge positive et la coloration rouge indique la charge négative. Les résidus acides de E6App sont également colorés en rouge. A droite : La poche hydrophobe de E6, colorée en rose, est formée principalement par les résidus hydrophobes Y32, Y70, L67, Q107, L50, V62, V53... etc., mais aussi par les chaînes latérales des arginines (R102, 129 et 131).

ceux de la zone d'interaction avec le zinc, se trouve en face du peptide E6App. La deuxième région d'hydrophobicité est située dans l'autre face de E6N formée par les résidus de l'hélice  $\alpha 2$  et ceux de la boucle qui s'étend entre  $\alpha 1$  et la région d'interaction avec le zinc. Le domaine E6C présente une seule surface hydrophobe exposée se trouvant entre l'hélice  $\alpha 3$  et la zone de liaison du zinc couvrant le brin  $\beta 5$  (Figure I.6).

### 2.3.2 Comparaison des structures de E6 bovine et humaine

La structure de E6 du type bovine (BPV1) a été aussi résolue par notre équipe (voir annexe 1) à partir de la fusion triple MBP-LxxLL-E6, dans lequel LxxLL est un peptide (séquence : M<sub>1</sub>D<sub>2</sub>D<sub>3</sub>L<sub>4</sub>D<sub>5</sub>A<sub>6</sub>L<sub>7</sub>L<sub>8</sub>A<sub>9</sub>D<sub>10</sub>) extrait de la protéine Paxilline ([Pax-p]). La dispersion spatiale de MBP par rapport à E6 est différente dans les deux complexes MBP-E6App/HPV16 E6 et MBP-Pax.p-BPV1 E6. Cependant, les structures des 2 complexes binaires (E6- LxxLL) sont très similaires dans les deux cas et se superposent avec une valeur RMSD (Ecart quadratique moyen global des distances) d'environ 2Å. Cette similarité entre les deux complexes E6-LxxLL, indépendante de MBP, exclut la possibilité d'une éventuelle perturbation de la structure de E6-LxxLL par MBP ou par les mutations F47R 4C4SΔ8CT de la protéine E6 humaine, sachant que l'autre structure est résolue à partir d'une protéine BPV1 E6 sauvage. De plus, toutes ces

## I.2 Structure cristallographique de la protéine E6 des Papillomavirus humains (HPV16) en complexe avec le peptide E6APp



**Figure I.7 – Comparaison des deux structures 3D de E6 (humaine et bovine)**

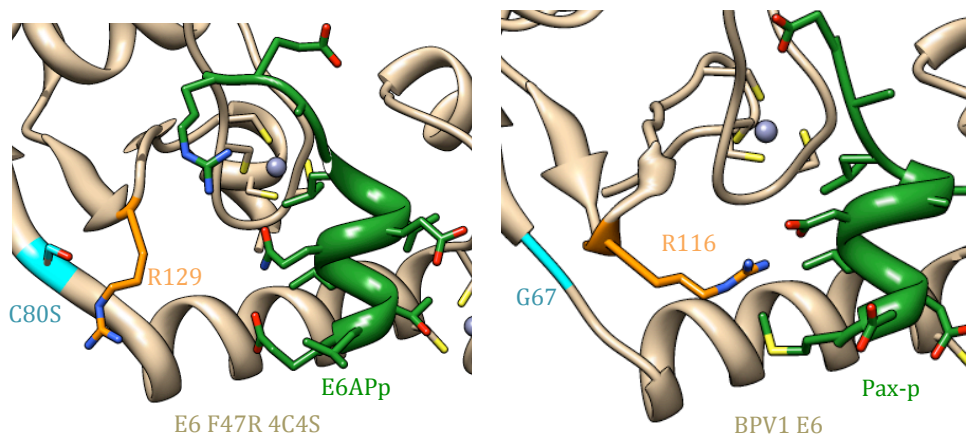
Représentation "cartoon" de la structure 3D de HPV16 E6 avec son peptide ligand E6AP (à gauche) en comparaison avec celle de BPV1 E6 avec son peptide ligand Paxilline (à droite). Notons l'absence de la partie N-terminale qui forme l'hélice  $\alpha_0$  chez BPV1 E6. 1' et 2' sont des courtes hélices 3/10.

mutations sont au niveau des positions exposées au solvant et loin de de l'interface d'interaction avec le peptide LxxLL, sauf la mutation C80S qu'on discutera plus bas. Le domaine E6N de la structure bovine et naturellement plus court par 13 résidus que celui de la protéine E6 humaine. De plus, les 10 premiers résidus E6N bovin ne sont pas repliés et constituent une région de désordre non visible dans la carte de densité électronique de la structure de ce domaine. De ce fait, E6N bovin manque une hélice alpha par rapport à E6N humain ( $\alpha_0$ ) (Figure I.7). Hormis cette région, le reste des deux E6N est très similaire (rmsd : 0.4 Å). La structure du domaine E6C est très conservé entre les deux types de virus avec un rmsd de 1.2Å.

Malgré la différence de séquence entre E6AP-p et Pax-p, les deux peptides se replient en forme d'hélice alpha qui s'étend respectivement entre  $E_1$  et  $L_8$  et entre  $M_1$  et  $L_8$ . Ces hélices alpha sont ancrées à l'intérieur d'une poche formée par les deux domaines (E6N et E6C) et la zone inter-domaine (Figure I.7).

Curieusement, la chaîne latérale de l'arginine 129 du HPV16 E6 est orientée vers le domaine E6C contrairement à son homologue du BPV1 E6 qui est impliquée dans une interaction importante pour le complexe E6-LxxLL avec l'acide aspartique en position 5 (D5) du peptide de Paxilline. Cette orientation de HPV16 R129 pourrait être artificielle, sachant que la mobilité de ce résidu était très élevée dans le cristal. L'orientation de R129 pourrait avoir été perturbée, soit par la mutation HPV16 C80S touchant un résidu spatialement proche de R129, soit par la présence de MBP, sachant que dans le cristal R129 est orientée en direction de résidus de MBP (Figure I.8).





**Figure I.8 – Interfaces d’interactions HPV16 E6-E6AP et BPV1 E6-Paxilline**

Représentation qui montre une éventuelle déviation de l’arginine 129 chez HPV16 E6 sous l’effet de la mutation C80S (à gauche), en comparaison avec son homologue R116 chez BPV1 E6 qui constitue un élément important dans l’interface d’interaction E6-LxxLL.

### 2.3.3 Structure de la forme libre de E6

Les deux structures cristallographiques des complexes E6-LxxLL montrent l’existence de certains résidus très conservés (BPV1 : R89 et R121) qui servent à connecter/liar les deux domaines de E6 entre eux et avec la région inter-domaines tout en contribuant à l’interaction avec le peptide ligand. Ces résidus, pourraient aussi jouer un rôle important dans la forme libre (en cas d’absence de ligand) de E6. Cependant, la nature faible de ces liaisons et la large poche hydrophobe exposée au solvant (en absence du peptide) laissent penser que E6 libre adopterait un autre repliement. Une autre alternative serait que E6 s’oligomériser en conservant le même repliement de la forme monomérique tout en cachant les zones hydrophobes. De nombreuses études effectuées au sein de notre équipe ont en effet démontré que E6 subit un processus d’oligomérisation spontanée après purification (Zanier *et al.* [2010]).

### 2.3.4 Un nouveau mode de reconnaissance des motifs LxxLL

Dans la structure HPV16 E6 et celle de BPV1 E6 nous observons un système d’interaction, dans lequel un motif LxxLL globalement acide et amphipathique se loge sous sa forme d’hélice alpha dans une poche hydrophobe et chargée positivement. La similarité entre les séquences des protéines E6 de plusieurs variants PV et la conservation de certains résidus suggèrent que la structure de la protéine E6 et son mode de liaison des peptides en hélice devraient être conservés dans la plupart des variant de E6 (voir alignement à l’introduction). Cette liaison

## I.2 Structure cristallographique de la protéine E6 des Papillomavirus humains (HPV16) en complexe avec le peptide E6APp

---

constituerait un nouveau mode de reconnaissance des motifs LxxLL par rapport aux modes déjà observés dans les structures résolues des complexes cellulaires formés par ce motif et qui montrent une fixation de ce dernier sur des régions superficielles de la protéine (Hoellerer *et al.* [2003]; Lorenz *et al.* [2008]) plutôt qu'un ancrage du motif LxxLL dans une poche comme nous l'observons pour l'interaction E6-LxxLL.

### 2.3.5 Conclusion

Nous avons présenté dans cette section des résultats intéressants tant au niveau méthodologique que pratique. Nous avons suivi une démarche logique qui a abouti à la résolution de la structure de la protéine E6 du type HPV 16 qui constitue la première information structurale publiée sur la protéine entière de E6 en complexe avec un peptide d'origine cellulaire. Cette démarche a consisté en deux étapes chronologiquement parallèles, à savoir la mutagenèse couplée à l'analyse comparative de qualité et le développement d'une technique de production capables de produire des protéines à haute qualité.

Les données structurales présentées que nous avons présentées démontrent le mode général de reconnaissance des cibles cellulaires des protéines E6, via le recrutement des motifs LxxLL. Il faut préciser ici que ce résultat (Résolution de la structure cristallographique de la protéine HPV16 E6) a été le fruit d'un travail de collaboration avec plusieurs membres de l'équipe Oncoprotéines, ainsi que la plate forme de cristallographie et l'équipe de Mr Jean CAVARELLI de l'Institut de Génétique et de Biologie Moléculaire et Cellulaire à Illkirch (IGBMC).

Ma contribution a consisté à la préparation des échantillons cristallisables depuis le clonage jusqu'à la purification des constructions MBP-E6APp et E6. J'ai également participé aux tests de cristallisation et au suivi des cristaux à la plateforme de cristallographie.



---

---

# Chapitre II

---

## Analyse du système d'interaction de la protéine E6

### 1 Problématique

#### 1.1 Préambule

Comme nous l'avons décrit dans la partie introductive, la protéine E6 reconnaît ses protéines-cibles cellulaires par deux voies principales connues, la fixation du motif X[T/S]X[L/V] du C-terminus de E6 sur les protéines à domaine PDZ ou la reconnaissance, par les deux domaines repliés de E6, de motifs LxxLL. Nous nous intéresserons dans ce chapitre au deuxième mode de reconnaissance (E6-LxxLL). Si les protéines E6 reconnaissent toutes certains de leurs ligands cellulaires via le même motif consensus général LxxLL, il semble que les E6 de chaque type de HPV montrent des variations fines de spécificité pour des sous-classes de motifs LxxLL.

Par exemple, la protéine HPV16 E6 préfère lier la séquence PESSLETLQELLGEER issue de l'ubiquitine ligase E6AP, tandis que la protéine BPV1 E6, phylogénétiquement éloignée de HPV16 E6, préfère lier la séquence DDL DALLADKE issue de la protéine d'adhésion focale Paxilline (Sidi *et al.* [2011]).

Pour effectuer cette étude, nous avons choisi un panel de 7 types viraux qui couvrent six genres de l'arbre phylogénétique des HPV (HPV1,3,5,9,16 et 18) ainsi que la souche bovine BPV1. Il s'agit de trois types du genre  $\alpha$  dont deux à haut risque de tropisme muqueux (HRM) (HPV 16 et 18) et un type à bas risque de tropisme mixte (HPV 03), deux types  $\beta$  cutanés dont l'un à haut risque (HPV05) et l'autre à bas risque (HPV 09), un type  $\mu$  cutané (HPV01), et enfin le type  $\delta$  bovin cutané à haut risque BPV1 (voir classement fig et Tableau). Après avoir cloné, produit et purifié les protéines E6 appartenant à ces 7 types, nous avons testé l'interaction de chaque E6 avec une liste de 93 peptides, dont la moitié est composée de variants

du peptide E6AP (PESELTLQELLGEER) mutés ponctuellement sur des positions-clés (Série 1), et l'autre moitié est issue des protéines cellulaires cibles ou potentiellement cibles de E6 (Série 2).

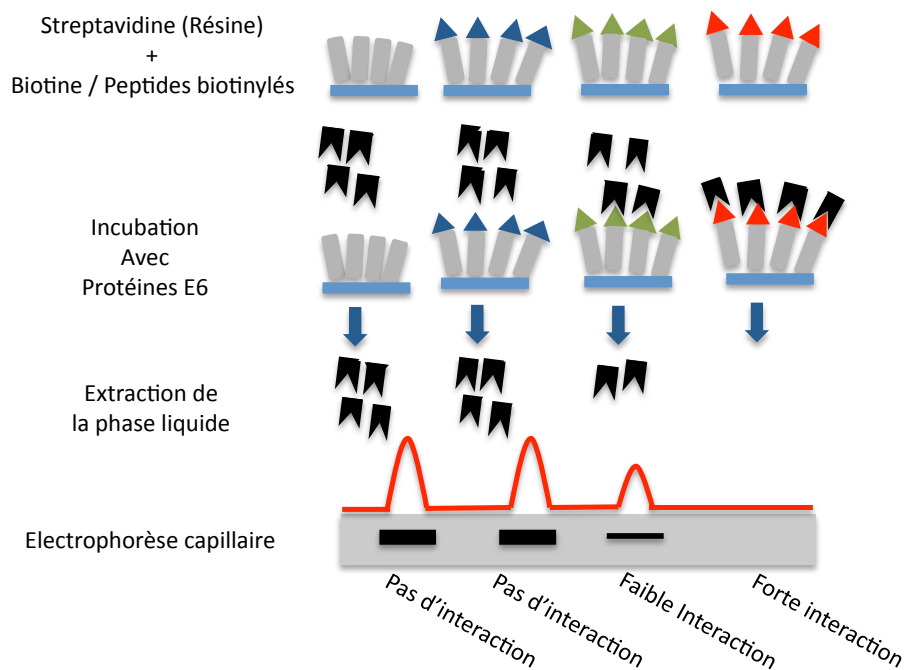
Les résultats de ces tests d'interaction ont fourni une base de données très riche qui constitue avec les données structurales précédemment discutées une bonne assise pour l'analyse des bases séquence-structure de la spécificité et de la sélectivité de chaque variant de E6. Les tests d'interactions ont été faits en utilisant une technique chromatographique développée au sein de notre laboratoire : la technique "hold-up". Avant d'entamer la description des résultats, je me permets ici de faire une petite introduction sur cette technique.

### 1.2 Introduction sur la technique de chromatographie Hold-up

Le Hold-up est une technique de chromatographie comparative de rétention, développée au sein de notre laboratoire ([Charbonnier \*et al.\* \[2006\]](#)). Cette méthode permet d'analyser, dans un format "haut-débit" et de façon relativement quantitative, les affinités de complexes protéine-protéine ou peptide-protéine à l'équilibre. En fait, en éliminant les étapes de lavage utilisées dans d'autres techniques comme le "Pulldown", le Hold-up permet la détection des interactions à dissociation rapide, souvent non détectées par d'autres techniques. En plus de cette particularité, le Hold-up possède d'autres caractéristiques importantes telles que la rapidité, la robustesse, l'adaptation avec les systèmes automatisés à haut débit et la possibilité d'une estimation de la constante de dissociation ( $K_d$ ).

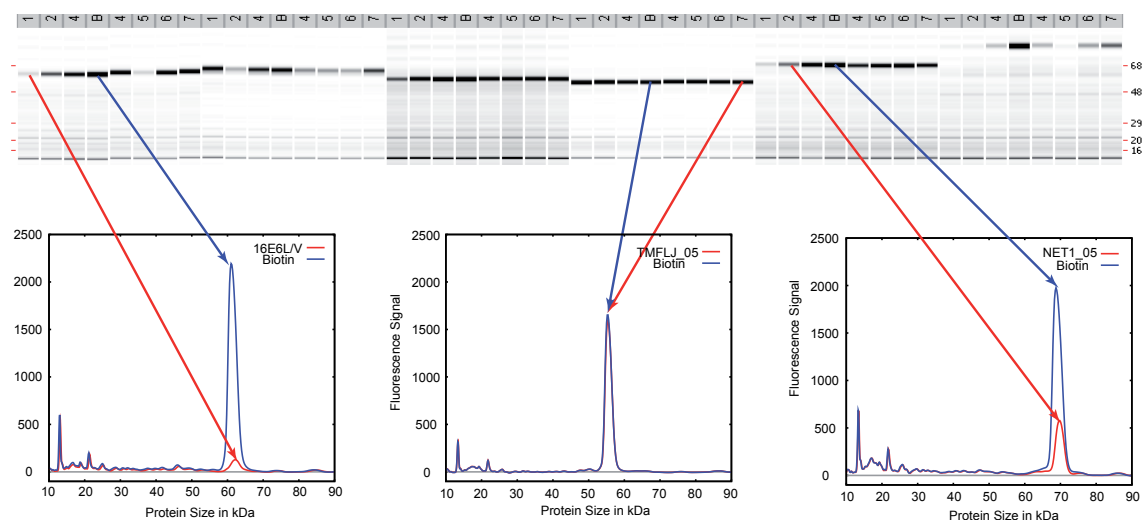
Pour réaliser les tests multiplex présentés dans le préambule (7 protéines contre 93 peptides), nous avons utilisé une version du Hold-up basée sur le système Avidine-Biotine. Les peptides testés sont synthétisés avec une extrémité N-terminale fusionnée à la Biotine, ce qui permet leur immobilisation sur une résine de Streptavidine-sépharose à haute performance (GE-Healthcare). Après la fixation des peptides biotinylés sur la résine Streptavidine, une étape d'incubation avec un excès de Biotine est effectuée pour saturer les sites libres. En parallèle, une autre quantité de résine libre (sans peptides) est incubée avec la Biotine pour servir de contrôle. Les deux types de résines (avec ou sans peptides) sont incubés par la suite pendant 20 min avec les protéines MBP-E6 fraîchement purifiées.

Finalement, le mélange résine-protéine est transféré dans un filtre à centrifuger pour filtrer la phase liquide. Pour chaque test, des échantillons d'un même volume du filtrat liquide sont analysés par électrophorèse. Pour chaque peptide considéré, la comparaison entre la quantité de E6 révélée dans le filtrat passé à travers la résine saturée en peptide et dans le filtrat passé à travers la résine saturée en biotine permet alors la quantification de la fraction de E6 qui



**Figure II.1 – Principe de la technique de Hold-up**

Schéma simplifié explicatif de la technique de chromatographie Hold-up, montrant les étapes principales : L'incubation de la résine streptavidine (gris) avec les peptides biotinylés (bleu, vert et rouge), l'incubation de la résine saturée avec la protéine (noire), l'extraction de la phase liquide et l'analyse de l'output grâce à l'électrophorèse. (d'après Luck K./Charbonnier S. et al. manuscrit en cours de préparation).



**Figure II.2 – Technique de Hold-up automatisée**

Figure explicative du Hold-up automatisé couplé à l'électrophorèse capillaire à haut débit, qui offre la possibilité d'utilisation de faibles quantités, une séparation plus rapide que l'électrophorèse classique et une détermination numérique (sous forme de pics) de la taille et de la quantité de chaque protéine testée. à titre explicatif nous montrons ici trois exemples différents : d'un très bon ligand (à gauche), un non-ligand (au milieu) et un bon ligand (à droite) (d'après Luck K. / Charbonnier S. et al. manuscrit en cours de préparation).

a interagi spécifiquement avec le peptide (Figure II.1). L'électrophorèse pour déterminer les quantités de protéines MBP-E6 dans les filtrats peut être effectuée soit par SDS-PAGE suivi d'une quantification par scanner, soit d'une façon plus automatisée et plus précise par électrophorèse capillaire à haut débit. L'électrophorèse capillaire offre une possibilité d'utilisation de faibles quantités, une séparation plus rapide que l'électrophorèse classique et une détermination numérique de la taille et de la quantité de chaque protéine testée. Dans les travaux montrés ici nous avons principalement utilisé l'électrophorèse capillaire, effectuée sur un appareil Caliper " LabChip GX " (Figure II.2).



## 2 Résultats

### 2.1 Interactome E6 - LxxLL

Nous présentons ici deux Séries de tests d'interactions Hold-up appliqués à un panel de protéines E6 de six types HPV (HPV1,3,5,9,16 et 18) ainsi qu'à la protéine E6 du type bovin BPV1 contre les listes, précédemment décrites, de 93 peptides présentant le motif LxxLL ou un motif proche. La version de la protéine HPV16 E6 utilisée dans cette étude est celle du mutant E6 F47R 4C4S pour les raisons de solubilité précédemment discutées. Nous avons aussi utilisé une version de HPV05 E6 tronquée de 22 résidus sur l'extrémité N-terminale (HPV05 $\Delta$ 22). En effet, nous avons précédemment constaté que HPV05 E6 avait une forte tendance à s'agréger, rendant sa purification très difficile. Or, HPV05 E6 possède un domaine N-terminal plus long que tous les autres types de E6 (voir alignement) et nous avons pensé que cette extension pourrait être l'une des causes de la faible solubilité de HPV05 E6. Nous avons donc conçu cette délétion du bras N-terminal en espérant qu'elle pourrait augmenter la solubilité de HPV05 E6. Hormis ces deux mutants (HPV16 E6 F47R 4C4S et HPV05 $\Delta$ 22), toutes les autres protéines E6 utilisées dans ces tests présentent des séquences sauvages.

Toutes les protéines E6 ont été produites et purifiées sous forme de fusion à la protéine solubilisante bactérienne MBP, et soumises, juste avant d'effectuer les mesures, à une ultracentrifugation à haute vitesse et de longue durée permettant d'éliminer les agrégats solubles inactifs de fusions MBP-E6. Il faut noter que les résultats de certains tests ne sont pas présentés, ni discutés ici suite à des problèmes de qualité ou d'indisponibilité des données. En particulier, une grande partie des tests effectués sur les protéines E6 de deux types HPV16 et BPV1 de la deuxième Série n'ont pu être analysés.

De plus, deux autres Séries de tests effectués sur la protéine HPV05 E6 entière ainsi que sur la protéine HPV11 E6 (bas risque muqueux) n'ont pu être analysés suite à des problèmes récurrents sur leur traitement par l'appareil Calliper. Notons enfin que les données montrées ici sont préliminaires puisqu'elles n'ont pas encore été reproduites en triplicat. Toutefois, nous pensons que ces données sont relativement fiables car les résultats concernant les différentes variations de peptides et les différentes E6 sont extrêmement cohérents.

### 2.1.1 Série 1 : mutants de E6APp

La première Série de ces tests a été effectuée sur les protéines E6 des six types HPV (HPV1,3,5,9,16 et 18) contre une panoplie de 46 peptides LxxLL résultant de la mutagenèse du peptide de liaison à E6 de E6APp (16 acides aminés). Chaque peptide porte une mutation unique par rapport à la version sauvage du peptide E6AP. La version du peptide E6AP utilisée pour la mutagenèse compte 16 résidus (séquence : PESS<sub>E1</sub>L<sub>2</sub>T<sub>3</sub>L<sub>4</sub>Q<sub>5</sub>E<sub>6</sub>L<sub>7</sub>L<sub>8</sub>G<sub>9</sub>E<sub>10</sub>E<sub>11</sub>R<sub>12</sub>). et est donc plus longue de quatre résidus du côté N-terminal par rapport au peptide E6AP présent dans la structure de HPV16 E6/E6APp discutée dans le chapitre précédent. Cette version longue lie la protéine HPV16 E6 mieux que la version courte (résultats non publiés).

Nous avons joué sur plusieurs paramètres dans cette carte de mutations, à savoir la charge, la taille et l'hydrophobicité des résidus, etc... . Les résultats de cette Série de tests Hold-up présentés sous forme de représentation "Heat map" (Figure II.3) montrent que tous les types de E6 testés interagissent bien avec la version sauvage du peptide E6AP sauf la protéine HPV05 E6Δ22. La faible liaison de l'ensemble des peptides par HPV 05 E6Δ22 pourrait avoir différentes origines, peut-être combinées. Tout d'abord, nous avons noté lors de la purification que la protéine HPV05 E6Δ22 a tendance à s'agréger, comme la protéine HPV05 E6 complète. Bien que l'étape d'ultracentrifugation ait permis d'éliminer une grande partie des particules agrégées, de nouveaux agrégats potentiellement inactifs peuvent s'être formés dans le laps de temps (entre 30 mn et 3h) séparant la fin de la centrifugation du début des mesures. Il se peut donc que la préparation utilisée contienne une fraction inactive, qui pourrait atteindre 50% du matériel puisqu'aucun peptide n'a retenu plus de 50% de la préparation de HPV05 E6Δ22. Cependant, il est également possible que l'ensemble des peptides testés ne correspondent pas aux motifs préférentiellement liés par HPV05 E6. Enfin il serait même envisageable que HPV05 E6 ne reconnaisse aucun motif peptidique en hélice. Du point de vue phylogénétique, cette dernière hypothèse semble cependant peu probable, sachant que toutes les autres protéines E6 que nous avons testées, incluant la E6 de BPV1 dont la séquence est très éloignée de toutes les séquences de E6 de HPV, reconnaissent des peptides de type LxxLL.

Les "Heat map" (Figure II.3 et II.4) ainsi que les histogrammes (Figure II.5) montrent que les E6 des types HPV01 et HPV03 lient bien voire très bien presque tous les mutants de E6AP. L'explication la plus plausible est que ces deux protéines E6, en particulier HPV01 E6, lient le peptide E6AP sauvage avec une très forte affinité. Or, la méthode holdup, dans les conditions standard utilisées ici, génère une intensité maximale pour toute interaction présentant une constante de dissociation à l'équilibre  $K_d < 10 \mu\text{M}$ . Si, par exemple, l'on suppose qu'une protéine E6 lie le peptide E6AP avec un  $K_d = 35 \text{ nM}$ , comme observé pour le complexe BPV1 E6 /

## Chapitre II. Analyse du système d'interaction de la protéine E6

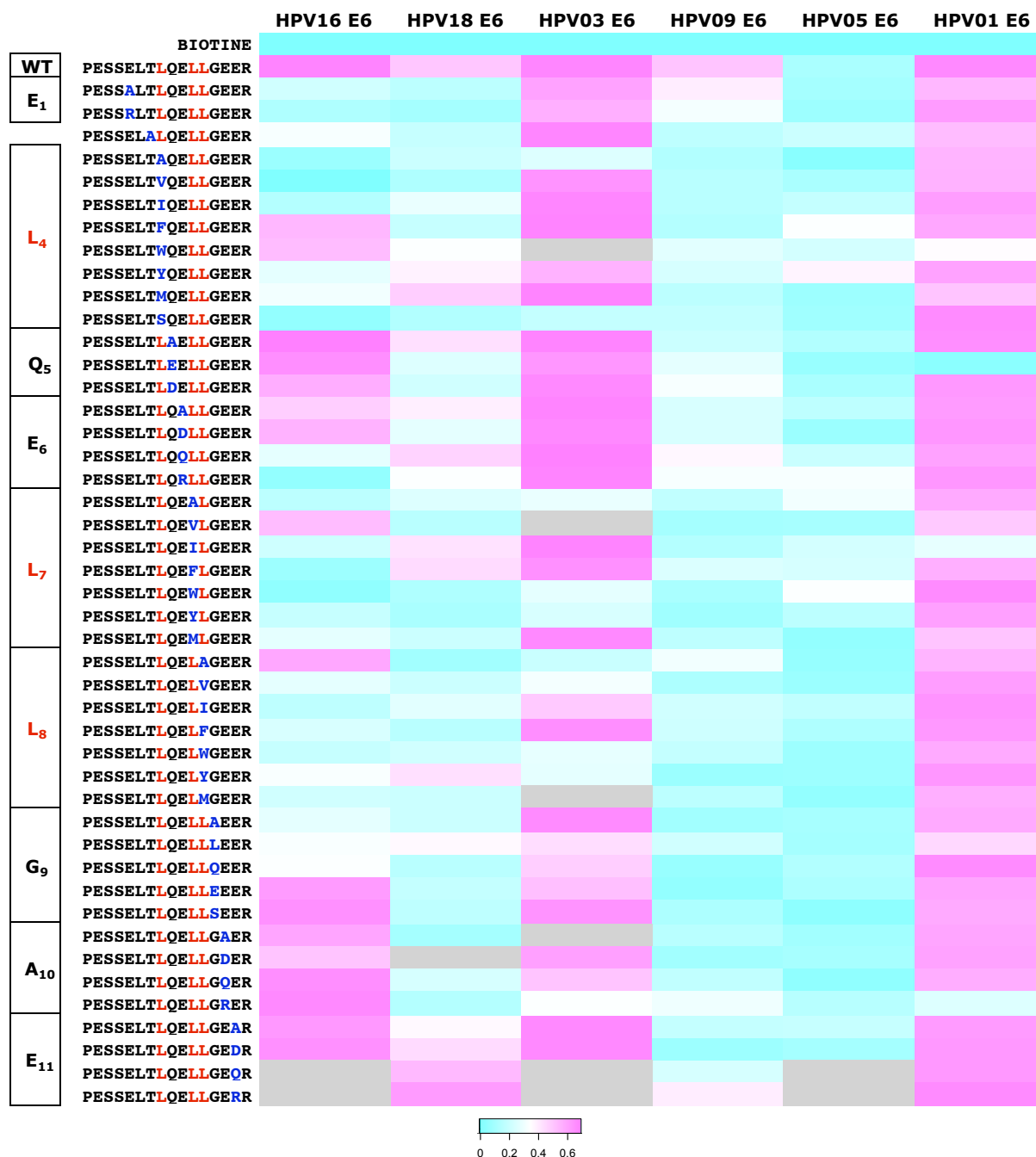


Figure II.3 – Heat map des résultats de Holdup - Série 1 (types confondus)

Cette Heat map montre les résultats des tests d'interactions entre les protéines E6 issues de 6 types viraux (HPV16, 18, 3, 9, 5 et 1) et le peptide E6AP (wt) et ses 45 mutants (mutations indiquées en bleu). La couleur cyan signifie la non-interaction ou interaction faible, le blanc une interaction moyenne et le rouge une bonne interaction. Les cases correspondant de données manquantes sont colorées en gris.

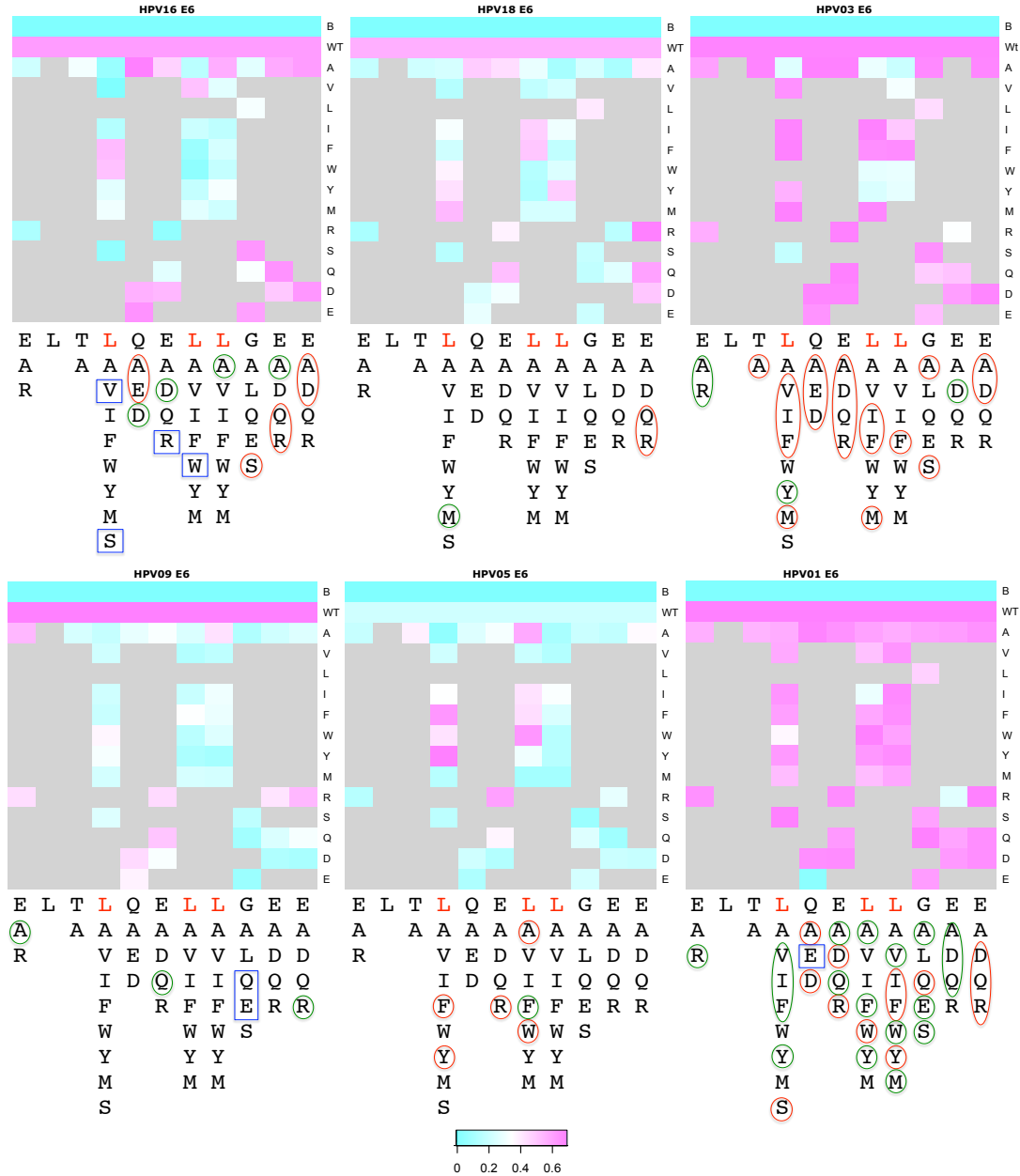


Figure II.4 – Heat map des résultats de Holdup - Série 1 (types séparés)

Nous présentons ici, les cartes "Heat map" de chaque type HPV, séparées et rangées selon l'ordre de mutations par résidus de E6APp. Les mutations les plus importantes sont indiquées par des cercles rouges (> 90%) ou verts (80% - 90%) et des carrés bleus pour les très faibles ligands (< 10%). Dans le cas des E6 des types HPV05 et HPV09 nous avons considéré comme superligands les mutants qui lient E6 avec un pourcentage de 40% - 50% à cause de l'interaction faible observée chez ces deux types.

## Chapitre II. Analyse du système d'interaction de la protéine E6

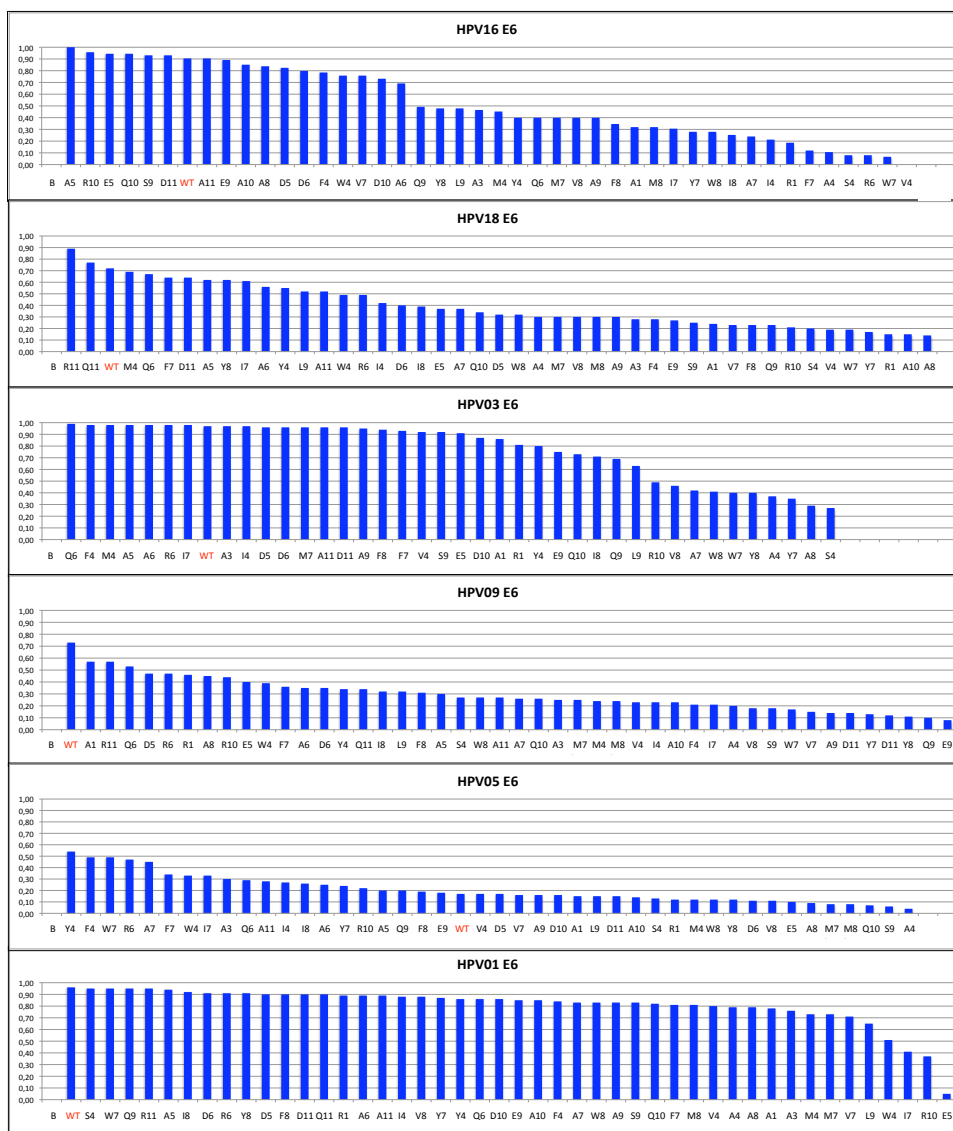


Figure II.5 – Résultats de "Holdup" - Série 1 (Histogrammes)

Dans cette figure, les résultats de Série 1 sont présentés sous forme d'histogrammes décroissants. Cette représentation permettra une meilleure visualisation de la sélectivité de chaque type de E6. Les peptides sont nommés par leurs mutations (A1, R11, Q6... etc.). Le peptide E6AP sauvage est indiqué par un WT rouge.

paxilline (Sidi *et al.* [2011]), seules les mutations induisant une perte d'affinité d'au moins deux unités Log ( $K_d' / K_d > 100$ ,  $K_d'$  étant la constante d'affinité du mutant) montreront un début de perte de signal en test holdup. Il apparaît donc probable que les interactions de HPV01 E6 et HPV03 E6 avec le peptide E6AP sont relativement fortes, avec un  $K_d < 500$  nM, peut-être même de l'ordre de  $K_d = 50$  nM pour HPV01 E6 au vu des fortes intensités de liaison observées pour ce variant avec la majorité des peptides E6AP mutés.

Par contre, si une protéine E6 lie le peptide E6AP sauvage avec une affinité moyenne, de l'ordre de  $K_d = 5 \mu\text{M}$ , alors toute perte d'affinité, même minimale ( $K_d' / K_d > 5$ ) sera détectée par le test holdup. Ce principe se vérifie dans le cas du mutant HPV16 E6 F47R 4C4S. Nous avons montré par BIAcore et ITC (Zanier, communication personnelle) que le mutant HPV16 E6 F47R 4C4S lie le peptide E6AP avec une affinité de l'ordre de  $K_d = 5 \mu\text{M}$ . EN accord avec ces données, dans nos tests holdup, HPV16 E6 F47R 4C4S montre une bonne interaction avec a peu près un tiers des mutants lie moyennement un tiers, et se lie très faiblement ou pas au dernier tiers.

Quand à HPV18 E6, de type à haut risque muqueux comme HPV16, elle montre, vis à vis des peptides mutants de E6AP (Figures) des intensités d'interaction plus faibles en moyenne que celle de HPV16 E6 4C4S F47R, mais avec un "spectre" d'interaction assez comparable. L'intensité d'interaction plus faible pourrait être imputable, comme dans le cas de HPV05 mais dans une moindre mesure, à une inactivité partielle qui pourrait affecter 10-20% de la préparation de HPV18 E6.

Concernant la E6 du type HPV09, elle lie faiblement voire très faiblement la plupart des peptides variants de E6AP y compris E6APp sauvage (WT) (Figure II.3 et II.4). Ceci ne peut être imputé à un problème de qualité d'échantillon car HPV09 E6, contrairement à HPV05 E6 et à HPV18 E6, lie assez fortement plusieurs peptides de la Série 2 (Figure II.6). HPV09 E6 a donc visiblement une faible affinité pour le motif LxxLL de E6AP et ses proches variants.

Pour rendre plus claires les différences d'affinités, nous avons présenté séparément les cartes "Heat map" de chaque type HPV et nous les avons rangées selon l'ordre de mutations par résidus de E6APp. Pour mieux visualiser l'effet de chaque mutation nous avons indiqué les mutations les plus marquées par des cercles ou ellipsoïdes rouges (très bons ligands,  $I > 90\%$ ) ou verts (bons ligands,  $80\% < I < 90\%$ ) et des carrés bleus pour les très faibles ligands ( $I < 10\%$ ). Dans le cas des E6 des types HPV05 et HPV09 nous avons abaissé les seuils d'intensités et considéré comme super-ligands les mutants qui lient E6 avec respectivement un pourcentage de 40% et de 50% à cause de l'interaction faible observée chez ces deux types (Figure II.4).

### 2.1.2 Série 2 : autres peptides LxxLL d'origine cellulaire

En plus des tests effectués sur les mutants de E6AP<sub>p</sub>, nous avons réalisé une deuxième Série de tests Hold-up d'interaction entre les protéines E6 issues des types HPV1,3,5,9,16,18 et BPV1 et une liste de 47 peptides d'origine cellulaire. Ces peptides représentent des séquences existant naturellement dans des protéines cellulaires, et qui comportent le motif LxxLL ou des motifs similaires (qui ont un résidu hydrophobe à la place d'une ou plus des 3 leucines du motif LxxLL). Nous les avons sélectionnés de la façon suivante. Nous avons d'abord dressé une liste des protéines publiées comme cibles de E6, et ne contenant pas de domaines PDZ. Dans les protéines de cette liste, nous avons cherché des séquences consensus LxxLL ou séquences proches, en tenant bien sûr compte des motifs LxxLL déjà identifiés dans ces protéines. Dans la plupart des cas, nous n'avons retenu que des séquences LxxLL ou assimilées présentes dans des zones prédites ou avérées en tant que régions non repliées, en suivant le principe que des motifs présents dans un domaine replié ne sont pas à priori disponibles pour s'engager dans une interaction (Diella *et al.* [2008])<sup>1</sup>.

En plus de ces 20 peptides nous avons rajouté 10 autres peptides potentiellement reconnus par E6, provenant d'une étude prédictive des ligands de BPV1 E6 (voir annexe 1), ou de communications personnelles non publiées :

#### Ligands publiés de E6 :

- ADA3 (Transcriptional adapter 3 [IPI00220486]) ;
- BRCA1(Breast cancer type 1 susceptibility protein [IPI00945705]) ;
- CBP (CREB-binding protein [IPI00023339]) ;
- E6AP (Ubiquitin-protein ligase E3A [IPI00011609]) ;
- E6BP (Reticulocalbin-2 [IPI00029628]) ;
- E6TP1(Signal-induced proliferation-associated 1-like protein 1[IPI00453688]) ;
- GPS2 (G protein pathway suppressor 2 [IPI00012301]) ;
- IRF3 (Interferon regulatory factor 3 [IPI00291901]) ;
- MAML1(Mastermind-like protein 1[IPI00006108]) ;
- MED13 (Mediator of RNA polymerase II transcription subunit 13 [IPI00021388]) ;
- P300 (Histone acetyltransferase p300 [IPI00020985]) ;
- p53 (humain, bovin et aviaire) ;
- Piasy (E3 SUMO-protein ligase PIAS4 [IPI00166500]) ;
- Pxn ([IPI00335727]) ;
- SGT1 (Protein ecdysoneless homolog [IPI00027034]) ;

---

1. la prédiction des régions de désordre a été réalisé par le logiciel libre accès en ligne : IUPred

- TGFB1(TGFB1-induced anti-apoptotic factor 1[IPI00294690]);
- TRIP-BR1(SERTA domain-containing protein 1[IPI00305990]);
- TSC2 (Tuberin [IPI00028493]);
- XRCC1(DNA repair protein XRCC1[IPI00002564]);
- ZyXIN ([IPI00926625]).

### **Ligands potentiels de E6 :**

#### **- prédictions E6 :**

- AP1B1(Adaptor-related protein complex 1, bêta 1 subunit [IPI00384489]);
- AP1G1(Adaptor-related protein complex 1, gamma 1 subunit [IPI00816366]);
- DBC1(DBIRD complex subunit KIAA1967 [IPI00182757]);
- EBP1(Proliferation-associated protein 2G4 [IPI00299000]).

#### **- Communication personnelles et congrès :**

- DNJA1(DnaJ homolog subfamily A member 1[IPI00012535]);
- EMERIN (EDMD [IPI00032003]);
- PAB1(Polyadenylate-binding protein 1[IPI00008524]);
- WIZ (Widely-interspaced zinc finger-containing protein [IPI00295502]).

La plupart de ces peptides sont des 12-mères, à l'exception des quatre peptides p53bos-2, p53chick-2 et TRIP-Br1-5 qui sont plus longs (13 à 16 résidus) car ils possédaient deux motifs LxxLL possibles partiellement chevauchants, tandis que le peptide MAML1 est plus court (9 résidus) car son motif LxxLL est situé à l'extrémité C-terminale de la protéine MAML1 (Tableau II.1). En plus de ces peptides d'origines cellulaires nous avons conçu et rajouté à cette liste le peptide ' E6APxn2 '. Il s'agit d'une "chimère" entre les peptides E6AP et PXN2, qui étaient apparus parmi les meilleurs ligands de HPV16 E6 lors de tests préliminaires.

Nous avons obtenu des données d'interaction avec la majorité des 47 peptides de la Série 2 pour les protéines E6 de HPV1, 3, 5 ( $\Delta 22$ ), 9 et 18. Par contre, les protéines BPV1 E6 et HPV16 E6 F47R4C4S n'ont pu être testées que contre 16 peptides. Les résultats de cette Série, présentés sous forme d'un tableau de signes (Tableau II.2) ou sous forme d'histogrammes (Figure II.6) montrent que globalement les protéines E6 interagissent moins fortement avec ces peptides d'origine cellulaire (Série 2) qu'avec la Série de mutants de E6AP (Série 1).

Ceci était attendu à cause de la diversité plus large des peptides de cette Série provenant d'origines différentes par rapport à la première Série composée seulement de mutants ponctuels uniques de E6APp. Par exemple, pour la protéine HPV03 E6, le nombre de "bons ligands" est très faible dans la deuxième Série de peptides LxxLL par rapport à la première.



## Chapitre II. Analyse du système d'interaction de la protéine E6

Tableau II.1 – Liste de peptides utilisés en Série 2

PROTEINES	PEPTIDES	
ADA3	ADA3_1	QLELETLLSSAS
	ADA3_2	VRTLEEELKPPPE
AP1B1	AP1B1	GGGLDSLMGDEP
AP1G1	AP1G1_1	GGELLDLLGDIN
	AP1G1_2	ANDLLDLLGGND
BRCA1	BRCA1_1	IQKVNEWFSRSD
	BRCA1_2	FSRSEDELLGSDD
	BRCA1_3	ERTLKYFLGIAG
CBP	CBP	HKQLESELLRGGGS
DBC1	DBC1	KSQQLQRLLQELR
DNJA1	DNJA1	LSLEKLLPERK
E6AP	E6AP	PESSELTLOELLGEER
E6BP	E6BP	GFVSLSEFLGDY
E6TP1	E6TP1	PTKLSDFLITGG
EBP1	EBP1	DAELKALLQSSA
Emerin	Emerin	DTELTTLRRYN
Gps2	Gps2	FLQLKKVLHEEE
IRF3	IRF3	EDILDELLGNMV
MAML1	MAML1	MSDLDDLGSQ
MED13	MED13	YTDLDNLFNSDE
p300	p300	QQALQNLLRTLRL
p53	p53	FSDLWKLLEN
p53bos	p53bos_1	FSDLWNLLEN
	p53bos_2	SOETFSDLWNLLEN
p53chick	p53chick_1	FMDLWSMLPYSM
	p53chick_2	TEVFMDLWSMLPYSMQQ
PAB1	PAB1	DNSELHMLLESP
Piasy	Piasy_1	VSDLQMLLGFVG
	Piasy_2	FNMLDELLKPTE
PXN	PXN_1	MDDLDAALLADLE
	PXN_2	LSELDRLLELN
	E6APxn2	PESSELDRLLELN
	PXN_4	TRELDLMLASLS
SGT1	SGT1	PDQLDQLQEA
TGFB1	TGFB1_1	MEDDALLSDLE
	TGFB1_2	LCELDRLQELN
	TGFB1_3	TLELDRLMLASLS
TRIP-Br1	TRIP-Br1_1	SASMASLLEDLS
	TRIP-Br1_2	DDGLEGLFEDID
	TRIP-Br1_3	EAELDYLMDVLV
	TRIP-Br1_4	LDYLMDVLVGTQ
	TRIP-Br1_5	EAELDYLMDVLVGTQ
TSC2	TSC2_1	LQTLQDILGDPG
	TSC2_2	ELQTLQDILGDP
WIZ	WIZ	INILQELLATSA
XRCC1	XRCC1	PEELGKILQGVV
Zyxin	Zyxin	IDSLSLLDDMT

Ce tableau présente les séquences des 47 peptides utilisés dans la deuxième Série de tests Holdup, ainsi que les noms de leurs protéines d'origine (30 protéines).

Tableau II.2 – Résultats de "Holdup" - Série 2 (Tableau)

PEPTIDES	$\alpha$		$\beta$			$\mu$	$\delta$																
	Muqueuses		Mixte	Cutanées																			
	HPV16	HPV18	HPV03	HPV09	HPV05	HPV01	BPV01																
ADA3_2	////	+/-	+/-	+	+/-	+++	////																
Piasy_2	////	+/-	+	+++	+/-	++	////																
p300	////	+	+	+	+	+++	////																
WIZ	////	+	+	+	+/-	+	////																
CBP	////	////	+/-	+	+/-	+	////																
E6APxn2	PESL	SELD	RLL	LELN	-	+	++																
PXN_2	----	SELD	RLL	LELN	-	+++	++																
TGFB1_2	----	LCELD	RLL	QELN	-	++	++																
DBC1	----	KSQ	LQR	LL	QELR	-	////																
SGT1	----	P	DQ	L	DQ	LL	QEAV	-	////														
XRCC1	----	P	EEL	GK	I	L	Q	G	V	-	////												
E6AP	PESSEL	T	L	Q	E	L	G	E	E	R	-	+++	++										
Gps2	----	F	L	Q	L	K	K	V	L	H	E	E	-	////									
p53bos_2	SQET	F	S	D	L	W	N	L	L	P	E	N	-	////									
p53	----	F	S	D	L	W	K	L	L	P	E	N	-	-									
p53chick_1	----	F	M	D	L	W	S	M	L	P	Y	S	M	-	////								
p53chick_2	-	T	E	V	F	M	D	L	W	S	M	L	P	Y	S	M	Q	-	////				
AP1G1_2	----	A	N	D	L	D	L	G	G	N	D	-	-	-	++	+++	////						
IRF3	----	E	D	I	L	D	E	L	L	G	N	M	V	-	+/-	+/-	+	+	+/-	++	++		
MAML1	----	M	S	D	L	D	D	L	L	G	S	Q	-	+	++	+	++	+	+++	+++	////		
Piasy_1	----	V	S	D	L	Q	M	L	L	G	F	V	G	-	////	////	////	////	////	////	////		
EBP1	----	D	A	E	L	K	A	L	L	Q	S	S	A	-	-	////	++	+/-	-	+	////		
Emerin	----	D	T	E	L	T	T	L	R	R	Y	N	-	////	+	+	-	+/-	+++	////	////		
ADA3_1	----	Q	L	E	L	E	T	L	L	S	S	A	S	-	////	+	+/-	+	-	-	////		
PXN_4	----	T	R	E	L	D	E	L	M	A	S	L	S	-	-	+	+/-	+++	+/-	+	++	////	
TGFB1_3	----	T	L	E	L	D	R	L	M	A	S	L	S	-	-	+/-	+	+	+/-	+++	++	////	
TRIP-Br1_4	----	L	D	Y	L	M	D	V	L	V	G	T	Q	-	////	+	+	+	+/-	+++	////	////	
TRIP-Br1_5	----	E	A	E	L	D	Y	L	M	D	V	L	V	G	T	Q	-	////	++	+/-	+++	////	////
TRIP-Br1_3	----	E	A	E	L	D	Y	L	M	D	V	L	V	-	-	+	+	+/-	+/-	++	++	////	
PXN_1	----	M	D	D	L	D	A	L	L	A	D	L	E	-	++	-	+	++	+/-	++	+++	////	
TGFB1_1	----	M	E	D	L	D	A	L	L	S	D	L	E	-	++	+	+	+/-	+/-	++	+++	////	
PAB1	----	D	N	S	E	L	L	H	M	L	E	S	P	-	////	+/-	+	+	+/-	++	+++	////	
p53bos_1	----	F	S	D	L	W	N	L	L	P	E	N	-	////	+	+/-	+/-	+	+++	+++	////	////	
AP1B1	----	G	G	L	D	S	L	M	G	D	E	P	-	-	+	+++	++	+/-	+++	++	////	////	
AP1G1_1	----	G	G	E	L	D	L	L	G	D	I	N	-	-	+/-	+++	++	+/-	++	+++	+++	////	
TRIP-Br1_2	----	D	D	G	L	E	G	L	F	E	D	I	D	-	////	////	+	+	+/-	++	++	////	////
Zyxin	----	I	D	S	L	S	S	L	L	D	D	M	T	-	////	+/-	+	+/-	+/-	++	++	////	////
DNJA1	----	L	S	L	L	E	K	L	L	P	E	R	K	-	////	+/-	+	++	+/-	++	++	////	////
TRIP-Br1_1	----	S	A	S	M	A	S	L	L	E	D	L	S	-	////	+	+/-	+/-	+/-	++	++	////	////
BRCA1_3	----	E	R	T	L	K	Y	F	L	G	I	A	G	-	////	+/-	+	++	+/-	-	-	////	////
E6TP1	----	P	T	K	L	S	D	F	L	I	T	G	G	-	////	+	+	+	+/-	+++	+++	////	////
TSC2_1	----	L	Q	T	L	Q	D	I	L	G	D	P	G	-	////	++	+	+/-	-	++	++	////	////
TSC2_2	----	E	L	Q	T	L	Q	D	I	L	G	D	P	-	////	+	+	+/-	+/-	+	+	////	////
E6BP	----	G	F	V	S	L	E	E	F	L	G	D	Y	-	////	+	++	++	+	+++	+++	////	////
BRCA1_1	----	I	Q	K	V	N	E	W	F	S	R	S	D	-	////	-	+	+	+	+++	+++	////	////
BRCA1_2	----	F	S	R	S	D	E	L	L	G	S	D	D	-	////	+/-	+/-	+/-	+/-	+++	+++	////	////
MED13	----	Y	T	D	L	D	N	L	F	N	S	D	E	-	-	+/-	+	+/-	-	++	+++	////	////

Tableau récapitulatif des résultats des tests d'interaction entre les protéines E6 ( des types HPV16, 18, 3, 5, 1 et BPV1) et la liste de 47 peptides LxxLL (Série 2). Les peptides LxxLL sont rangés selon leur appariement de séquence tandis que l'ordre des protéines E6 est fait sur la base de leurs tropismes (muqueux, cutané ou mixte) et leurs genres ( $\alpha$ ,  $\beta$ ,  $\mu$  ou  $\delta$ ). La qualité d'interaction est présenté par signes : +++ (> 80%), ++ (80% - 50%), + (50% - 20%), +/- (20% - 10%) et - (<10%). Les cases correspondant de données manquantes sont rayées.

## Chapitre II. Analyse du système d'interaction de la protéine E6

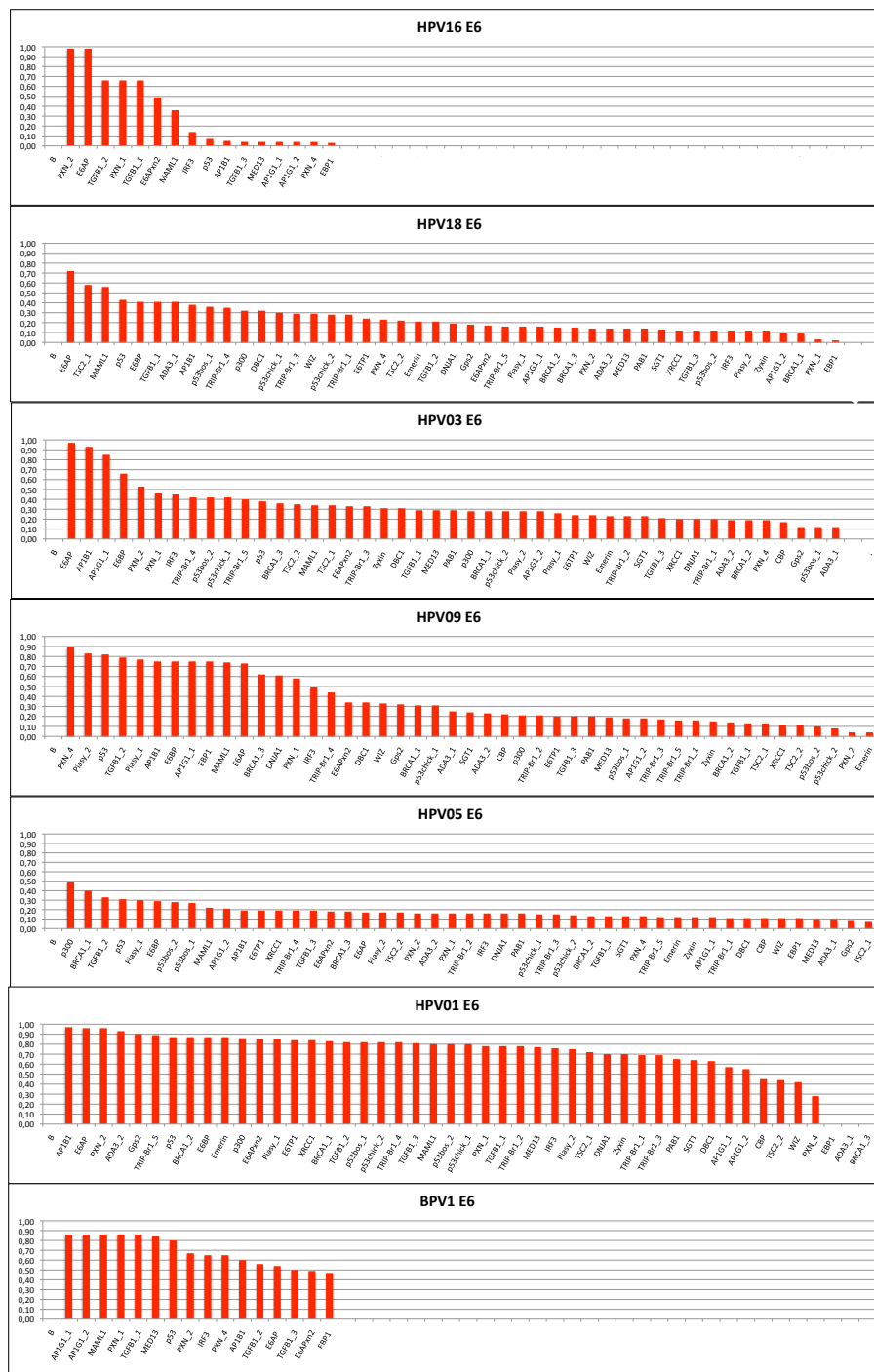


Figure II.6 – Résultats de "Holdup" - Série 2 (Histogrammes)

Histogrammes décroissantes qui présentent les résultats de la deuxième Série de testes d'interactions Holdup entre les protéines E6 ( des types HPV16, 18, 3, 5, 1 et BPV1) et la liste de 47 peptides LxxLL. Notons le manque de certains données à cause des problèmes logistiques.

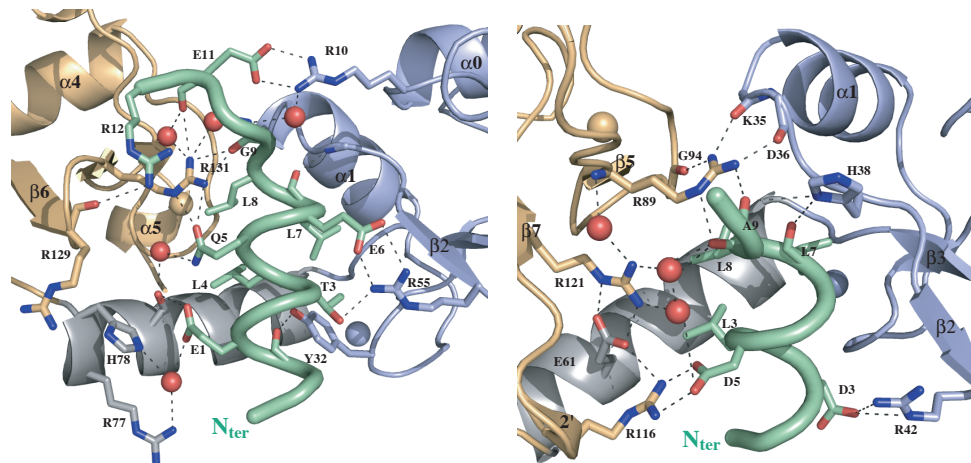
Dans la Série 2 seuls cinq peptides (E6AP, AP1B1, AP1G1-1, E6BP et PXN-2) arrivent à lier HPV03 E6 avec un pourcentage plus élevé à 50% contre 30 peptides dans la Série 1, sachant que le nombre des peptides testés est à peu près égal dans les deux Séries (environ 46 peptides). La protéine du type HPV09 constitue une exception à cette tendance puisqu'elle interagit plus fortement avec un plus grand nombre de peptides dans la Série 2 que dans la Série 1. Alors qu'elle liait, dans la première Série, seulement 4 peptides dans la Série 1 avec un pourcentage plus élevé que 50%, HPV09 E6 interagit avec 14 peptides la Série 2, dont 11 avec un pourcentage plus élevé que 70%. Donc, si HPV09 E6 possède bien un site de liaison aux motifs LxxLL, elle semble mal adaptée à la séquence LxxLL de E6AP et à ses dérivés. La protéine du type HPV01 E6 possède un profil remarquable, en ce sens qu'elle lie fortement une grande partie des peptides de la Série 2. Ceci suggère que HPV01 est une protéine qui lie fortement les motifs LxxLL et leurs dérivés en général, d'une manière peu dépendante des variations fines de séquence. Cette notion est renforcée par le fait que HPV01 E6 liait aussi très fortement la plupart des peptides de la Série 1, en semblant être peu influencée par les mutations ponctuelles, en particulier par les changements de résidus hydrophobes ou de charges sur des positions clés. Donc, HPV01 E6 apparaît comme une protéine qui reconnaît avec une forte affinité, mais aussi avec une faible spécificité, des motifs  $\Phi_{xx}\Phi$  où  $\Phi$  est un résidu hydrophobe.

## 2.2 Etude d'interface d'interaction E6 - LxxLL

### 2.2.1 Complexes HPV16 E6/E6APp et BPV1 E6 - Pax.p (Structures cristallographiques)

Les données extraites des deux structures cristallographiques des complexes HPV16 E6/E6AP (Chapitre I) et BPV1 E6/Pax-p (annexe 1) nous ont fourni suffisamment de données pour déjà esquisser les ébauches de la fresque du système d'interaction de la protéine E6 avec les protéines cellulaires à motif linéaire LxxLL. Comme nous l'avons déjà expliqué brièvement au chapitre premier, les trois Leucines ( $L_4, L_7$  et  $L_8$ ) du motif LxxLL sont ancrées dans la poche de E6 sur des résidus hydrophobes appartenant principalement au domaine E6N et à la région inter-domaines. Les périphéries de la poche hydrophobe, en particulier au niveau du E6C, contiennent des résidus basiques qui entrent en complexe avec les résidus négativement chargés contenus par les motifs LxxLL liant E6, tels que ceux de E6APp et Pax-p. En sus de ces contacts hydrophobes et électrostatiques, un réseau de contacts médiés par des molécules d'eau s'observe entre les résidus basiques du domaine E6C et les résidus acides du peptide LxxLL (Figure II.7).

La nécessité ou l'importance de certains de ces résidus dans l'interaction E6-LxxLL a été



**Figure II.7 – Interfaces d'interactions HPV16 E6-E6AP et BPV1 E6-Paxilline**

Représentation structurale qui montre les divers interactions formées entre les résidus de la protéine E6 et ceux du peptide LxxLL dans deux contextes différents; HPV16 E6-E6AP (à gauche) et BPV1 E6-Pax (à droite). ce sont essentiellement des liaisons de charge, liaisons polaires et interactions médiées par les molécules d'eau.

confirmée par la mutagenèse couplée aux tests d'interaction avec le peptide E6AP ou par des tests d'activité de dégradation de p53 (voir annexe 2, Figure 2).

### 2.2.2 Concernant les autres types de E6 (Modélisation par homologie)

Pour étudier l'interface d'interaction E6-LxxLL d'une façon plus large et pour pouvoir interpréter du point de vue structural la banque de données générée par les tests Hold-up, nous avons d'abord réalisé une modélisation par homologie pour prédire des modèles des structures des autres types de E6 non encore résolues, utilisées dans notre étude interactomique (HPV1,3,5,9,18) en se basant sur les deux structures déjà résolues de E6 (HPV16 E6 et BPV1 E6) et les alignements de séquences comme support de modélisation<sup>2</sup>. Nous pensons que ces modèles sont relativement fiables. En effet, les deux structures que nous avons résolues sont elles-mêmes très similaires (rmsd 2Å sur les régions alignées) alors qu'elles sont issues des deux Papillomavirus (HPV16 et BPV01) les plus distants phylogénétiquement de toute la Série considérée. En particulier, dans les deux structures la conformation des chaînes principales de E6 et du peptide est très conservée au niveau de la zone d'interaction E6-LxxLL (Figure II.7), ce qui nous induit à penser qu'elle va être également conservée dans les autres cinq types de

2. La modélisation par homologie a été effectuée par le logiciel MODELLER 9.10

HPV considérés. Par contre, et de façon intéressante, du point de vue des séquences primaires, cette zone connaît des variations importantes présentées ci-après.

### 2.2.3 Détermination de la zone d'interaction

L'inventaire des résidus de E6 qui sont en contact atomique avec les résidus des peptides LxxLL dans les deux structures cristallographiques a été fait en calculant l'exposition au solvant des résidus en présence et en absence du peptide lié. En se basant sur les résultats de ces calculs, on peut réduire la zone d'interaction à 24 résidus de E6 impliqués dans des contacts directs avec le peptide et accessibles spatialement à ce dernier. Il s'agit de 14 résidus de E6N (R<sub>10</sub>K<sub>11</sub>V<sub>31</sub>Y<sub>32</sub>D<sub>49</sub>L<sub>50</sub>C<sub>51</sub>V<sub>53</sub>R<sub>55</sub>V<sub>62</sub>L<sub>67</sub>Y<sub>70</sub>S<sub>71</sub> et I<sub>73</sub> dans HPV16) et de 10 résidus de E6C (S<sub>74</sub>R<sub>77</sub>H<sub>78</sub>L<sub>100</sub>R<sub>102</sub>Q<sub>107</sub>R<sub>129</sub>G<sub>130</sub>R<sub>131</sub> et T<sub>133</sub> dans HPV16). Notons que dans cette nomenclature nous attribuons à E6N et à E6C les trois premiers tours et le dernier tour de l'hélice de connection, respectivement. Nous nous basons en cela sur les découpages que nous avons fait pour produire les deux domaines séparément sous forme soluble (Zanier *et al.* [2012]). L'alignement de ces 24 résidus pour les différents types de E6 montre une conservation globale des régions qui forment les piliers du système de reconnaissance E6-LxxLL observé dans les deux structures cristallographiques (région hydrophobe, région basique) (Figure II.8). Au sein de cette conservation globale apparente, des variations de séquences sont cependant observées dans cette région spatiale de E6 en terme de nombre et distribution des résidus hydrophobes, chargés et polaires, etc... Ce sont ces variations qui impliquent les variations fines de préférences de chaque type de E6 par rapport au motif consensus général LxxLL (Tableau II.3, Figure II.9 et II.10).

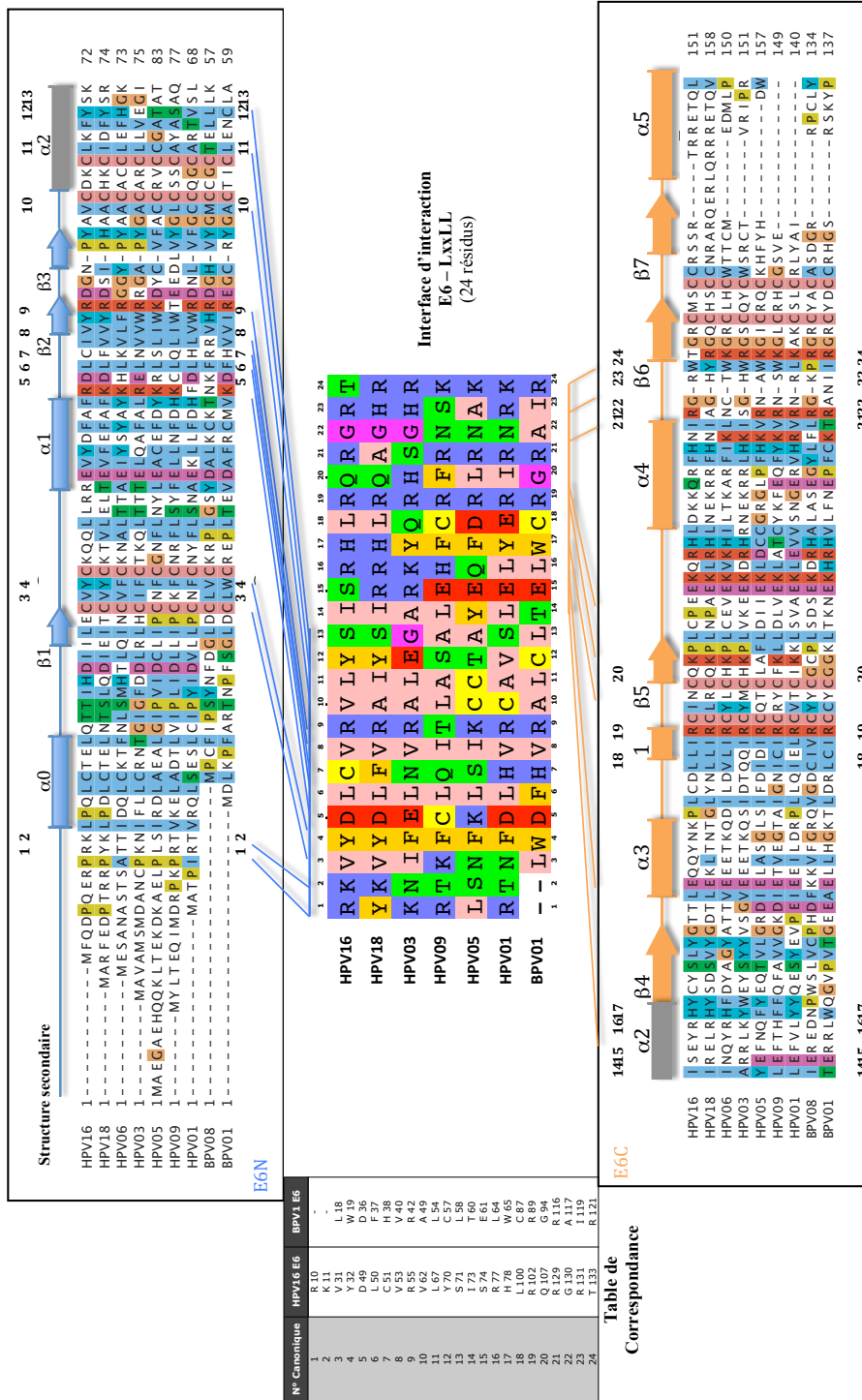


Figure II.8 – Alignement des séquences d'interfaces d'interactions de divers types E6

Alignement des séquences des deux domaines de E6 ; E6N (bleu, en haut) et E6C (orange, en bas), appartenant aux 7 HPV (16, 18, 6, 3, 5, 9 et 1) et 2 BPV (8 et 1), avec la structure secondaire de HPV16 E6. Au milieu : Alignement des 24 résidus impliqués dans l'interaction E6-LxxLL (à droite), des types de protéines E6 utilisées dans cette études (HPV16, 18, 3, 9, 5, 1 et BPV1) et la table de correspondance entre l'ordre des séquence canonique et celui de HPV16 E6 et BPV1 E6.

Tableau II.3 – Variations de séquences de la zone d'interaction E6-LxxLL

Domaines	Résidus négatifs		Résidus positifs		Résidus hydrophobes		Résidus polaires	
	E6N	E6C	E6N	E6C	E6N	E6C	E6N	E6C
HPV 16	1	0	3	5	8	3	1	3
HPV 18	1	0	2	6	9	4	1	1
HPV 03	2	0	2	6	7	2	2	2
HPV 09	0	1	2	4	7	4	4	2
HPV 05	0	2	2	3	7	4	4	2
HPV 01	1	2	3	4	6	4	3	1
BPV 01	1	1	2	3	8	6	0	1

Variations du nombre et de la distribution des résidus chargés (négativement [D e E] et positivement [K, H et R]), hydrophobes (I, V, L, F, C, M, A, G et W, selon : [Kyte et Doolittle \[1982\]](#)) et polaires (T, S, N et Q) au niveau de la région d'interaction E6-LxxLL selon l'alignement des séquences des protéines E6 issues de divers types viraux (HPV16, 18, 3, 9, 5, 1 et BPV1 E6).



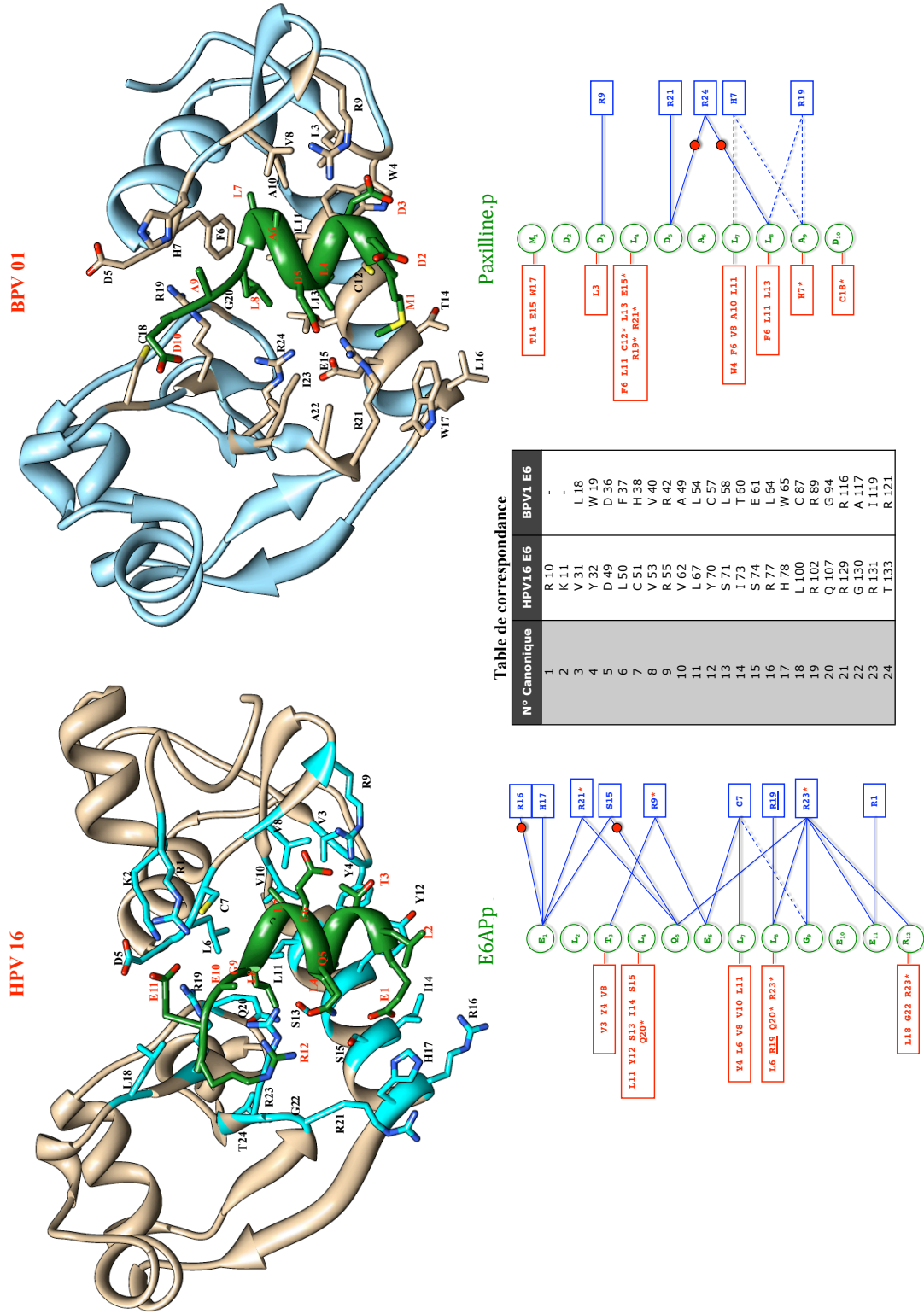


Figure II.9 – Réseaux de contacts E6-peptide (HPV16 et BPV1)

En haut : Structures des complexes ; HPV16 E6-E6AP (à gauche) et BPV1 E6-Pax (à droite). Les résidus impliqués en liaison avec le peptide LxxLL sont indiqués par numéros canoniques (1 à 24). En haut : Schéma représentatif des interactions E6-LxxLL (HPV16 [à gauche] et BPV1 [à droite]) : contacts de chaîne principale (pointillés bleus), polaires (lignes bleues), hydrophobes (lignes rouges) ou médiés par l'eau (boules rouges).

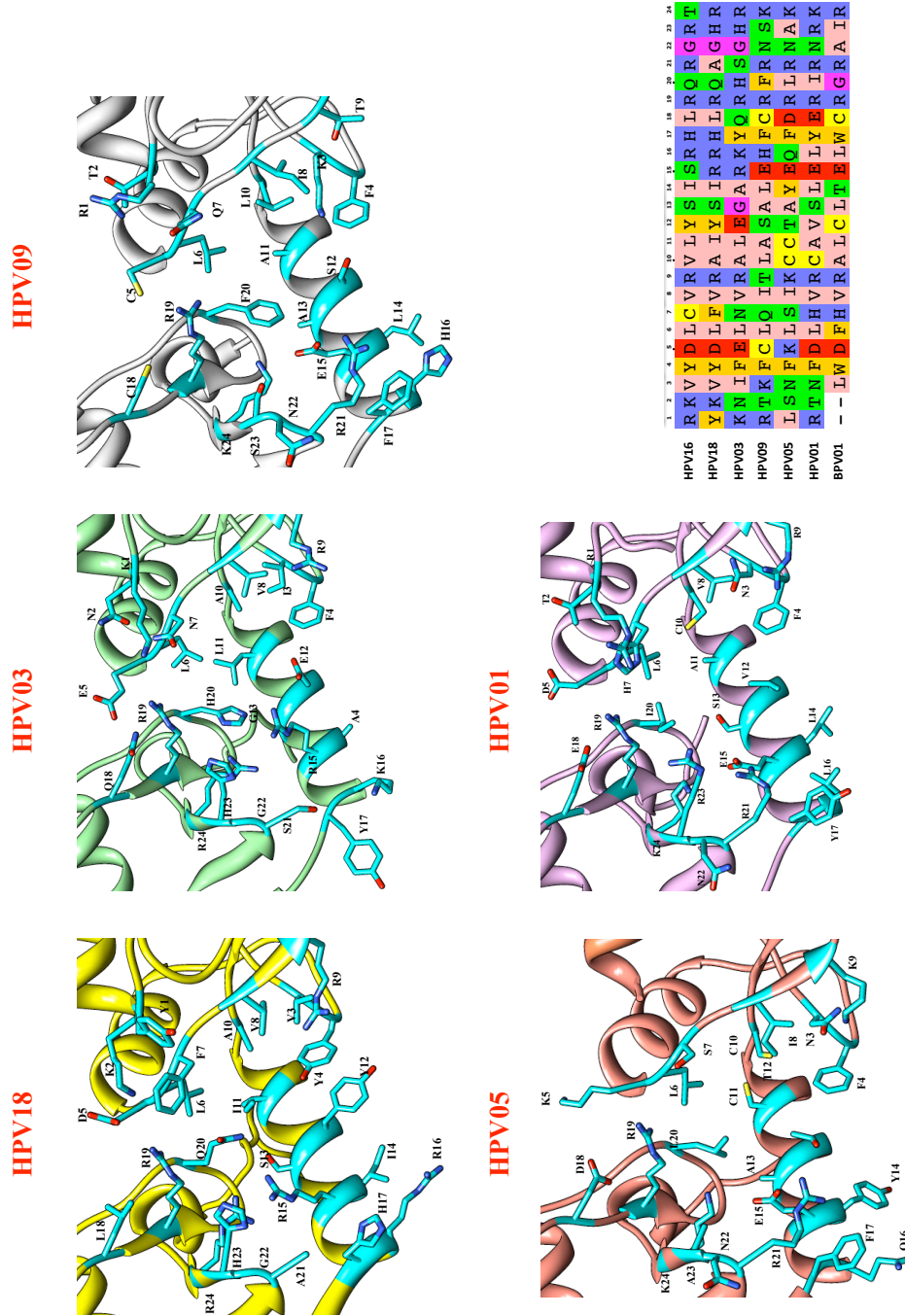


Figure II.10 – Interface d'interaction des modèles HPV1.3,5,9 et 18

Zone d'interaction E6-LxxLL prédites des protéines E6 des types viraux HPV18, 3, 9, 5 et 1 (modélisation par homologie avec HPV16 E6 et BPV1 E6) avec l'alignement des séquences de ces 5 types de E6 ainsi que celles de HPV16 E6 et BPV1 E6.

### 3 Discussion

#### 3.1 Modèle de mécanisme d'interaction E6-LxxLL

Malgré la conservation globale apparente de la zone d'interaction présentée si dessus, une observation plus focalisée montre une haute diversité de cette interface confirmée par les résultats des deux Séries de tests Hold-up qui révèlent clairement des champs de sélectivité propres à chaque type. Nous suggérons que cette diversité est dûe aux régions non conservées dans les séquences des 24 résidus qui forment l'interface de contact avec les motifs LxxLL, chez les différents types de E6. Comme nous l'avons déjà évoqué, cette diversité n'agit probablement pas sur la conformation de la chaîne principale, mais essentiellement au niveau de la nature des chaînes latérales de ces 24 résidus. Les informations structurales et interactomiques nous incitent à suggérer ici un modèle qui essaye d'expliquer les mécanismes mutationnels qui permettent à E6 de s'adapter à diverses variations des motifs LxxLL des protéines cellulaires. Selon ce modèle on peut classer les 24 résidus qui forme la zone de liaison en 3 catégories :

##### **- les résidus attractifs :**

Cette catégorie comporte les résidus hydrophobes (V31Y32L50C51V53V62L67Y70I73 ... etc.), qui ne jouent qu'une fonction attractive générale des motif hydrophobes (LxxLL) sans ciblage précis des résidus qui entourent les 3 Leucines. Ceci ressemble à un "Fusil de chasse" qui vise globalement l'objectif sans précision. Le nombre de ces résidus hydrophobes entre 13 et 8 dont la majorité se trouvent au niveau du domaine E6N.

##### **- Les résidus sélectifs :**

Contrairement à la catégorie précédente, cette catégorie comporte les résidus chargés positivement (majoritairement les Arginines), localisés principalement au niveau du côté C-terminal de l'interface. Selon le modèle que nous suggérons, ces résidus sont responsables d'une grande partie des contacts protéine-peptides déterminants de la sélectivité via leurs contacts directes avec les résidus peptidiques acides ou polaires (comme Fusil de précision), formant ainsi des "agrafes" à Arginines comme par exemples les "agrafes" R10-E11, R55-E6 qui se forment entre HPV16 E6 et le peptide E6AP ou l'agrafe R116-D5 et R42-D3 entre BPV1 E6 et le peptide Paxilline (Figure II.7). Le raisonnement basé sur l'effet sélectif de ces résidus est renforcé par le fait que certains contacts de ces Arginines peuvent varier de positionnement relativement aux résidus peptidiques. Par exemple le résidu HPV16 R129 interagit avec le résidu en position 1(E1), contrairement à son homologue BPV1 R116 qui lie le résidu en position 5 (D5). Ce décalage montre la capacité de E6 à s'adapter à des variations du peptide concernant les résidus polaires et acides. Il faut toutefois également noter que les résidus arginines apportent,

en plus de leur charge, une contribution hydrophobe à l'interaction, qui peut être importante. Le nombre de ces résidus arginine varie entre 8 et 5 (R10R55R77R129R131 ... etc).

#### - Les résidus guides :

Ces résidus acides ou polaires servent à recruter et orienter les chaînes latérales des résidus sélectifs (lunette de visée), vers des coordonnées spatiales précises calquant ainsi la poche de reconnaissance sous formes des moules compatibles avec les motifs cellulaires LxxLL variés. Par exemple, selon ce modèle, le résidu acide ou polaire en position 15 (E61[BPV1], S74 [HPV16]) oriente en direction du peptide les résidus en position 21 (R116 [BPV1], R129 [HPV16]). Ce raisonnement est cohérent avec les non-conservations simultanées du résidu sélectif en position 21 et de son résidu guide en position 15 chez les deux protéines HPV03 et HPV18, pourtant phylogénétiquement proches (genres  $\alpha 2$  et  $\alpha 7$ ).

### 3.2 Spécificité et affinité

Hormis des observations générales sur la spécificité et l'affinité globale des types E6, cette étude n'a probablement pas couvert assez de séquences peptidiques pour permettre l'attribution à chaque variant de E6 d'une sous-famille précise de motifs LxxLL. Les peptides testés dans la première Série n'étaient sans doute pas assez diversifiés, car tous déduits d'un même peptide, ce qui permettait d'identifier des résidus délétères aux interactions, mais uniquement dans le cadre de la séquence de départ (E6AP). Les peptides de la deuxième Série étaient au contraire trop différents les uns des autres, ce qui ne nous permettait pas d'attribuer aisément les gains ou les pertes d'affinité dus à une position et/ou un résidu particuliers. L'interaction de tous les variants hormis HPV05 E6 avec la version sauvage du peptide E6AP reflète l'importance de cette dernière comme clé des mécanismes pathologiques des Papillomavirus. Le très faible taux d'interaction de la protéine HPV05 avec tous les peptides dans les deux Séries de tests, est peut être lié (en sus des problèmes de qualité déjà évoqués) à la particularité de ce type de E6 d'avoir un résidu positivement chargé à la position 5 qui peut créer une répulsion avec l'Arginine en position 19 conservée chez tous les types E6. Le résultat de cette répulsion R19-K05 est peut être drastique sur la forme de la poche d'interaction qui peut devenir spécifique pour des motifs autres que LxxLL. Quand à la tolérance large de E6 du type HPV01 dans les deux Séries de tests Hold-up, nous suggérons qu'elle est liée à la bonne distribution de résidus sélectifs sur la poche d'interaction, qui donnerait plus de chance de lier fortement le peptide.

### 3.3 Différences de comportement entre le Peptide et la Protéine

Les résultats de cette étude présentent quelques éléments d'incohérence apparente avec la plupart des études publiées dans ce domaine qui montrent une interaction exclusive pour les protéines E6 issues des Papillomavirus Muqueux à Haut Risque (HRM) avec des protéines cellulaires comme Paxilline, E6TP1, ADA3, IRF3 et MAML1 (voir revue [Klingelutz et Roman \[2012\]](#)). Or, nous démontrons ici que des peptides extraits de ces protéines interagissent non seulement avec des E6 de PV muqueux à haut risque mais aussi avec des E6 issues de PV muqueux à bas risque, voire de PV cutanés à haut ou bas risque. Tout d'abord, le fait que la plus part de ces études ont été réalisées sur des protéines entières et que la nôtre est effectuée sur des peptides linéaires courts, peut déjà expliquer au moins en partie cette incohérence, puisque l'incompatibilité de ces protéines avec E6 à BRM est peut être dûe à un entourage non permissif du point de vue structural (repliement) ou séquentiel (nature des résidus) du motif LxxLL au sein de la protéine. De plus, le fait que certaines de ces études ont été faites in vivo, ce qui laisse la possibilité d'existence d'une interaction via une troisième protéine reconnue par les E6 à HRM et non par les BRM, d'une façon similaire à la triple interaction E6-E6AP-p53.

### 3.4 Mode compétitif ou enzymatique

La reconnaissance des motifs linéaires courts comme celui de LxxLL par les protéines modulaires constitue un système de reconnaissance très utilisé dans les mécanismes de la régulation et signalisation cellulaire ([Diella et al. \[2008\]](#)). Il est bien démontré maintenant que ce système est utilisé comme une porte d'attaque fragile par les virus ([Davey et al. \[2011\]](#)). Dans ce contexte, il paraît que la protéine E6 a développé sa poche de reconnaissance pour cibler spécialement une famille de motifs hydrophobes et chargés négativement existants sur les protéines cellulaires associées à des fonctions sensibles pour la cellule (E6AP, BAK, Paxilline... etc.). Cependant la disparité d'affinités entre ces motifs et les protéines E6 appartenant aux types différents des Papillomavirus laisse penser à l'existence de deux types de stratégies utilisées pour s'emparer des protéines cellulaires :

#### **- mode compétitif :**

C'est le cas de la protéine BPV1 E6 qui lie les motifs LxxLL avec une constante de dissociation ( $K_d$ ) assez petite de l'ordre des micromolaires ou des nanomolaires. Par exemple BPV1 E6 peut lier le motif LxxLL de la protéine Paxilline avec un  $K_d$  de 35,1 nM ([Sidi et al. \[2011\]](#), [Zanier et al. \[2005b\]](#)). Ce qui constitue une affinité très forte en comparaison avec les interactions des protéines cellulaires avec ces motifs qui ont un  $K_d$  de l'ordre des millimolaires ou submillimolaires ([Hoellerer et al. \[2003\]](#), [Lorenz et al. \[2008\]](#), [Savkur et Burris \[2004\]](#), [Hur](#)

*et al.* [2004]).

### **- mode enzymatique :**

Contrairement à BPV1 E6, la protéine HPV16 E6 lie les motifs LxxLL avec un  $K_d$  d'environ 4 $\mu$ M, bien moins forte que celui du complexe BPV1 E6/LxxLL (Sidi *et al.* [2011]). Par contre, HPV16 E6 est connue par sa capacité à induire la dégradation d'une partie importante de ses cibles cellulaires par voie de protéasome. L'efficacité de ce mécanisme enzymatique propre à HPV16 E6 compense l'affinité faible par rapport à BPV1 E6 et explique aussi peut être, le taux faible de cette protéine dans les cellules infectées par le virus HPV16 (Nomine *et al.* [2006]). Ce modèle basé sur la disparité d'affinité pour LxxLL entre HPV16 E6 et BPV1 E6 pourrait toutefois être remis en cause par le fait que les mesures de  $K_d$  du complexe HPV16 E6/LxxLL ont été faites sur un mutant de E6 (E6 F47R 4C4S). Par exemple, nous savons l'Arginine 129 du HPV16 E6 est importante pour l'interaction, et nous avons discuté dans le chapitre 1 le fait que le positionnement de cette arginine pourrait être influencé par la mutation C80S. On peut penser que l'affinité réelle du complexe du HPV16 E6 sauvage avec le peptide LxxLL pourrait être, sinon aussi forte que celle observée pour le complexe BPV01 E6 - paxilline, du moins plus forte que celle mesurée en utilisant nos mutants solubilisés de HPV16 E6. Mais, d'un autre côté, le nombre de peptides super-ligands (>80%) très élevé de BPV01 E6 par rapport à celui de HPV16 E6 observé dans les résultats des interactions Hold-up discutés plus haut, concorde bien avec ce modèle proposant deux mécanismes d'action distincts des protéines E6.



---

---

## Chapitre III

---

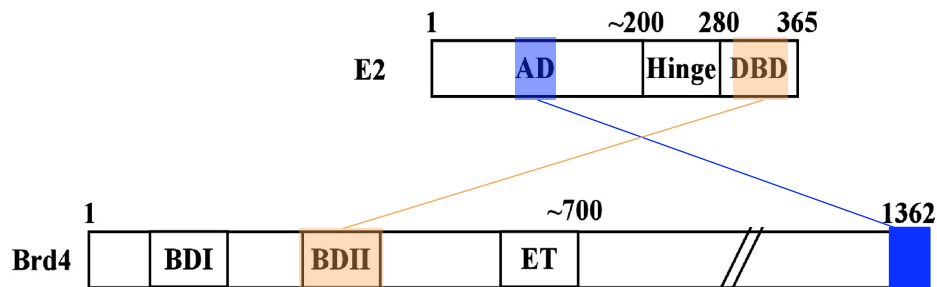
# Mise en évidence d'une nouvelle interaction entre la protéine virale E2 et la protéine cellulaire Brd4

### 1 Problématique

Cette partie des résultats, bien qu'elle ne concerne pas l'étude des oncoprotéines E6, constitue cependant une application des procédés méthodologiques développés durant ce projet de recherche. En effet, nous avons profité des compétences acquises grâce aux travaux effectués sur la protéine E6, pour débiter un nouveau projet en collaboration avec le Pr Cheng-Ming Chiang (University of Texas Southwestern Medical School), avec pour but de comprendre la régulation de la protéine virale E2, médiée par la protéine cellulaire Brd4. Nous avons vu dans l'introduction que la protéine E2 constitue un facteur de transcription clé dans la régulation du cycle viral des HPV qui régule la transcription, la réplication et la ségrégation de l'ADN viral. Nous avons vu également qu'elle est composée de deux domaines fonctionnels et structuraux ; le domaine N-terminal (200 résidus) qui représente le domaine de transactivation (TAD ou AD) et le domaine d'interaction avec l'ADN (DBD) (90 résidus) ([Androphy \*et al.\* \[1987b\]](#), [Kurg \*et al.\* \[2006\]](#)).

Il a été démontré que la protéine virale E2 assure la maintenance du génome viral en interagissant avec le fuseau mitotique, permettant ainsi la co-multiplication du génome du virus avec celui de l'hôte, la ségrégation équitable du génome viral et le maintien des formes épisomales de l'ADN viral. La co-localisation de E2 avec le fuseau mitotique est garantie par la protéine cellulaire à bromodomaine Brd4, capable d'interagir avec les deux partenaires (E2 et fuseau mitotique). Il a été démontré, par des études fonctionnelles et structurales, que la liaison E2-Brd4 se fait précisément entre le domaine d'activation N-terminal (AD) de E2 et les





**Figure III.1 – Surface électrique et hydrophobe de E6**

Dans cette figure : Figure descriptive de l'interaction déjà connue entre le domaine d'activation N-terminal (AD) de E2 et les 20 derniers résidus du domaine C-terminal de Brd4 (bleu), et celle que nous présentons dans ce travail, qui se fait entre le domaine d'interaction avec l'ADN (DBD) de la protéine HPV16 E2 et la région bromodomaine II (BDII) de la protéine Brd4 (orange) (Figure adaptée de [Abbate \*et al.\* \[2006\]](#)).

20 derniers résidus du domaine C-terminal de Brd4 ([Abbate \*et al.\* \[2006\]](#); [You \*et al.\* \[2004\]](#)).

Nous démontrons ici l'existence d'une autre liaison entre les deux protéines E2 et Brd4, médiée par le domaine d'interaction avec l'ADN (DBD) de la protéine E2 issue du virus HPV16 et la région bromodomaine II (BDII) de la protéine Brd4 (Figure III.1). Cette interaction est conditionnée par la phosphorylation de certains résidus de Brd4. Nos données structurales ont été effectuées sur la base de travaux fonctionnels préliminaires qui avaient été effectués par le Pr Cheng-Ming Chiang.

## 2 Résultats

### 2.1 Interface de reconnaissance

Pour mettre en évidence cette interaction, nous avons cloné, exprimé et purifié le domaine C-terminal (DBD) de la protéine HPV16 E2 (résidus : 285-365), uniformément marqué à l'azote N15 pour pouvoir effectuer des analyses en RMN. Le profil de purification montre que cette protéine adopte bien une forme dimérique. En parallèle, nous avons fait synthétiser le peptide synthétique sauvage (Brd4wt) issu de la région bromodomaine II (BDII) de la protéine Brd4, ainsi que son variant (3xPBrd4) phosphorylé au niveau des Sérines en position 5, 10 et 14. Nous indiquons ci-après les séquences des peptides. Les sérines phosphorylées sont indiquées par la lettre "s" rouge et minuscule :

- Brd4 wt (séquence : S<sub>1</sub>D<sub>2</sub>S<sub>3</sub>S<sub>4</sub>S<sub>5</sub>D<sub>6</sub>S<sub>7</sub>D<sub>8</sub>S<sub>9</sub>S<sub>10</sub>T<sub>11</sub>D<sub>12</sub>D<sub>13</sub>S<sub>14</sub>E<sub>15</sub>E<sub>16</sub>E<sub>17</sub>R<sub>18</sub>) ;

- 3xPBrd4 (séquence : S<sub>1</sub>D<sub>2</sub>S<sub>3</sub>S<sub>4</sub>S<sub>5</sub>D<sub>6</sub>S<sub>7</sub>D<sub>8</sub>S<sub>9</sub>S<sub>10</sub>T<sub>11</sub>D<sub>12</sub>D<sub>13</sub>S<sub>14</sub>E<sub>15</sub>E<sub>16</sub>E<sub>17</sub>R<sub>18</sub>) ;

Nous avons ensuite effectué une analyse par RMN des titrations entre la protéine E2 (DBD) et chacun des peptides mentionnés ci-dessus, avec une stoechiométrie protéine/peptide croissante (1/0, 1/0.25, 1/0.5, 1/0.75, 1/1 et 1/5). Les données RMN ont été collectées et traitées par Katia Zanier, spectroscopiste RMN dans notre équipe. Le spectre HSQC (Heteronuclear Simple Quantum Correlation) à deux dimensions ( $H^1N^{15}$ ) de la protéine E2 montre des variations de déplacements chimiques induites par la présence du peptide triplement phosphorylé (3xPBrd4) (Figure III.2). Les variations affectent des pics correspondant à plusieurs résidus. Les variations les plus importantes sont observées au niveau des résidus K<sub>15</sub>, R<sub>18</sub>, K<sub>22</sub>, S<sub>33</sub>, W<sub>35</sub>, H<sub>36</sub> et F<sub>78</sub> (Figure III.2 et III.3). Ces variations sont proportionnelles à la concentration du peptide (Figure III.4). Contrairement au peptide phosphorylé, le peptide sauvage (Brd4wt) n'induit aucun changement au niveau des pics correspondant aux résidus de E2 sur le spectre HSQC. L'attribution des résidus du domaine DBD E2 HPV16 sur les spectres HSQC, ont été faites à l'aide de celle du celui de HPV06 publiée par Brown et ses collègues (Brown *et al.* [2011]).

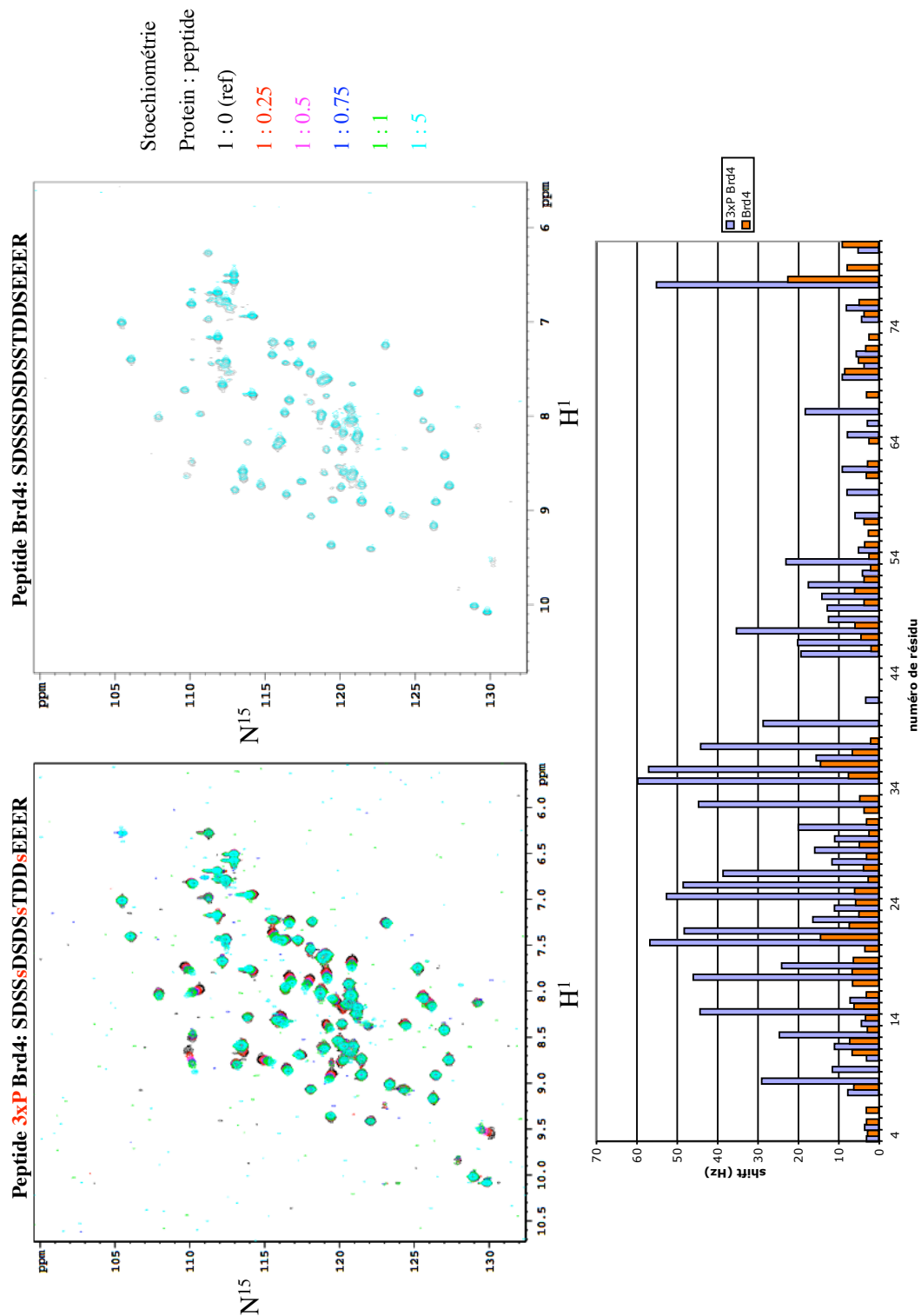


Figure III.2 – Analyse biophysique de l'interaction Brd4(BDII)-E2 (DBD)

En haut : Les spectre HSQC à deux dimensions ( $^1\text{H}^1\text{N}^{15}$ ) de la protéine E2 montre des variations de déplacements chimiques induites par la présence du peptide triplement phosphorylé (3xPBrd4) (à gauche), alors que le peptide sauvage (à droite) n'a pas d'effet sur le spectre de E2. (En bas : Les variations de déplacements chimiques représentées sous forme d'histogrammes.

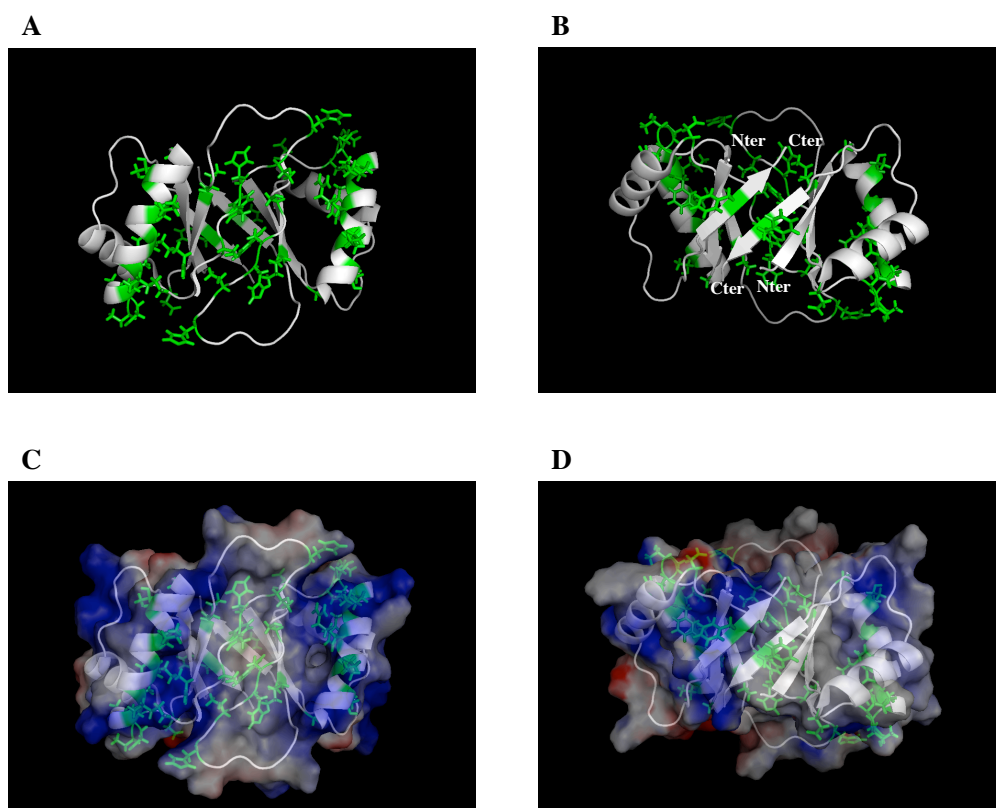
## 2.2 Rôle de la phosphorylation et de la salinité du solvant

Nous avons effectué les mêmes tests d'interactions par titrations en RMN en utilisant des peptides mono-phosphorylés pour analyser séparément la contribution de chaque sérine phosphorylée dans cette interaction :

- P1 Brd4 (séquence : S<sub>1</sub>D<sub>2</sub>S<sub>3</sub>S<sub>4</sub>S<sub>5</sub>D<sub>6</sub>S<sub>7</sub>D<sub>8</sub>S<sub>9</sub>S<sub>10</sub>T<sub>11</sub>D<sub>12</sub>D<sub>13</sub>S<sub>14</sub>E<sub>15</sub>E<sub>16</sub>E<sub>17</sub>R<sub>18</sub>) ;
- P2 Brd4 (séquence : S<sub>1</sub>D<sub>2</sub>S<sub>3</sub>S<sub>4</sub>S<sub>5</sub>D<sub>6</sub>S<sub>7</sub>D<sub>8</sub>S<sub>9</sub>S<sub>10</sub>T<sub>11</sub>D<sub>12</sub>D<sub>13</sub>S<sub>14</sub>E<sub>15</sub>E<sub>16</sub>E<sub>17</sub>R<sub>18</sub>) ;
- P3 Brd4 (séquence : S<sub>1</sub>D<sub>2</sub>S<sub>3</sub>S<sub>4</sub>S<sub>5</sub>D<sub>6</sub>S<sub>7</sub>D<sub>8</sub>S<sub>9</sub>S<sub>10</sub>T<sub>11</sub>D<sub>12</sub>D<sub>13</sub>S<sub>14</sub>E<sub>15</sub>E<sub>16</sub>E<sub>17</sub>R<sub>18</sub>).

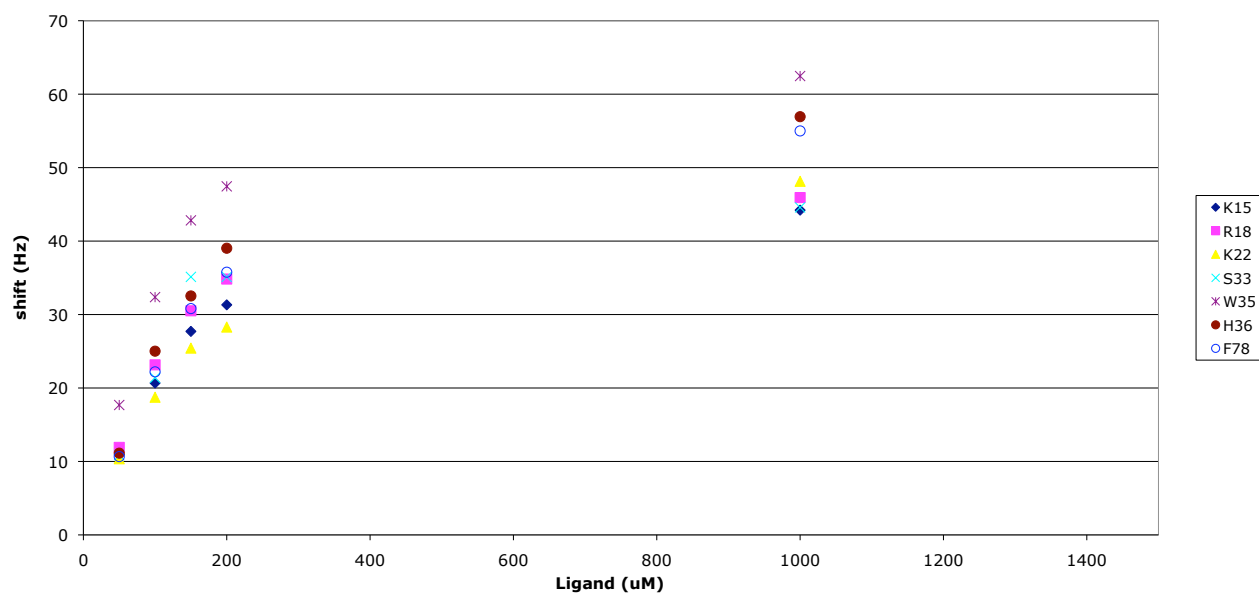
Ces analyses montrent que la contribution de chacune des sérines phosphorylées en position 5, 10 et 14 dans l'interaction Brd4(BDII)-E2 (DBD), est à peu près équivalente (Figure III.5).

D'autre part, pour estimer l'effet de la force ionique sur cette interaction, nous avons répété les mêmes titrations protéine/peptide avec deux concentrations en NaCl différentes (200 mM et 400 mM). Nous avons remarqué que les variations de déplacement chimiques induits par les titrations diminuent globalement quand on augmente la teneur en NaCl dans le tampon (Figure III.6).



**Figure III.3 – Interface d'interaction Brd4(BDII)-E2 (DBD)**

Les chaînes latérales des résidus impliqués dans cette interaction (vert), représentés sur une structure en mode "cartoon" (deux vues A et B) et en mode surface (C et D).



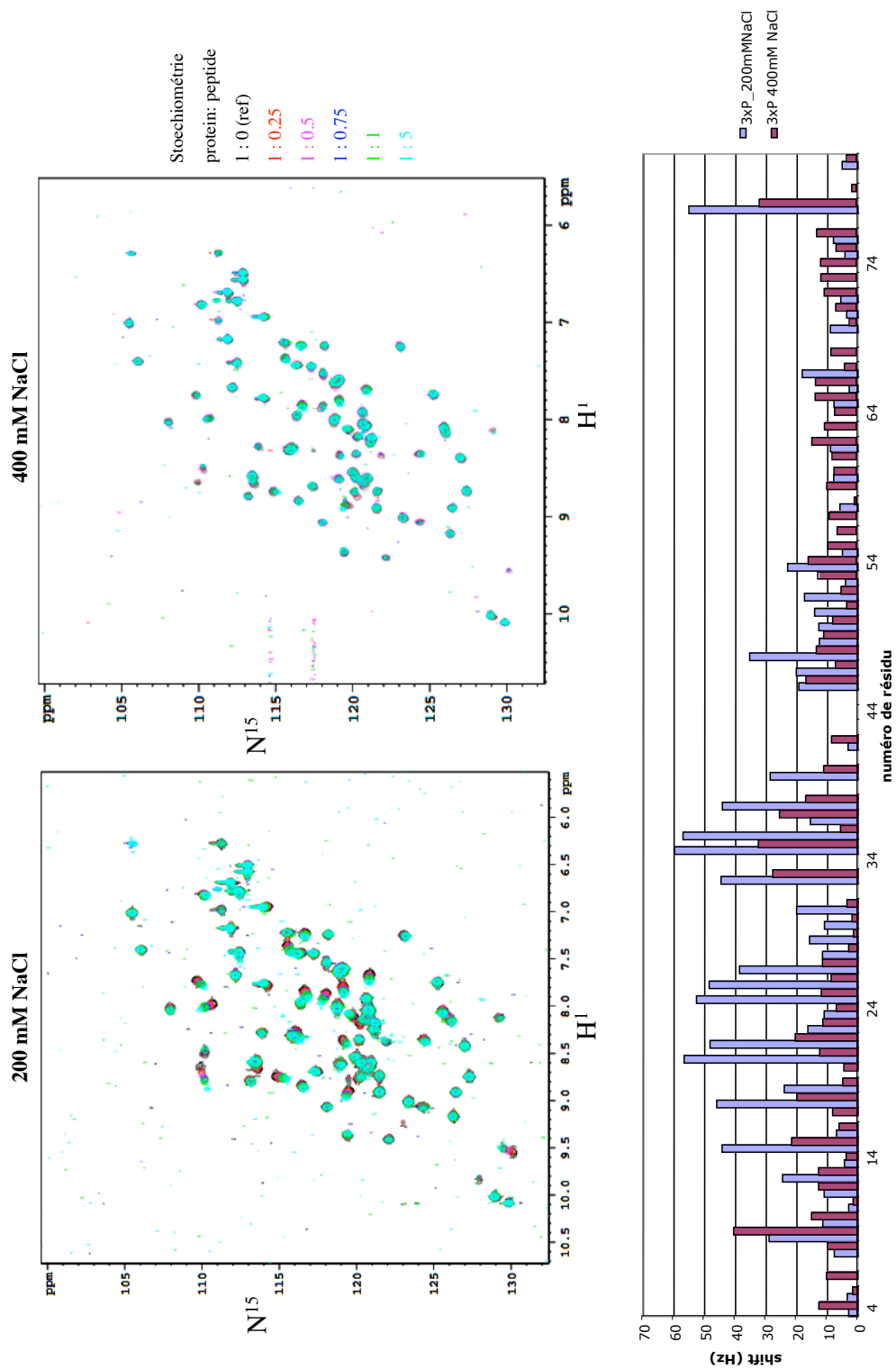
**Figure III.4 – Relation proportionnelle entre la concentration du peptide et les déplacements chimiques des résidus**

Courbes montrant la relation proportionnelle entre la concentration du peptide 3xPBrd4 et les déplacements chimiques des résidus (observées sur la carte HSQC) des résidus K<sub>15</sub>, R<sub>18</sub>, K<sub>22</sub>, S<sub>33</sub>, W<sub>35</sub>, H<sub>36</sub> et F<sub>78</sub>.

### Chapitre III. Mise en évidence d'une nouvelle interaction entre la protéine virale E2 et la protéine cellulaire Brd4



**Figure III.5** – Comparaison d'effets induits par les différents peptides phosphorylés  
Dans cette figure : Histogrammes montrant la contribution de chaque sérine phosphorylée dans l'interaction Brd4(BDII)-E2 (DBD).



**Figure III.6 – Robustesse de l'interaction en fonction de la force ionique**  
 En haut : Les spectre HSQC de la protéine E2 avec deux concentrations en NaCl différentes, soit 200 mM (à droite) et 400 mM (à gauche), en présence des mêmes concentrations de peptide 3xPBrd4. En bas : Histogramme comparatif des variations de déplacements chimiques induites par le peptide dans les deux concentrations de NaCl.



### 3 Discussion

Nous avons démontré dans ce chapitre par des analyses biophysiques (RMN), l'existence d'une nouvelle interaction liant un peptide phosphorylé issu de la région bromodomain II de la protéine cellulaire Brd4 et le domaine DBD de la protéine E2 du virus HPV16. Il a été publié que ce même domaine de E2 (DBD) reconnaît, sous sa forme dimérique, des séquences ADN double brin présentant le consensus ACCN(6)GGT ([Androphy \*et al.\* \[1987b\]](#)). Ceci laisse penser que le domaine BDII agit à l'état phosphorylé comme compétiteur de l'ADN sur le domaine DBD de E2, ce qui permettrait à E2 de se détacher de l'ADN après avoir complété sa fonction de recrutement de E1 sur l'origine de réplication. En effet, pour que la réplication du génome viral ait lieu, il faut que le dimère de E2 interagisse dans une première étape à la fois avec l'ADN et avec E1 permettant ainsi la fixation de cette dernière sur l'origine de réplication de l'ADN viral. Les étapes suivantes sont l'homo-dimérisation de E1 et l'hydrolyse de l'ATP. Le dimère de E2 se détache après l'hydrolyse de l'ATP, laissant E1 s'auto-associer sous sa forme fonctionnelle hexamérique avant de commencer son rôle catalyseur de séparation des deux brins ADN à partir de l'origine de réplication dans le sens 3'-5' ([Fouts \*et al.\* \[1999\]](#); [White \*et al.\* \[2001\]](#)). On peut suggérer donc que l'interaction de la région BDII phosphorylée de la protéine Brd4 avec le domaine DBD de E2, pourrait jouer un rôle dans la réplication du génome viral en permettant la rupture de la liaison entre le domaine DBD(E2) et l'ADN par le biais d'une interaction compétitive avec le même domaine (DBD). Cependant, le fait, mis en évidence par nos travaux, que l'interaction E2-phosphopeptides est fortement dépendante de la force ionique, pourrait mettre en cause sa signification biologique. D'un autre côté, la capacité contributive des 3 sérines phosphorylées dans cette interaction et le fait que le peptide non phosphorylé ne lie pas E2, ce sont des raisons seraient en faveur de l'existence d'un rôle biologique de cette liaison Brd4(BDII)-E2 (DBD). Un travail supplémentaire structural et biochimique sera nécessaire pour confirmer ou infirmer cette interaction.

---

# CONCLUSIONS GENERALES ET PERSPECTIVES

L'étude biochimique et structurale des oncoprotéines d'origine cellulaire ou virale, couplée à la modélisation moléculaire et au criblage à haut débit constitue une voie d'attaque essentielle dans l'arsenal disponible actuellement pour lutter contre le cancer. Cette approche a permis le développement des molécules anticancéreuses, dont certaines sont actuellement utilisées de manière systématique en thérapie anticancéreuse, telles que l'Imatinib (STI571 ou Glivec®) qui inhibe spécifiquement l'oncogène Bcr-Abl retrouvé dans la majorité des Leucémies Myéloïdes Chroniques, et agit également sur d'autres cancers du tube digestif ([Druker et Lydon \[2000\]](#)). L'utilisation de l'arsenic et l'acide rétinoïque qui ciblent et dégradent spécifiquement l'oncoprotéine de fusion PML/RAR impliquée dans la Leucémie aiguë promyélocytaire a donné aussi des résultats intéressants sur le plan de la thérapie des leucémies ([Zhu et al. \[1997\]](#)).

Nous nous sommes intéressés dans ce projet de recherche à réaliser des études biochimiques et structurales sur les oncoprotéines E6 produites par les *Papillomavirus* ainsi que leurs systèmes d'interactions avec les protéines cellulaires. Cette étude a pour but principal de contribuer à la recherche visant la lutte contre le cancer du col de l'utérus et autres cancers et pathologies dépendantes des *Papillomavirus*, où la protéine E6 joue un rôle crucial dans les processus pathogènes.

A l'issue de ce travail, des progrès considérables autant sur l'étude structurale de la protéine E6 que sur l'étude de son système d'interaction ont été réalisées. Nous avons également travaillé durant cette thèse sur l'interaction entre la protéine virale E2 et la protéine cellulaire Brd4, ce qui nous a permis l'identification d'une nouvelle interaction entre ces deux protéines.

**Stratégie efficace pour produire des protéines E6 de bonne qualité :** Comme dans toute étude structurale, nous avons commencé par optimiser le protocole général d'expression et de purification protéique et l'adapter aux difficultés de production rencontrées avec la protéine E6. Nous nous sommes particulièrement attardés sur cette étape en raison de la tendance de cette protéine à précipiter aux premières étapes de purifications.

Les méthodes de production des protéines recombinantes utilisées en routine dans notre laboratoire consistaient à exprimer E6 à une température relativement basse (15 C°), dans un système bactérien classique (*E.coli* BL21, DE3), en fusion avec la protéine MBP connue par sa capacité à stabiliser et solubiliser les protéines difficiles à purifier (Kapust et Waugh [1999]). On a ajouté ultérieurement, à ce protocole une étape d'ultracentrifugation à très haute vitesse et de longue durée pour sédimenter les fractions de MBP-E6 agrégées en forme d'oligomères, et pour récupérer les fractions monodisperses solubles. Ce protocole général a beaucoup amélioré la qualité des mutants de E6 et a même permis, la résolution des structures solubles des deux domaines séparés de E6 du type HPV16, cependant la qualité des échantillons purifiés de E6 entier restait assez limitée.

Nous avons alors apporté une modification importante à ce protocole, à savoir l'ajout d'un peptide LxxLL durant l'étape d'expression de E6, ce qui a amélioré considérablement la solubilité de E6, en raison, peut être de l'élimination de l'exposition au solvant de la grande poche de reconnaissance des motifs LxxLL, de nature hydrophobe, qui rentre en liaison avec le peptide LxxLL. Parallèlement, nous avons effectué une étude de mutagenèse comparative, visant à limiter l'oxydation de la protéine E6 riche naturellement en cystéines (14 cystéines pour HPV16 E6), par la mutation de certains résidus cystéines non conservés en résidus sérines. Cette modification du protocole a permis rapidement l'obtention de la structure du complexe E6-E6APp, comportant le mutant E6 F47R 4C4S  $\Delta$ 8CT du type HPV16 et le peptide LxxLL dérivé de E6AP.

De plus, cette étude de mutagenèse comparative a permis l'identification de certains mutants de E6 de qualité moyenne, qui ne portent que deux mutations (au lieu de 5 ou plus dans les mutants E6F47R4C4S et E6F47R 4C4SD144-151). Ces mutants pourraient s'avérer des candidats préférentiels pour des éventuelles études biochimiques supplémentaires, étant donné qu'ils sont plus proches de la forme sauvage de E6. En effet, le mutant E6F47R4C4S qui porte une mutation proche quoique non directement impliquée dans l'interface d'interaction E6-LxxLL (C80S), ne possède pas la fonction essentielle des protéines E6, à savoir la dégradation de la protéine pro-apoptotique p53 (Zanier *et al.* [2012]). Ceci a justifié notre recherche d'autres constructions portant moins de mutations, telles que E6F47RC111S et E6F47RC97S.

Cette stratégie de production protéique a bien prouvé son efficacité, puisqu'elle a mené

à la résolution des structures cristallographiques de deux protéines E6 issues de deux types viraux distincts (HPV16 et BPV1). Nous pensons donc que cette stratégie de solubilisation des protéines recombinantes par des peptides cibles pourrait être utilisable dans des cas similaires à la problématique de E6, où la protéine étudiée est instable sans son peptide de reconnaissance.

**Structure cristallographique de l'oncoprotéine E6 :** La structure cristallographique de E6 HPV16E6 en complexe avec son peptide ligand E6AP, présentée dans cette thèse constitue la première structure résolue des protéines E6 entières issues des virus HPV. Elle constitue une information structurale très attendue depuis environ une vingtaine d'années par la communauté scientifique, en raison de la place cruciale de E6 dans les processus pathogènes des HPV, notamment dans ceux du cancer du col de l'utérus. Cette structure est composée de deux domaines à zinc homologues (E6N et E6C), connectés par une longue hélice *alpha*. Le peptide, lui même replié en hélice *alpha*, s'introduit dans une poche constituée par les deux domaines de E6 et leur hélice de connection. Ce mode de liaison entre la protéine E6 et le peptide LxxLL, est une originalité par rapport aux structures de protéines cellulaires en complexe avec des motifs LxxLL, qui montrent une interaction plutôt superficielle (Hoellerer *et al.* [2003]; Lorenz *et al.* [2008]). La forte similarité entre les repliements des deux domaines à zinc, E6N et E6C montrée dans cette structure cristallographique de E6 entier et dans les structures en solution des domaines isolés publiées par notre équipe (Nomine *et al.* [2006]; Zanier *et al.* [2012]), pourrait soutenir le scénario évolutif qui suggère que les deux domaines de E6 proviennent d'une duplication d'un simple domaine ancestrale. Cette théorie est aussi soutenu par le fait que le type viral FLPV qui infecte les oiseaux possède une protéine E6 dont la séquence montre l'existence d'un seul domaine E6 semblable aux deux domaines E6N et E6C des autres types viraux. Une théorie évolutive plus large suggère que les deux oncoprotéines E6 et E7 ont une origine évolutive commune (Cole et Danos [1987]). Cette théorie s'appuie sur la similarité entre les séquences des deux domaines E6N et E6C et la séquence du domaine CR3 de E7 (2 copies de motifs CxxC séparés par 29 résidus), et sur la coexistence possible des mêmes motifs fonctionnels sur les deux protéines E6 et E7 issues des types viraux différents. A savoir, le motif LxCxE qui constitue classiquement le site d'interaction avec la protéine pRB chez beaucoup de protéines E7, qui est curieusement retrouvé dans la séquence de la protéine E6 du type MmuPV1 (*Mus musculus Papillomavirus* 1), récemment identifié (Joh *et al.* [2011]). Réciproquement le motif C-terminal qui forme la région de liaison avec les protéines à PDZ chez les protéines E6 des HPV à HRM, est retrouvé dans la séquence protéique de E7 du type viral RhPV (*Rhesus macaque Papillomavirus*) (Tomaic *et al.* [2009]). Cependant, les données structurales que nous avons présentées dans cette étude démontrent que malgré l'homologie de séquences entre E6 et E7

discutée ci-dessus, les repliements des E6N et E6C sont bien différents de celui de la structure publiée de E7.

**Analyse du système utilisé par E6 pour lier les protéines cellulaires :** Ce volet important de ma thèse avait pour objectif de répondre à la question suivante : comment certaines protéines E6, tout en reconnaissant la plupart de leurs ligands cellulaires via le même motif consensus général, LxxLL établissent-elles des préférences spéciales de chaque type E6 (sous-classes de motifs) ?

Par exemple, deux protéines E6 de deux souches phylogénétiquement éloignées HPV 16 (humaine) et BPV 1(bovine) reconnaissent majoritairement leurs cibles cellulaires par le biais d'un motif consensus LxxLL. Cependant, HPV16 E6 préfère lier la séquence ELTLQELLGEER issue de l'ubiquitine ligase E6AP, tandis que BPV1 E6, et non HPV16 E6, reconnaît la séquence DDL DALLADKE issue de la protéine d'adhésion focale Paxilline. Pour répondre à cette question, et analyser les variations entre les séquences consensus des protéines E6 appartenant à différentes souches nous avons choisi un panel de 7 souches qui couvrent les principaux genres de l'arbre phylogénétique des HPV (HPV1,3,5,9,16 et 18), plus la souche bovine BPV1. Après avoir produit et purifié les protéines E6 appartenant à ces 7 souches nous avons testé l'interaction de chaque E6 avec une liste de 93 peptides, dont la moitié issus de protéines cellulaires cibles de E6 extraites de la bibliographie et l'autre moitié issus de la mutagenèse du peptide E6AP.

Les tests d'interactions étaient faits par la technique chromatographique Hold-up, développée au sein de notre laboratoire.

Les résultats de ces environ 700 tests d'interaction ont fourni une base de données très riche, ce qui nous a permis d'attribuer à chaque variant de E6 analysé un spectre de préférence au sein du motif général LxxLL, et d'analyser les bases séquence-structure de la spécificité et de la sélectivité de chaque variant.

La richesse et la diversité de ces données interactomiques, ainsi que les données structurales structurales discutées plus haut, pourraient ouvrir la voie vers le criblage et la sélection *in silico*, *in vitro* et *in vivo* des molécules chimiques d'interférence a fin de concevoir des médicament anti-HPV capables notamment à traiter :

- les tumeurs du col de l'utérus induites par les HPV à haut risque (16 et 18) ;
- les tumeurs cutanées induites par les HPV cutanés à haut risque (05) ;
- les lésions cutanées des HPV cutanés à bas risques (1, 3 et 9).

Cependant cette étude reste préliminaire, et insuffisante pour en tirer des conclusion précises et exhaustives, en raison qu'elle est la première dans son genre (multiplex entre 7 protéines E6

et 93 peptides LxxLL) à notre connaissance et en raison du fait que les tests ont été effectués une seule fois pour une grande partie de ces résultats. Les perspectives qui sont ouvertes après ce travail sont essentiellement liées à une compréhension plus profonde des systèmes d'interactions propres à chaque variant.

Dans un proche avenir, nous comptons effectuer une autre étude à la lumière des informations collectées à partir de cette première étude, en utilisant une banque de peptides plus large, plus variée et plus ciblée et qui tiendra compte les différences de préférences entre les types de E6 déjà entrevue lors de cette première étude. Cette banque de peptides sera composée principalement, des peptides issues de la mutagenèse (alanine scanning mutagenesis) appliquée sur le meilleur peptide ligand pour chaque type de E6 selon les résultats de première étude. Parallèlement, nous allons utiliser la technique SPOT comme test comparatif, pour confirmer les résultats de la technique Hold-up.

D'autre part, nous avons suggéré un modèle comparatif, qui explique le système général d'interaction des protéines E6 avec les motifs cellulaires LxxLL, en proposant un classement des 24 résidus qui forment l'interface de reconnaissance selon la nature de leurs contribution dans la liaison, en 3 catégories :

- les résidus attractifs (résidus hydrophobes), qui servent à lier d'une façon générale les motifs LxxLL hydrophobes, sans jouer un rôle sélectif;
- les résidus sélectifs (résidus chargés positivement), responsables de la sélectivité qui détermine les champs de préférence propres à chaque type de E6 ;
- les résidus guides (résidus acides ou polaires), qui servent à recruter et orienter les chaînes latérales des résidus sélectifs vers les résidus partenaires dans le peptide.

Ce modèle a été basé sur l'observation des deux structures cristallographiques des complexes E6-LxxLL (HPV16 E6 et BPV1 E6), plus cinq autres structures modélisées de différents types de E6 (HPV 1, 3, 5, 9 et 18), et sur une analyse d'alignement des séquences. Une fois réalisée l'étude perspective décrite ci-dessus, nous serons très probablement en mesure de confirmer ou plutôt de raffiner le modèle d'interaction E6-LxxLL, en utilisant les outils de la Bioinformatique pour mieux analyser les contacts possibles entre tous les résidus du peptide LxxLL et chacun des 24 résidus qui forment l'interface de reconnaissance au sein de chaque type E6 (analyse par paires "pairwise"). Cela permettra sans doute la prédiction des listes des protéines cellulaires cibles pour telle ou telle protéine E6. L'exemple le plus proche de cette étude perspective, dans la littérature est celui de la méthode avancée par Chen et ses collègues qui ont suggéré une méthode de prédiction des peptides ligands des domaines PDZ ([Chen \*et al.\* \[2008\]](#)), en se basant sur une banque d'interaction multiplex entre 82 domaines de PDZ et 93 peptides.

**Nouvelle interaction entre la protéine virale E2 et la protéine cellulaire Brd4 :** En profitant de notre compétence méthodologique accumulée grâce aux travaux effectués sur la protéine E6, nous avons démarré un projet en collaboration avec M. Cheng-Ming Chiang (University of Texas Southwestern Medical School) dans le but de mieux comprendre la régulation de la protéine virale E2, médiée par la protéine cellulaire Brd4. Dans ce cadre, nous avons démontré par des analyses biophysiques (RMN), l'existence d'une nouvelle interaction entre le domaine C-terminal dimérisé de la protéine E2 du type HPV16 et des peptides phosphorylés issues de la région centrale BDII de la protéine cellulaire Brd4. La zone de reconnaissance identifiée constitue aussi l'interface de liaison entre l'ADN la protéine E2 (Abbate *et al.* [2006]). Cependant la forte influence sur cette interaction de la force ionique du solvant pourrait mettre en cause sa signification biologique. Un travail supplémentaire structural et biochimique sera nécessaire pour mieux caractériser cette interaction.

En conclusion, nous espérons que les résultats obtenus durant cette thèse fourniront des nouvelles connaissances cruciales pour étudier la co-évolution des oncoprotéines virales et de leurs cibles cellulaires, mais aussi pour comprendre les processus de la tumorigénèse induite par des virus oncogènes tels que les HPV.

---

# MATÉRIELS ET METHODES

## Production des protéines :

**Constructions d'ADN :** En plus de la protéine sauvage de HPV16 E6 et BPV1 E6, nous avons travaillé durant cette thèse sur plusieurs mutants de de HPV16 E6, à savoir : E6F47R, E6F47RC80S, E6F47RC97S, E6F47RC111S, E6F47RC140S, E6F47R4C4S et le mutant E6 F47R 4C4S  $\Delta$ 8CT qui porte en sus des mutations E6F47R4C4S une troncation des 8 derniers résidus du E6. Nous avons travaillé également sur le domaine C-terminal (DBD) de la protéine E2 (résidus : 285-365) du type HPV16.

**Bactéries d'expression :** La souche bactérienne que nous utilisons habituellement pour les expressions protéiques est la souche *E.coli* BL21 (DE3) (Invitrogen life technologies). Notre stock de cette souche bactérienne à été développé à partir d'une colonie offerte par M. Christophe ROMIER.

**Milieu de culture :** Nous utilisons deux types de milieux de culture bactérienne :

- Le LB (Lysogeny broth) : classique non marqué dont 1 L contient les ingrédients suivants : 10 g de digestion pancréatique de caséine, 5 g d'extrait de levures, 5 g de chlorure de sodium. Pour l'expression des protéines E6, nous rajoutons à LB 100  $\mu$ M de  $ZnSO_4$  au moment de l'induction de cultures, puisque E6 est une protéine à zinc, raison pour laquelle nous n'utilisons pas l'EDTA (acide éthylène diamine tétraacétique), pour ne pas chélater les ions du zinc nécessaires pour le bon repliement de la protéine.

- Milieu minimum M9 : Pour les expressions destinées aux analyses biophysiques en RMN, nous utilisons le milieu minimum M9 N15, qui permet le marquage des protéines par l'azote 15.



1 L du milieu contient : 8 g  $\text{Na}_2\text{HPO}_4$  (2  $\text{H}_2\text{O}$ ), 3 g  $\text{KH}_2\text{PO}_4$ , 0.5 g  $\text{NaCl}$ , 0.5 g  $\text{NH}_4\text{Cl}$ , 0.85 mg  $\text{ZnSO}_4$ , 8.4 g  $\text{FeCl}_3$  (6  $\text{H}_2\text{O}$ ), 0.10 mg  $\text{CuCl}_2$ , 0.1 mg  $\text{CoCl}_2$  (6  $\text{H}_2\text{O}$ ), 11 mg  $\text{MnCl}_2$  (2  $\text{H}_2\text{O}$ ), 1 mM  $\text{MgSO}_4$ , 0.33 mM  $\text{CaCl}_2$ , 4% D-glucose, 1 mg Biotine, 1 mg Thiamine (pH 7.5). Pour un marquage M9,  $\text{N14H}_4\text{Cl}$  est substitué par  $\text{N15NH}_4\text{Cl}$  (M9 15N).

**Préparation des BL21 électro-compétentes :** Pour préparer des cellules BL21 électro-compétentes, nous incubons une préculture en agitation à 37°C sur une nuit (13 à 15 h). Le lendemain, cette préculture sera diluée au 1/40ème (inoculation) sur un volume d'un litre du milieu de culture LB (Lysogeny broth) dans un fiole de 5 L, pour être incubée de nouveau à 37°C jusqu'à qu'elle atteigne une densité optique de 0,6 à 600 nm. Une fois cette densité optique est atteinte, nous effectuons la suite de la préparation à froid, nous plaçons d'abord les fioles sur la glace pendant 30 min et puis la culture sera répartie dans 4 tubes de 250 ml pour être centrifugé pendant 15 min à 4°C avec accélération de 3500 g pendant 15 minutes. Le culot est lavé rapidement deux fois avec de l'eau stérile (250 ml puis 125 ml). Et puis, un troisième lavage est effectué avec 5 ml de l'eau bidistillée stérile contenant 10% de glycérol. Finalement, le culot est resuspendu dans 500  $\mu\text{l}$  de l'eau contenant 10% de glycérol, reparti sur des aliquotes de 60  $\mu\text{l}$  par tube Eppendorf de 1,5 ml et transféré immédiatement au congélateur - 80°C.

**Transformation par électroporation :** Pour transformer les gènes codant pour les protéines que nous travaillons dessus (HPV16 E6, HPV16 E2... etc.), nous utilisons l'appareil Gene Pulser™ (BioRad), réglé à 200 $\Omega$ , 2,5 V et 25  $\mu\text{Fa}$ , pour une quantité de 60  $\mu\text{l}$  de bactéries. Les bactéries sont ensuite reprises dans 1 ml de LB, incubées 30 min à 37°C et diluées par la suite dans un volume de 10 ml de LB.

**Précultures :** La préculture est formée à partir d'une dilution de 1/200 de la solution de transformation, dans un volume de 10 ml de LB (ou 50 ml en cas de préculture destinée au marquage en milieu minimum M9 N15) incubé sur une nuit à 37 C° en agitation.

**Expression des protéines E6 :** Les précultures laissées sur une nuit à 37 C°, sont diluées au 1/500ème dans un volume 500 ml de LB qui contient la concentration adéquate d'antibiotique utilisée (15  $\mu\text{g}/\text{ml}$  Kanamycine). La culture est incubée sur un agitateur à 37C° jusqu'à qu'elle atteigne une densité optique de 0.7 à une longueur d'onde de 600 nm. L'induction de l'expression se fait grâce à l'IPTG à une concentration finale de 0,5 mM, simultanément nous rajoutons comme déjà mentionné le zinc sous forme de  $\text{Zn}_2+\text{SO}_4^{2-}$  à une concentration finale de 100  $\mu\text{M}$  final.

Pour favoriser le repliement de E6, l'incubation des cultures se fait avec une température de 15 C°, en agitation sur une nuit. le lendemain, 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). Finalement le culot est stocké dans un congélateur à -80°C jusqu'à son utilisation. En cas d'expressions des échantillons marqués par N15, les précultures (50 ml) sont récupérées et centrifugées dans des flacons de 50 ml à 3500 g pendant 20 min. Après avoir éliminé le surnageant, le culot est re-suspendu dans 10 ml de LB pour être ensuite dilué dans un volume de 500 ml de M9 et incubé en agitation sur nuit à 15C°.

**Lyse bactérienne et purification du complexe MBP-E6 :** Après culture et centrifugation, le culot cellulaire 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, plus une tablette d'un cocktail d'inhibiteur de protéases (Roche) pour chaque 50 ml de tampon). En moyenne un culot cellulaire obtenu à partir d'un litre de culture est suspendu dans 50 ml de tampon de lyse. La sonication est effectuée sur des volumes de 25 à 75 ml avec une sonde 50 mm (BRANSON, Digital Sonifier®) à amplitude de 20 à 50% par cycle de 1 sec d'ultrasons et 3 sec de pause pendant 2 minutes. L'échantillon est replacé dans glace pour éviter un échauffement trop fort. Les extraits bruts sont alors centrifugés pendant 1 heure à 80 000 x g dans une centrifugeuse L-70 (Beckman). Après filtration (filtre à 0,22 µM), le surnageant est déposé sur une colonne amylose, résine spécifique de la protéine MBP (Maltose Binding Protein), préalablement équilibrée dans un tampon A (50 mM Tris-HCl pH 6,8, 400 mM NaCl, 2 DTT). La colonne est ensuite lavée avec 8 à 10 volumes de tampon B (A + cocktail d'inhibiteurs de protéases). L'élution du complexe MBP-E6, réalisée par addition de maltose 10 mM dans le tampon B est suivie par spectrophotométrie.

L'éluat est placés dans des tubes UltraClear® de 12,5 ml dans un rotor SW41 Ti puis centrifugé à 38 000 rpm (200 000 x g) sur une centrifugeuse L-70 (Beckman) réfrigérée à 4° pendant 15 heures. Le surnageant (solution protéique) est prélevé très délicatement pour éviter la suspension du culot. Par précaution, les 2 ml séparant le surnageant et le culot ne sont pas prélevés.

**Digestion du complexe MBP-E6 :** La séparation de MBP et de la protéine E6 est réalisée par digestion avec la TEV à 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. La protéine E6 éluée est alors concentrée par centrifugation dans un concentrateur Vivaspin qui retient les molécules de PM > 5kDa (VIASCIENCE, 5/15) jusqu'à la concentration voulue pour les mesures.

**Précautions particulières pour optimiser la purification de E6 :** Dans la purification de E6, certaines précautions doivent être prises pour minimiser la tendance de cette protéine à se précipiter.

**Tampon :** Tous les tampons doivent être filtrés et soigneusement dégazés

**Température :** la protéine et tous les tampons doivent rester à 4°C ou dans la glace : à défaut de pouvoir travailler en chambre froide, un bac de glace doit être constamment à disposition.

**Sonication :** La température est d'autant plus critique dans cette procédure car l'énergie de vibration des molécules peut entraîner une élévation locale de température assez forte en plus de la formation de bulles d'air (mousse). La sonication doit être donc rapide et de puissance adaptée à la taille de l'échantillon. Si le temps de sonication doit être long pour effectuer la lyse des cellules, il est plus adapté de faire plusieurs coups de sonication séparés de temps d'arrêt pour éviter une longue exposition de la protéine à une température élevée.

**L'oxydation :** Après la lyse cellulaire, la protéine devient plus fragile. En plus de la température qui devient plus critique, elle doit être maintenue dans un milieu réducteur (comme DTT 1mM) pour éviter l'oxydation due à l'air ambiant ou un dégazage non efficace d'un tampon. Cette précaution doit être prise pendant et après l'étape de purification pendant le stockage. Un agent antibactérien (azide de Sodium) est parfois utile pendant la conservation.

**Concentration :** Dans le milieu cellulaire la majorité des protéine est à haute concentration, ce qui permet, entre autre, une meilleure stabilité. Pour favoriser cet état, la protéine doit être concentrée tout en s'assurant que la concentration saline initiale (400 mM NaCl) soit maintenue. Il faudrait penser à utiliser un système de concentration qui ne retient que la protéine (système Amicon ou autre). Cependant, il faut éviter une concentration excessive car en plus de la perte de protéine inhérente au système de concentration une concentration élevée de la protéine peut entraîner son agrégation.

**Dosage et pureté :** Cette étape bien que nécessaire pourrait être une source de perte ou dégradation de la protéine. Pour éviter cela, il suffit de prendre un échantillon de la protéine (dans un eppendorf) qu'on place dans la glace et qu'on utilisera. Le reste de la protéine est alors stocké sans risque.

### Cristallographie du HPV16 E6/E6AP) :

**Principe :** Le complexe HPV16 E6/E6AP a été produit en fusionnant un mutant de la maltose binding protein (MBP) à l'extrémité N-terminale du Peptide ligand LXXLL pour favoriser la cristallisation. Des mutations ponctuelles ( D83A, K84A, K240A, E360A, K363A, D364A) de la MBP ont choisies car elles augmentent la propension de cette protéine à cristalliser ([Kobe \*et al.\* \[1999\]](#), [Ullah \*et al.\* \[2008\]](#)). Parallèlement, un oligomère codant pour les résidus 403-414 (ELTLQELLGEER) de E6AP a été cloné par l'intermédiaire d'un linker de trois alanine à l'extrémité C-terminale du mutan MBP, générant ainsi une constuction de fusion MBP-E6AP. Pour la protéine HPV16 E6, les résidus 1-143 du mutant E6 F47R 4C/4S (E6 F47R 4C4S  $\Delta$ 8CT) ont été amplifiés par PCR et clonés dans les sites NcoI et KpnI du vecteur PETM-41 contenant une tag His6-MBP en position N-terminale suivie par un site de clivage TEV. Ce qui a engendré une constuction de fusion (His6-MBP-tev-F47R E6 4C/4S) où le tag His6-MBP peut être séparé de E6 par digestion par la protéase TEV.

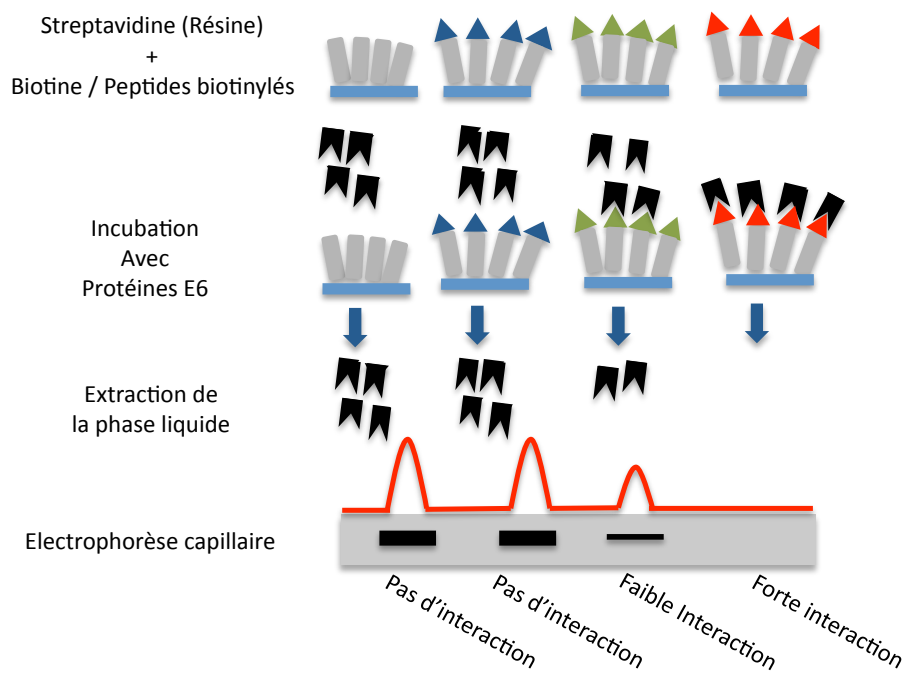
Les constructions MBP-E6AP et His6- MBP-tev-E6 F47R 4C4S  $\Delta$ 8CT ont été exprimées selon le protocole décrit plus haut. Les échantillons MBP-E6AP sont chargés soit sur une colonne de filtration sur gel de forte capacité (Superdex 75 HiLoad 16/60) . par contre, les échantillons de His6-MBP-tev-F47R E6 4C/4S sont digérés par la protéase TEV avant le chargement sur la colonne (Superdex 75 HiLoad 16/60). La protéine de fusion et F47R 4C/4S E6 éluent les deux sous forme de pics monomériques. Le complexe HPV16 E6/E6AP est reconstitué en mélangeant les échantillons de MBP-E6 E6AP et F47R 4C/4S dans un rapport de stœchiométrique 1 :1. Les conditions de tampon utilisées sont 50 mM Tris /pH6,8, 200 mM NaCl, 2 mM DTT et 5 mM maltose pour la 4C/4S F47R E6 / MBP-E6ap. les échantillons E6 F47R 4C/4S / MBP-E6AP ont été concentrés à 80 mg / ml avant la cristallisation.

**Cristallisation** Les conditions de cristallisation ont été testées en utilisant des kits commerciaux (Qiagen) par la méthode sitting-drop vapor-diffusion dans des plaques 96 puits MRC 2-drops employant un Robot cartésien. Après optimisation, les cristaux de HPV16 E6/E6AP (200 x 200 x 200  $\mu$ m) utilisables pour la diffraction des rayons X ont été obtenus en une semaine à partir de gouttes assises constituées par le mélange de 200 nl de solution de protéique à 80

mg / ml et 200 nl de la solution du réservoir formée 10% de PEG 8K, 0,05 M CHES pH 9,5. Les gouttes ont été équilibrées contre 50 ul de la solution du réservoir à 290K. Les cristaux ont été transférés de manière séquentielle à cinq cryo solutions contenant des concentrations croissantes de PEG 400 jusqu'à ce qu'une concentration finale de 25% PEG 400 ait été atteinte. Les cristaux ont été placés rapidement dans l'azote liquide.

**Collecte et traitement des données, détermination de la structure** Toutes les données de diffraction des rayons X ont été recueillies à l'European Synchrotron Radiation Facility (ESRF) sur la ligne ID14-4. Les données ont été acquises à partir d'un seul cristal cryo-refroidi (100K) sur un détecteur Q315rCCD (ADSC). 180 ° de données ont été recueillies en utilisant 0,4° de rotation et 0,3 s de temps d'exposition pour chaque image. Les données ont été traitées et mesurées en utilisant HKL2000. Les cristaux appartiennent au groupe d'espace  $P2_12_12_1$  (système orthorhombique), avec des paramètres de maille  $a = 106.3\text{\AA}$ ,  $b = 134.9\text{\AA}$ ,  $c = 138.8\text{\AA}$ . La structure a été résolue par deux remplacements moléculaires séquentiels utilisant MOLREP en utilisant d'abord la MBP et ensuite BPV1 E6 comme modèle. Deux copies de MBP-E6AP et une copie de HPV16 E6 ont été trouvées. E6 a été reconstruite et utilisé pour un troisième tour de remplacement moléculaire pour localiser la seconde molécule de E6. L'unité asymétrique contient deux copies du complexe MBP-E6AP/HPV16-E6, avec un coefficient de Matthews correspondant de  $4,21 \text{\AA}^3/\text{Da}$  et une teneur en solvant de 70,8%. L'affinement de la structure a été effectué en utilisant PHENIX et Buster, puis par un "iterative model building" (Arp/wARP). La qualité des modèles affinés a été évaluée à l'aide de MOLPROBITY (voir annexe 2, tableau 1).

**Tests d'interaction par la technique de chromatographie Hold-up :** Pour étudier in vitro l'interaction de la protéine E6 avec des diverses séquences LxxLL et évaluer leurs affinités relatives, nous avons utilisé une version modifiée de la technique de rétention chromatographique comparative hold-up ([Charbonnier \*et al.\* \[2006\]](#)), basée sur l'affinité avidine-biotine. pour réaliser une étude multiplexe entre 93 peptides biotinylés, du motif LxxLL (ou motif proche), dérivés de protéines humaines et qui ont été prédits comme liant E6 ont été chimiquement synthétisés (JPT Peptide Technologies, Berlin) (voir tableau), contre un panel de 7 types de protéine E6 issues de divers types viraux qui couvrent six genres de l'arbre phylogénétique des HPV (HPV1,3,5,9,16 et 18) ainsi que la souche bovine BPV1.



### Principe de la technique de Hold-up

Schéma simplifié explicatif de la technique de chromatographie Hold-up, montrant les étapes principales : L'incubation de la résine streptavidine (gris) avec les peptides biotylés (bleu, vert et rouge), l'incubation de la résine saturée avec la protéine (noire), l'extraction de la phase liquide et l'analyse de l'output grâce à l'électrophorèse. (d'après Luck K./Charbonnier S. et al. manuscrit en cours de préparation).

## Peptides LxxLL

Série 1			Série 2		
N°	Nom	Séquence	N°	Nom	Séquence
1	WT	PESELTLQELLGEEER	1	ADA3_1	QLELETLLSSAS
2	A <sub>1</sub>	PESSALTLQELLGEEER	2	ADA3_2	VRTLEELLKPPE
3	R <sub>1</sub>	PESSRLTLQELLGEEER	3	AP1B1	GGGLDSLMGDEP
4	A <sub>3</sub>	PESELALQELLGEEER	4	AP1G1_1	GGELLDLLGDIN
5	A <sub>4</sub>	PESELTAQELLGEEER	5	AP1G1_2	ANDLLDLLGGND
6	V <sub>4</sub>	PESELTVQELLGEEER	6	BRCA1_1	IQKVNEWFSRSD
7	I <sub>4</sub>	PESELTIQELLGEEER	7	BRCA1_2	FSRSDPELLGSDD
8	F <sub>4</sub>	PESELTFQELLGEEER	8	BRCA1_3	ERTLKYFLGIAG
9	W <sub>4</sub>	PESELTWQELLGEEER	9	CBP	HKQLSELLRGGSS
10	Y <sub>4</sub>	PESELTYQELLGEEER	10	DBC1	KSQLQRLLQELR
11	M <sub>4</sub>	PESELTMQELLGEEER	11	DNJA1	LSLLEKLLPERK
12	S <sub>4</sub>	PESELTSQELLGEEER	12	E6AP	PESELTLQELLGEEER
13	A <sub>5</sub>	PESELTLAELLGEEER	13	E6BP	GFVSLLEEFLLGDY
14	E <sub>5</sub>	PESELTLLEELLGEEER	14	E6TP1	PTKLSDFLLITGG
15	D <sub>5</sub>	PESELTLDEELLGEEER	15	EBP1	DAELKALLQSSA
16	A <sub>6</sub>	PESELTLQALLGEEER	16	Emerin	DTELTTLRRYN
17	D <sub>6</sub>	PESELTLQDLLGEEER	17	Gps2	FLQLKKVLLHEEE
18	Q <sub>6</sub>	PESELTLQQLGEEER	18	IRF3	EDILDELLGNMV
19	R <sub>6</sub>	PESELTLQRLLGEEER	19	MAML1	MSDLDLGLSQ
20	A <sub>7</sub>	PESELTLQEALGEEER	20	MED13	YTDLDNLFNSDE
21	V <sub>7</sub>	PESELTLQEVLLGEEER	21	p300	QQALQNLRLTLR
22	I <sub>7</sub>	PESELTLQEVLLGEEER	22	p53	FSDLWKLLEPENN
23	F <sub>7</sub>	PESELTLQEFLLGEEER	23	p53bos_1	FSDLWNLEPENN
24	W <sub>7</sub>	PESELTLQEWLLGEEER	24	p53bos_2	SQETFSDLWNLEPENN
25	Y <sub>7</sub>	PESELTLQEYLLGEEER	25	p53chick_1	FMDLWSMLPYSM
26	M <sub>7</sub>	PESELTLQEMLLGEEER	26	p53chick_2	TEVFMDLWSMLPYSMQQ
27	A <sub>8</sub>	PESELTLQEELGEEER	27	PAB1	DNSELLHMLLESP
28	V <sub>8</sub>	PESELTLQELVLLGEEER	28	Piasy_1	VSDLQMLLGFVG
29	I <sub>8</sub>	PESELTLQELLLGEEER	29	Piasy_2	FNMLDELLKPTE
30	F <sub>8</sub>	PESELTLQELFLGEEER	30	PXN_1	MDDL DALLADLE
31	W <sub>8</sub>	PESELTLQELWLLGEEER	31	PXN_2	LSELDRLLLELN
32	Y <sub>8</sub>	PESELTLQELYLLGEEER	32	E6APxn2	PESSLSELDRLLLELN
33	M <sub>8</sub>	PESELTLQELMLLGEEER	33	PXN_4	TRELDL MASLS
34	A <sub>9</sub>	PESELTLQEELLAEEER	34	SGT1	PDQLDQLLQEA
35	L <sub>9</sub>	PESELTLQEELLLLEER	35	TGFB1_1	MEDLDALLSDLE
36	Q <sub>9</sub>	PESELTLQEELLLQEEER	36	TGFB1_2	LCELDRLLQELN
37	E <sub>9</sub>	PESELTLQEELLLLEER	37	TGFB1_3	TLELDRLL MASLS
38	S <sub>9</sub>	PESELTLQEELLLSEER	38	TRIP-Br1_1	SASMALLEDLS
39	A <sub>10</sub>	PESELTLQEELLLGAER	39	TRIP-Br1_2	DDGLEGLFEDID
40	D <sub>10</sub>	PESELTLQEELLLGDER	40	TRIP-Br1_3	EAELDYLM DVLV
41	Q <sub>10</sub>	PESELTLQEELLLQGER	41	TRIP-Br1_4	LDYLM DVLVGTQ
42	R <sub>10</sub>	PESELTLQEELLLGRER	42	TRIP-Br1_5	EAELDYLM DVLVGTQ
43	A <sub>11</sub>	PESELTLQEELLLGEAR	43	TSC2_1	LQTLQDILGDPG
44	D <sub>11</sub>	PESELTLQEELLLGEDR	44	TSC2_2	ELQTLQDILGDP
45	Q <sub>11</sub>	PESELTLQEELLLGEQR	45	WIZ	INILQEELIATSA
46	R <sub>11</sub>	PESELTLQEELLLGERR	46	XRCC1	PEELGKILQGVV
			47	Zyxin	IDSLSLLDDMT

Deux peptides biotinylés supplémentaires ont également été synthétisés comme témoins négatifs : 16E6-CT (SSRTRRETQL) correspondant au fragment C-terminal de HPV16 E6 liant PDZ et CTRL (GSNSGNGNS) qui représente une séquence polaire non chargée non impliquée dans la liaison. Chaque peptide biotinylé est dissout dans du tampon PBS à une concentration de 1 mM, puis incubé avec 20  $\mu$ l streptavidine-sépharose de haute performance (GE-Healthcare). La Peptide saturé de résine est ensuite mis à incuber avec 2 mM biotine pour saturer tous les sites de streptavidine. Nous avons également préparé un lot supplémentaire de résine de contrôle ne contenant que la biotine. La concentration résultante de peptide biotinylé (estimée à partir des spécifications de la résine fournies par le fabricant) au sein de la résine humide est de 150-200 mM environ 15  $\mu$ l résine saturée de peptide ou résine saturée de biotine sont ensuite mélangés et incubés pendant 15 min avec 30 $\mu$ l de E6 MBP-BPV-1 monomérique fraîchement purifiés. Chaque mélange de résine-protéine contient environ une concentration de 50  $\mu$ M de peptide et 4  $\mu$ M de protéine MBP-E6. Les mélanges résine-protéine sont finalement transférés dans un dispositif filtrant centrifugeable. La phase liquide phases est rapidement extraite. Un volume égal de chaque éluat est testé par SDD PAGE. La quantité MBP-E6 liée à chaque peptide est quantifiée en soustrayant la quantité de MBP-E6 présente dans l'éluat de la résine saturée de peptide de la quantité de MBP-E6 trouve dans l'éluat de la résine saturée de biotine.

**Bioinformatique, Statistiques et Graphisme** Nous avons utilisé durant cette thèse une bonne panoplie de logiciels de bioinformatiques, de statistiques ou de graphisme. Ces logiciels ont permit la réalisation des figures de structures protéiques et calculs de surface électrique ou hydrophobe des protéines (Chimera UCSF et PyMOL), analyse de données et leur représentation en forme de 'heat maps' (Excel et R), alignement de séquences (CLUSTAL), prédiction (IUPRED), modélisation par homologie (MODELLER) en plus de la gestion de bibliographies et édition de textes ( $\text{\LaTeX}$ ).



## Références bibliographiques

---

---

## Références bibliographiques

- ABBATE, E. A., BERGER, J. M. et BOTCHAN, M. R. (2004). The x-ray structure of the papillomavirus helicase in complex with its molecular matchmaker e2. *Genes Dev*, 18(16):1981–1996. [27](#)
- ABBATE, E. A., VOITENLEITNER, C. et BOTCHAN, M. R. (2006). Structure of the papillomavirus dna-tethering complex e2:brd4 and a peptide that ablates hpv chromosomal association. *Mol Cell*, 24(6):877–889. [30](#), [106](#), [120](#)
- ACCARDI, R., RUBINO, R., SCALISE, M., GHEIT, T., SHAHZAD, N., THOMAS, M., BANKS, L., INDIVERI, C., SYLLA, B. S., CARDONE, R. A., RESHKIN, S. J. et TOMMASINO, M. (2011). E6 and e7 from human papillomavirus type 16 cooperate to target the pdz protein na/h exchange regulatory factor 1. *J Virol*, 85(16):8208–8216. [39](#), [43](#)
- ALPHS, H. H., GAMBHIRA, R., KARANAM, B., ROBERTS, J. N., JAGU, S., SCHILLER, J. T., ZENG, W., JACKSON, D. C. et RODEN, R. B. S. (2008). 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*, 105(15):5850–5855. [24](#)
- AN, J., MO, D., LIU, H., VEENA, M. S., SRIVATSAN, E. S., MASSOUMI, R. et RETTIG, M. B. (2008). Inactivation of the cyld deubiquitinase by hpv e6 mediates hypoxia-induced nf-kappab activation. *Cancer Cell*, 14(5):394–407. [39](#), [43](#)
- ANDROPHY, E. J., HUBBERT, N. L., SCHILLER, J. T. et LOWY, D. R. (1987a). Identification of the hpv-16 e6 protein from transformed mouse cells and human cervical carcinoma cell lines. *EMBO J*, 6(4):989–992. [47](#)
- ANDROPHY, E. J., LOWY, D. R. et SCHILLER, J. T. (1987b). Bovine papillomavirus e2 trans-activating gene product binds to specific sites in papillomavirus dna. *Nature*, 325(6099):70–73. [27](#), [105](#), [114](#)
- ANTINORE, M. J., BIRRER, M. J., PATEL, D., NADER, L. et McCANCE, D. J. (1996). The human papillomavirus type 16 e7 gene product interacts with and trans-activates the ap1 family of transcription factors. *EMBO J*, 15(8):1950–1960. [13](#), [35](#)
- ANTONSSON, A., PAYNE, E., HENGST, K. et MCMILLAN, N. A. J. (2006). The human papillomavirus type 16 e7 protein binds human interferon regulatory factor-9 via a novel pest domain required for transformation. *J Interferon Cytokine Res*, 26(7):455–461. [35](#)
- ANTSON, A. A., BURNS, J. E., MOROZ, O. V., SCOTT, D. J., SANDERS, C. M., BRONSTEIN, I. B., DODSON, G. G., WILSON, K. S. et MAITLAND, N. J. (2000). Structure of the intact transactivation domain of the human papillomavirus e2 protein. *Nature*, 403(6771):805–809. [27](#)
- ASHRAFI, G. H., HAGHSHENAS, M., MARCHETTI, B. et CAMPO, M. S. (2006). E5 protein of human papillomavirus 16 downregulates hla class i and interacts with the heavy chain via its first hydrophobic domain. *Int J Cancer*, 119(9):2105–2112. [32](#)

## Références bibliographiques

---

- AVVAKUMOV, N., TORCHIA, J. et MYMRYK, J. S. (2003). Interaction of the hpv e7 proteins with the pcaf acetyltransferase. *Oncogene*, 22(25):3833–3841. [35](#)
- BAKER, T. S., NEWCOMB, W. W., OLSON, N. H., COWSERT, L. M., OLSON, C. et BROWN, J. C. (1991). Structures of bovine and human papillomaviruses. analysis by cryoelectron microscopy and three-dimensional image reconstruction. *Biophys J*, 60(6):1445–1456. [9](#)
- BALEJA, J. D., CHERRY, J. J., LIU, Z., GAO, H., NICKLAUS, M. C., VOIGT, J. H., CHEN, J. J. et ANDROPHY, E. J. (2006). Identification of inhibitors to papillomavirus type 16 e6 protein based on three-dimensional structures of interacting proteins. *Antiviral Res*, 72(1):49–59. [45](#)
- BARBOSA, M. S., EDMONDS, C., FISHER, C., SCHILLER, J. T., LOWY, D. R. et VOUSDEN, K. H. (1990). 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*, 9(1):153–160. [35](#)
- BARNARD, P. et McMILLAN, N. A. (1999). The human papillomavirus e7 oncoprotein abrogates signaling mediated by interferon-alpha. *Virology*, 259(2):305–313. [35](#)
- BARRON, B. A. et RICHART, R. M. (1981). Screening protocols for cervical neoplastic disease. *Gynecol Oncol*, 12(2 Pt 2):S156–67. [17](#)
- BEGUE, P., HENRION, R., BLANC, B., GIRARD, M. et SANCHO-GARNIER, H. (2007). Vaccination against human papillomavirus. implementation and efficacy against cervical cancer control]. *Bull Acad Natl Med*, 191(9):1805–1816. [20](#)
- BELNAP, D. M., OLSON, N. H., CLADEL, N. M., NEWCOMB, W. W., BROWN, J. C., KREIDER, J. W., CHRISTENSEN, N. D. et BAKER, T. S. (1996). Conserved features in papillomavirus and polyomavirus capsids. *J Mol Biol*, 259(2):249–263. [9](#)
- BEN-SAADON, R., FAJERMAN, I., ZIV, T., HELLMAN, U., SCHWARTZ, A. L. et CIECHANOVER, A. (2004). 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*, 279(40):41414–41421. [34](#)
- BEREZUTSKAYA, E. et BAGCHI, S. (1997). The human papillomavirus e7 oncoprotein functionally interacts with the s4 subunit of the 26 s proteasome. *J Biol Chem*, 272(48):30135–30140. [35](#)
- BERNARD, H.-U. (2002). Gene expression of genital human papillomaviruses and considerations on potential antiviral approaches. *Antivir Ther*, 7(4):219–237. [9](#)
- BERNARD, H.-U., BURK, R. D., CHEN, Z., van DOORSLAER, K., HAUSEN, H. z. et de VILLIERS, E.-M. (2010). Classification of papillomaviruses (pvs) based on 189 pv types and proposal of taxonomic amendments. *Virology*, 401(1):70–79. [4](#), [5](#), [6](#), [7](#)
- BERNAT, A., AVVAKUMOV, N., MYMRYK, J. S. et BANKS, L. (2003). Interaction between the hpv e7 oncoprotein and the transcriptional coactivator p300. *Oncogene*, 22(39):7871–7881. [35](#)
- BISCHOF, O., NACERDDINE, K. et DEJEAN, A. (2005). Human papillomavirus oncoprotein e7 targets the promyelocytic leukemia protein and circumvents cellular senescence via the rb and p53 tumor suppressor pathways. *Mol Cell Biol*, 25(3):1013–1024. [35](#)
- BISCHOF, O., SCHWAMBORN, K., MARTIN, N., WERNER, A., SUSTMANN, C., GROSSCHEDL, R. et DEJEAN, A. (2006). The e3 sumo ligase pi3y is a regulator of cellular senescence and apoptosis. *Mol Cell*, 22(6):783–794. [39](#), [44](#)
- BISHOP, B., DASGUPTA, J., KLEIN, M., GARCEA, R. L., CHRISTENSEN, N. D., ZHAO, R. et CHEN, X. S. (2007). Crystal structures of four types of human papillomavirus l1 capsid proteins: understanding the specificity of neutralizing monoclonal antibodies. *J Biol Chem*, 282(43):31803–31811. [9](#), [22](#)
- BODAGHI, S., JIA, R. et ZHENG, Z.-M. (2009). Human papillomavirus type 16 e2 and e6 are rna-binding proteins and inhibit in vitro splicing of pre-mrnas with suboptimal splice sites. *Virology*, 386(1):32–43. [39](#), [45](#)

- BODILY, J. M., MEHTA, K. P. M. et LAIMINS, L. A. (2011). Human papillomavirus e7 enhances hypoxia-inducible factor 1-mediated transcription by inhibiting binding of histone deacetylases. *Cancer Res*, 71(3):1187–1195. 34, 35
- BOHL, J., DAS, K., DASGUPTA, B. et VANDE POL, S. B. (2000). Competitive binding to a charged leucine motif represses transformation by a papillomavirus e6 oncoprotein. *Virology*, 271(1):163–170. 53
- BORBELY, A. A., MURVAI, M., KONYA, J., BECK, Z., GERGELY, L., LI, F. et VERESS, G. (2006). Effects of human papillomavirus type 16 oncoproteins on survivin gene expression. *J Gen Virol*, 87(Pt 2):287–294. 40
- BOSHART, M., GISSMANN, L., IKENBERG, H., KLEINHEINZ, A., SCHEURLEN, W. et zur HAUSEN, H. (1984). A new type of papillomavirus dna, its presence in genital cancer biopsies and in cell lines derived from cervical cancer. *EMBO J*, 3(5):1151–1157. 4
- BOUVARD, V., STOREY, A., PIM, D. et BANKS, L. (1994). Characterization of the human papillomavirus e2 protein: evidence of trans-activation and trans-repression in cervical keratinocytes. *EMBO J*, 13(22):5451–5459. 27
- BOYER, S. N., WAZER, D. E. et BAND, V. (1996). E7 protein of human papilloma virus-16 induces degradation of retinoblastoma protein through the ubiquitin-proteasome pathway. *Cancer Res*, 56(20):4620–4624. 35
- BREHM, A., NIELSEN, S. J., MISKA, E. A., MCCANCE, D. J., REID, J. L., BANNISTER, A. J. et KOUZARIDES, T. (1999). The e7 oncoprotein associates with mi2 and histone deacetylase activity to promote cell growth. *EMBO J*, 18(9):2449–2458. 35
- BRIMER, N., LYONS, C., WALLBERG, A. E. et VANDE POL, S. B. (2012). Cutaneous papillomavirus e6 oncoproteins associate with mam11 to repress transactivation and notch signaling. *Oncogene*. 39, 42
- BRINKMAN, J. A., HUGHES, S. H., STONE, P., CAFFREY, A. S., MUDERSPACH, L. I., ROMAN, L. D., WEBER, J. S. et KAST, W. M. (2007). Therapeutic vaccination for hpv induced cervical cancers. *Dis Markers*, 23(4):337–352. 20
- BROWN, C., CAMPOS-LEON, K., STRICKLAND, M., WILLIAMS, C., FAIRWEATHER, V., BRADY, R. L., CRUMP, M. P. et GASTON, K. (2011). Protein flexibility directs dna recognition by the papillomavirus e2 proteins. *Nucleic Acids Res*, 39(7):2969–2980. 107
- BUCK, C. B., CHENG, N., THOMPSON, C. D., LOWY, D. R., STEVEN, A. C., SCHILLER, J. T. et TRUS, B. L. (2008). Arrangement of l2 within the papillomavirus capsid. *J Virol*, 82(11):5190–5197. 9
- BURKHARDT, A., DiMAIO, D. et SCHLEGEL, R. (1987). Genetic and biochemical definition of the bovine papillomavirus e5 transforming protein. *EMBO J*, 6(8):2381–2385. 31
- CAIRNS, P., OKAMI, K., HALACHMI, S., HALACHMI, N., ESTELLER, M., HERMAN, J. G., JEN, J., ISAACS, W. B., BOVA, G. S. et SIDRANSKY, D. (1997). Frequent inactivation of pten/mmac1 in primary prostate cancer. *Cancer Res*, 57(22):4997–5000. 43
- CAMPO-FERNANDEZ, B., MORANDELL, D., SANTER, F. R., ZWERSCHKE, W. et JANSEN-DURR, P. (2007). Identification of the flh2 transcriptional coactivator as a new functional target of the e7 oncoprotein of human papillomavirus type 16. *J Virol*, 81(2):1027–1032. 35
- CEBALLOS, A., REMES LENICOV, F., SABATTE, J., RODRIGUEZ RODRIGUES, C., CABRINI, M., JANCIC, C., RAIDEN, S., DONALDSON, M., AGUSTIN PASQUALINI, R. J., MARIN-BRIGGILER, C., VAZQUEZ-LEVIN, M., CAPANI, F., AMIGORENA, S. et GEFFNER, J. (2009). Spermatozoa capture hiv-1 through heparan sulfate and efficiently transmit the virus to dendritic cells. *J Exp Med*, 206(12):2717–2733. 24
- CHAMBERS, G., ELLSMORE, V. A., O'BRIEN, P. M., REID, S. W. J., LOVE, S., CAMPO, M. S. et NASIR, L. (2003). Association of bovine papillomavirus with the equine sarcoid. *J Gen Virol*, 84(Pt 5):1055–1062. 3
- CHARBONNIER, S., ZANIER, K., MASSON, M. et TRAVE, G. (2006). Capturing protein-protein complexes at equilibrium: the holdup comparative chromatographic retention assay. *Protein Expr Purif*, 50(1):89–101. 78, 126
- CHEN, J. J., REID, C. E., BAND, V. et ANDROPHY, E. J. (1995). Interaction of papillomavirus e6 oncoproteins with a putative calcium-binding protein. *Science*, 269(5223):529–531. 39, 44

## Références bibliographiques

---

- CHEN, J. R., CHANG, B. H., ALLEN, J. E., STIFFLER, M. A. et MACBEATH, G. (2008). Predicting pdz domain-peptide interactions from primary sequences. *Nat Biotechnol*, 26(9):1041–1045. [119](#)
- CHEN, X. S., GARCEA, R. L., GOLDBERG, I., CASINI, G. et HARRISON, S. C. (2000). Structure of small virus-like particles assembled from the l1 protein of human papillomavirus 16. *Mol Cell*, 5(3):557–567. [9](#), [10](#), [21](#), [22](#), [23](#), [29](#)
- CHINAMI, M., MORIYAMA, K., FUKUMAKI, Y., TERADA, M. et SHINGU, M. (1993). Association of rna with human papillomavirus e7 protein of type 16 but not type 6b. *Biochem Biophys Res Commun*, 197(3):1609–1614. [34](#), [35](#)
- CHOE, J., VAILLANCOURT, P., STENLUND, A. et BOTCHAN, M. (1989). Bovine papillomavirus type 1 encodes two forms of a transcriptional repressor: structural and functional analysis of new viral cdnas. *J Virol*, 63(4):1743–1755. [8](#)
- CHOI, J., SOUTHWORTH, L. K., SARIN, K. Y., VENTEICHER, A. S., MA, W., CHANG, W., CHEUNG, P., JUN, S., ARTANDI, M. K., SHAH, N., KIM, S. K. et ARTANDI, S. E. (2008). Tert promotes epithelial proliferation through transcriptional control of a myc- and wnt-related developmental program. *PLoS Genet*, 4(1):e10. [41](#)
- CLERTANT, P. et SEIF, I. (1984). A common function for polyoma virus large-t and papillomavirus e1 proteins. *Nature*, 311(5983):276–279. [4](#)
- COLE, S. T. et DANOS, O. (1987). 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*, 193(4):599–608. [117](#)
- CRUM, C. P., NUOVO, G., FRIEDMAN, D. et SILVERSTEIN, S. J. (1988). Accumulation of rna homologous to human papillomavirus type 16 open reading frames in genital precancers. *J Virol*, 62(1):84–90. [12](#)
- CRUZ, M. R., CERQUEIRA, D. M., CRUZ, W. B., CAMARA, G. N. L., BRIGIDO, M. M., SILVA, E. O., CARVALHO, L. G. S. et MARTINS, C. R. F. (2004). Prevalence of human papillomavirus type 16 variants in the federal district, central brazil. *Mem Inst Oswaldo Cruz*, 99(3):281–282. [4](#)
- DALAL, S., GAO, Q., ANDROPHY, E. J. et BAND, V. (1996). Mutational analysis of human papillomavirus type 16 e6 demonstrates that p53 degradation is necessary for immortalization of mammary epithelial cells. *J Virol*, 70(2):683–688. [38](#)
- DANOS, O., KATINKA, M. et YANIV, M. (1982). Human papillomavirus 1a complete dna sequence: a novel type of genome organization among papovaviridae. *EMBO J*, 1(2):231–236. [8](#)
- DAVEY, N. E., TRAVE, G. et GIBSON, T. J. (2011). How viruses hijack cell regulation. *Trends Biochem Sci*, 36(3):159–169. [102](#)
- DAVIES, R., HICKS, R., CROOK, T., MORRIS, J. et VOUSDEN, K. (1993). Human papillomavirus type 16 e7 associates with a histone h1 kinase and with p107 through sequences necessary for transformation. *J Virol*, 67(5):2521–2528. [35](#)
- DAY, P. M., RODEN, R. B., LOWY, D. R. et SCHILLER, J. T. (1998). The papillomavirus minor capsid protein, l2, induces localization of the major capsid protein, l1, and the viral transcription/replication protein, e2, to pml oncogenic domains. *J Virol*, 72(1):142–150. [14](#)
- DE GEEST, K., TURYK, M. E., HOSKEN, M. I., HUDSON, J. B., LAIMINS, L. A. et WILBANKS, G. D. (1993). Growth and differentiation of human papillomavirus type 31b positive human cervical cell lines. *Gynecol Oncol*, 49(3):303–310. [12](#)
- de SANJOSE, S., QUINT, W. G., ALEMANY, L., GERAETS, D. T., KLAUSTERMEIER, J. E., LLOVERAS, B., TOUS, S., FELIX, A., BRAVO, L. E., SHIN, H.-R., VALLEJOS, C. S., de RUIZ, P. A., LIMA, M. A., GUIMERA, N., CLAVERO, O., ALEJO, M., LLOMBART-BOSCH, A., CHENG-YANG, C., TATTI, S. A., KASAMATSU, E., ILJAZOVIC, E., ODIDA, M., PRADO, R., SEUD, M., GRCE, M., USUBUTUN, A., JAIN, A., SUAREZ, G. A. H., LOMBARDI, L. E., BANJO, A., MENENDEZ, C., DOMINGO, E. J., VELASCO, J., NESSA, A., CHICHAREON, S. C. B., QIAO, Y. L., LERMA, E., GARLAND, S. M., SASAGAWA, T., FERRERA, A., HAMMOUDA, D., MARIANI, L., PELAYO, A., STEINER, I., OLIVA, E., MEIJER, C. J., AL-JASSAR, W. F., CRUZ, E., WRIGHT, T. C., PURAS, A., LLAVE, C. L., TZARDI, M., AGORASTOS, T., GARCIA-BARRIOLA, V., CLAVEL, C., ORDI, J., ANDUJAR, M., CASTELLSAGUE, X., SANCHEZ, G. I., NOWAKOWSKI, A. M., BORNSTEIN, J., MUNOZ, N. et BOSCH, F. X. (2010). Human papillomavirus genotype attribution in invasive cervical cancer: a retrospective cross-sectional worldwide study. *Lancet Oncol*, 11(11):1048–1056. [15](#), [16](#)

- de VILLIERS, E.-M., FAUQUET, C., BROKER, T. R., BERNARD, H.-U. et zur HAUSEN, H. (2004). Classification of papillomaviruses. *Virology*, 324(1):17–27. [3](#), [4](#), [5](#)
- DEGENHARDT, Y. Y. et SILVERSTEIN, S. (2001a). Interaction of zyxin, a focal adhesion protein, with the e6 protein from human papillomavirus type 6 results in its nuclear translocation. *J Virol*, 75(23):11791–11802. [39](#), [44](#)
- DEGENHARDT, Y. Y. et SILVERSTEIN, S. J. (2001b). Gps2, a protein partner for human papillomavirus e6 proteins. *J Virol*, 75(1):151–160. [39](#), [40](#)
- DELL, G. et GASTON, K. (2001). Human papillomaviruses and their role in cervical cancer. *Cell Mol Life Sci*, 58(12-13):1923–1942. [15](#)
- DIELLA, F., HASLAM, N., CHICA, C., BUDD, A., MICHAEL, S., BROWN, N. P., TRAVE, G. et GIBSON, T. J. (2008). Understanding eukaryotic linear motifs and their role in cell signaling and regulation. *Front Biosci*, 13:6580–6603. [88](#), [102](#)
- DOORBAR, J. (2005). The papillomavirus life cycle. *J Clin Virol*, 32 Suppl 1:S7–15. [11](#)
- DOORBAR, J., CAMPBELL, D., GRAND, R. J. et GALLIMORE, P. H. (1986). Identification of the human papilloma virus-1a e4 gene products. *EMBO J*, 5(2):355–362. [31](#)
- DOORBAR, J., FOO, C., COLEMAN, N., MEDCALF, L., HARTLEY, O., PROSPERO, T., NAPHTINE, S., STERLING, J., WINTER, G. et GRIFFIN, H. (1997). Characterization of events during the late stages of hpv16 infection in vivo using high-affinity synthetic fabs to e4. *Virology*, 238(1):40–52. [14](#)
- DOORBAR, J., PARTON, A., HARTLEY, K., BANKS, L., CROOK, T., STANLEY, M. et CRAWFORD, L. (1990). Detection of novel splicing patterns in a hpv16-containing keratinocyte cell line. *Virology*, 178(1):254–262. [36](#)
- DOUVIER S, D. S. (2004). Infection à papillomavirus. *Encycl Méd Chir (Elsevier, Paris). Gynécologie*, 8-054-A-10. [17](#), [19](#)
- DRUKER, B. J. et LYDON, N. B. (2000). Lessons learned from the development of an abl tyrosine kinase inhibitor for chronic myelogenous leukemia. *J Clin Invest*, 105(1):3–7. [115](#)
- DU, M., FAN, X., HONG, E. et CHEN, J. J. (2002). Interaction of oncogenic papillomavirus e6 proteins with fibulin-1. *Biochem Biophys Res Commun*, 296(4):962–969. [39](#), [44](#)
- DUENSING, S. et MUNGER, K. (2004). Mechanisms of genomic instability in human cancer: insights from studies with human papillomavirus oncoproteins. *Int J Cancer*, 109(2):157–162. [32](#)
- DURST, M., GISSMANN, L., IKENBERG, H. et zur HAUSEN, H. (1983). A papillomavirus dna from a cervical carcinoma and its prevalence in cancer biopsy samples from different geographic regions. *Proc Natl Acad Sci U S A*, 80(12):3812–3815. [4](#)
- DYSON, N., GUIDA, P., MUNGER, K. et HARLOW, E. (1992). Homologous sequences in adenovirus e1a and human papillomavirus e7 proteins mediate interaction with the same set of cellular proteins. *J Virol*, 66(12):6893–6902. [35](#)
- DYSON, N., HOWLEY, P. M., MUNGER, K. et HARLOW, E. (1989). The human papilloma virus-16 e7 oncoprotein is able to bind to the retinoblastoma gene product. *Science*, 243(4893):934–937. [34](#)
- EL MEHDAOUI, S., TOUZE, A., LAURENT, S., SIZARET, P. Y., RASSCHAERT, D. et COURSAGET, P. (2000). Gene transfer using recombinant rabbit hemorrhagic disease virus capsids with genetically modified dna encapsidation capacity by addition of packaging sequences from the 11 or 12 protein of human papillomavirus type 16. *J Virol*, 74(22):10332–10340. [24](#)
- ELBEL, M., CARL, S., SPADERNA, S. et IFTNER, T. (1997). A comparative analysis of the interactions of the e6 proteins from cutaneous and genital papillomaviruses with p53 and e6ap in correlation to their transforming potential. *Virology*, 239(1):132–149. [40](#)
- ENEMARK, E. J., CHEN, G., VAUGHN, D. E., STENLUND, A. et JOSHUA-TOR, L. (2000). Crystal structure of the dna binding domain of the replication initiation protein e1 from papillomavirus. *Mol Cell*, 6(1):149–158. [25](#)

## Références bibliographiques

---

- ENEMARK, E. J., STENLUND, A. et JOSHUA-TOR, L. (2002). Crystal structures of two intermediates in the assembly of the papillomavirus replication initiation complex. *EMBO J*, 21(6):1487–1496. [25](#)
- ENZENAUER, C., MENGUS, G., LAVIGNE, A., DAVIDSON, I., PFISTER, H. et MAY, M. (1998). Interaction of human papillomavirus 8 regulatory proteins e2, e6 and e7 with components of the tfiid complex. *Intervirology*, 41(2-3):80–90. [35](#)
- FAUQUET, C. M. et FARGETTE, D. (2005). International committee on taxonomy of viruses and the 3,142 unassigned species. *Virology*, 2:64. NLM: Original DateCompleted: 20051014. [4](#)
- FAVRE-BONVIN, A., REYNAUD, C., KRETZ-REMY, C. et JALINOT, P. (2005). Human papillomavirus type 18 e6 protein binds the cellular pdz protein tip-2/gipc, which is involved in transforming growth factor beta signaling and triggers its degradation by the proteasome. *J Virol*, 79(7):4229–4237. [13](#), [39](#), [43](#)
- FERLAY J, Bray F, P. P. P. D. (2004). Globocan 2002: Cancer incidence, mortality and prevalence worldwide. *Lyon: IARC Press, editor.*, (5). [15](#)
- FILIPPOVA, M., JOHNSON, M. M., BAUTISTA, M., FILIPPOV, V., FODOR, N., TUNGTEAKKHUN, S. S., WILLIAMS, K. et DUERKSEN-HUGHES, P. J. (2007). The large and small isoforms of human papillomavirus type 16 e6 bind to and differentially affect procaspase 8 stability and activity. *J Virol*, 81(8):4116–4129. [39](#), [40](#)
- FILIPPOVA, M., PARKHURST, L. et DUERKSEN-HUGHES, P. J. (2004). The human papillomavirus 16 e6 protein binds to fas-associated death domain and protects cells from fas-triggered apoptosis. *J Biol Chem*, 279(24):25729–25744. [39](#), [40](#)
- FILIPPOVA, M., SONG, H., CONNOLLY, J. L., DERMODY, T. S. et DUERKSEN-HUGHES, P. J. (2002). The human papillomavirus 16 e6 protein binds to tumor necrosis factor (tnf) r1 and protects cells from tnf-induced apoptosis. *J Biol Chem*, 277(24):21730–21739. [39](#), [40](#)
- FIRZLAFF, J. M., GALLOWAY, D. A., EISENMAN, R. N. et LUSCHER, B. (1989). The e7 protein of human papillomavirus type 16 is phosphorylated by casein kinase ii. *New Biol*, 1(1):44–53. [35](#)
- FLORES, E. R., ALLEN-HOFFMANN, B. L., LEE, D. et LAMBERT, P. F. (2000). The human papillomavirus type 16 e7 oncogene is required for the productive stage of the viral life cycle. *J Virol*, 74(14):6622–6631. [34](#)
- FLORIN, L., BECKER, K. A., LAMBERT, C., NOWAK, T., SAPP, C., STRAND, D., STREECK, R. E. et SAPP, M. (2006). Identification of a dynein interacting domain in the papillomavirus minor capsid protein l2. *J Virol*, 80(13):6691–6696. [24](#)
- FOUTS, E. T., YU, X., EGELMAN, E. H. et BOTCHAN, M. R. (1999). Biochemical and electron microscopic image analysis of the hexameric e1 helicase. *J Biol Chem*, 274(7):4447–4458. [25](#), [114](#)
- FUNK, J. O., WAGA, S., HARRY, J. B., ESPLING, E., STILLMAN, B. et GALLOWAY, D. A. (1997). Inhibition of cdk activity and pcna-dependent dna replication by p21 is blocked by interaction with the hpv-16 e7 oncoprotein. *Genes Dev*, 11(16):2090–2100. [13](#), [34](#), [35](#)
- GAMMOH, N., GRM, H. S., MASSIMI, P. et BANKS, L. (2006). Regulation of human papillomavirus type 16 e7 activity through direct protein interaction with the e2 transcriptional activator. *J Virol*, 80(4):1787–1797. [35](#)
- GAO, Q., KUMAR, A., SINGH, L., HUIBREGTSE, J. M., BEAUDENON, S., SRINIVASAN, S., WAZER, D. E., BAND, H. et BAND, V. (2002). Human papillomavirus e6-induced degradation of e6tp1 is mediated by e6ap ubiquitin ligase. *Cancer Res*, 62(11):3315–3321. [39](#), [41](#)
- GAO, Q., KUMAR, A., SRINIVASAN, S., SINGH, L., MUKAI, H., ONO, Y., WAZER, D. E. et BAND, V. (2000). Pkn binds and phosphorylates human papillomavirus e6 oncoprotein. *J Biol Chem*, 275(20):14824–14830. [39](#), [43](#)
- GAO, Q., SINGH, L., KUMAR, A., SRINIVASAN, S., WAZER, D. E. et BAND, V. (2001). Human papillomavirus type 16 e6-induced degradation of e6tp1 correlates with its ability to immortalize human mammary epithelial cells. *J Virol*, 75(9):4459–4466. [39](#), [41](#)

- GAO, Q., SRINIVASAN, S., BOYER, S. N., WAZER, D. E. et BAND, V. (1999). The e6 oncoproteins of high-risk papillomaviruses bind to a novel putative gap protein, e6tp1, and target it for degradation. *Mol Cell Biol*, 19(1):733–744. [39](#), [41](#)
- GARCIA-ALAI, M. M., GALLO, M., SALAME, M., WETZLER, D. E., MCBRIDE, A. A., PACI, M., CICERO, D. O. et de PRAT-GAY, G. (2006). Molecular basis for phosphorylation-dependent, pest-mediated protein turnover. *Structure*, 14(2):309–319. [28](#)
- GARDIOL, D., KÜHNE, C., GLAUNSINGER, B., LEE, S. S., JAVIER, R. et BANKS, L. (1999). Oncogenic human papillomavirus e6 proteins target the discs large tumour suppressor for proteasome-mediated degradation. *Oncogene*, 18(40):5487–5496. [13](#), [39](#), [42](#)
- GENTHER, S. M., STERLING, S., DUENSING, S., MUNGER, K., SATTLER, C. et LAMBERT, P. F. (2003). Quantitative role of the human papillomavirus type 16 e5 gene during the productive stage of the viral life cycle. *J Virol*, 77(5):2832–2842. [13](#)
- GENTHER WILLIAMS, S. M., DISBROW, G. L., SCHLEGEL, R., LEE, D., THREADGILL, D. W. et LAMBERT, P. F. (2005). Requirement of epidermal growth factor receptor for hyperplasia induced by e5, a high-risk human papillomavirus oncogene. *Cancer Res*, 65(15):6534–6542. [31](#), [32](#)
- GEWIN, L., MYERS, H., KIYONO, T. et GALLOWAY, D. A. (2004). Identification of a novel telomerase repressor that interacts with the human papillomavirus type-16 e6/e6-ap complex. *Genes Dev*, 18(18):2269–2282. [39](#)
- GLAUNSINGER, B. A., LEE, S. S., THOMAS, M., BANKS, L. et JAVIER, R. (2000). Interactions of the pdz-protein magi-1 with adenovirus e4-orf1 and high-risk papillomavirus e6 oncoproteins. *Oncogene*, 19(46):5270–5280. [39](#), [43](#)
- GRASSO, F., NEGRI, D., MOCHI, S., ROSSI, A., CESOLINI, A., GIOVANNELLI, A., CHIANTORE, M., LEONE, P., GIORGI, C. et CARA, A. (2012). Successful therapeutic vaccination with integrase defective lentiviral vector expressing nononcogenic human papillomavirus e7 protein. *Int J Cancer*. [20](#)
- GROFF, D. E. et LANCASTER, W. D. (1985). Molecular cloning and nucleotide sequence of deer papillomavirus. *J Virol*, 56(1):85–91. [3](#)
- GROSS-MESILATY, S., REINSTEIN, E., BERCOVICH, B., TOBIAS, K. E., SCHWARTZ, A. L., KAHANA, C. et CIECHANOVER, A. (1998). Basal and human papillomavirus e6 oncoprotein-induced degradation of myc proteins by the ubiquitin pathway. *Proc Natl Acad Sci U S A*, 95(14):8058–8063. [39](#)
- GROSSMAN, S. R. et LAIMINS, L. A. (1989). E6 protein of human papillomavirus type 18 binds zinc. *Oncogene*, 4(9):1089–1093. [47](#)
- GUCCIONE, E., LETHBRIDGE, K. J., KILICK, N., LEPPARD, K. N. et BANKS, L. (2004). Hpv e6 proteins interact with specific pml isoforms and allow distinctions to be made between different pod structures. *Oncogene*, 23(27):4662–4672. [39](#)
- GUPTA, S., TAKHAR, P. P., DEGENKOLBE, R., KOH, C. H., ZIMMERMANN, H., YANG, C. M., GUAN SIM, K., HSU, S. I. et BERNARD, H. U. (2003). The human papillomavirus type 11 and 16 e6 proteins modulate the cell-cycle regulator and transcription cofactor trip-br1. *Virology*, 317(1):155–164. [39](#), [42](#)
- HABIG, M., SMOLA, H., DOLE, V. S., DERYNCK, R., PFISTER, H. et SMOLA-HESS, S. (2006). E7 proteins from high- and low-risk human papillomaviruses bind to tgfbeta-regulated smad proteins and inhibit their transcriptional activity. *Arch Virol*, 151(10):1961–1972. [35](#), [36](#)
- HAGENSEE, M. E., OLSON, N. H., BAKER, T. S. et GALLOWAY, D. A. (1994). Three-dimensional structure of vaccinia virus-produced human papillomavirus type 1 capsids. *J Virol*, 68(7):4503–4505. [9](#)
- HAMPSON, L., LI, C., OLIVER, A. W., KITCHENER, H. C. et HAMPSON, I. N. (2004). The pdz protein tip-1 is a gain of function target of the hpv16 e6 oncoprotein. *Int J Oncol*, 25(5):1249–1256. [13](#), [39](#), [43](#)
- HAN, R., CLADEL, N. M., REED, C. A. et CHRISTENSEN, N. D. (1998). Characterization of transformation function of cottontail rabbit papillomavirus e5 and e8 genes. *Virology*, 251(2):253–263. [8](#)



## Références bibliographiques

---

- HARRIS, S. F. et BOTCHAN, M. R. (1999). Crystal structure of the human papillomavirus type 18 e2 activation domain. *Science*, 284(5420):1673–1677. [27](#)
- HARRY, J. B. et WETTSTEIN, F. O. (1996). Transforming properties of the cottontail rabbit papillomavirus oncoproteins le6 and se6 and of the e8 protein. *J Virol*, 70(6):3355–3362. [8](#)
- HEATON, P. R., DEYRIEUX, A. F., BIAN, X. L. et WILSON, V. G. (2011). Hpv e6 proteins target ubc9, the sumo conjugating enzyme. *Virus Res*, 158(1-2):199–208. [39](#), [44](#)
- HEBNER, C. M. et LAIMINS, L. A. (2006). Human papillomaviruses: basic mechanisms of pathogenesis and oncogenicity. *Rev Med Virol*, 16(2):83–97. [14](#)
- HEGDE, R. S. (2006). Papillomavirus proteins and their potential as drug design targets. *Future Virology*, 1(6):795–810. [26](#)
- HEGDE, R. S. et ANDROPHY, E. J. (1998). Crystal structure of the e2 dna-binding domain from human papillomavirus type 16: implications for its dna binding-site selection mechanism. *J Mol Biol*, 284(5):1479–1489. [27](#)
- HEGDE, R. S., GROSSMAN, S. R., LAIMINS, L. A. et SIGLER, P. B. (1992). Crystal structure at 1.7 Å of the bovine papillomavirus-1 e2 dna-binding domain bound to its dna target. *Nature*, 359(6395):505–512. [27](#)
- HERMONAT, P. L., SANTIN, A. D. et ZHAN, D. (2000). Binding of the human papillomavirus type 16 e7 oncoprotein and the adeno-associated virus rep78 major regulatory protein in vitro and in yeast and the potential for downstream effects. *J Hum Virol*, 3(3):113–124. [35](#)
- HILDESHEIM, A., MANN, V., BRINTON, L. A., SZKLO, M., REEVES, W. C. et RAWLS, W. E. (1991). Herpes simplex virus type 2: a possible interaction with human papillomavirus types 16/18 in the development of invasive cervical cancer. *Int J Cancer*, 49(3):335–340. [11](#)
- HO, G. Y., BURK, R. D., KLEIN, S., KADISH, A. S., CHANG, C. J., PALAN, P., BASU, J., TACHEZY, R., LEWIS, R. et ROMNEY, S. (1995). Persistent genital human papillomavirus infection as a risk factor for persistent cervical dysplasia. *J Natl Cancer Inst*, 87(18):1365–1371. [17](#)
- HOELLERER, M. K., NOBLE, M. E. M., LABESSE, G., CAMPBELL, I. D., WERNER, J. M. et AROLD, S. T. (2003). Molecular recognition of paxillin Id motifs by the focal adhesion targeting domain. *Structure*, 11(10):1207–1217. [75](#), [102](#), [117](#)
- HOOLEY, E., FAIRWEATHER, V., CLARKE, A. R., GASTON, K. et BRADY, R. L. (2006). The recognition of local dna conformation by the human papillomavirus type 6 e2 protein. *Nucleic Acids Res*, 34(14):3897–3908. [27](#)
- HOWLEY, P. M., MUNGER, K., WERNESS, B. A., PHELPS, W. C. et SCHLEGEL, R. (1989). Molecular mechanisms of transformation by the human papillomaviruses. *Princess Takamatsu Symp*, 20:199–206. [32](#)
- HSU, C. H., PENG, K. L., JHANG, H. C., LIN, C. H., WU, S. Y., CHIANG, C. M., LEE, S. C., YU, W. C. et JUAN, L. J. (2012). The hpv e6 oncoprotein targets histone methyltransferases for modulating specific gene transcription. *Oncogene*, 31(18):2335–2349. [39](#), [44](#)
- HUANG, S.-M. et MCCANCE, D. J. (2002). Down regulation of the interleukin-8 promoter by human papillomavirus type 16 e6 and e7 through effects on creb binding protein/p300 and p/caf. *J Virol*, 76(17):8710–8721. [35](#)
- HUH, K., ZHOU, X., HAYAKAWA, H., CHO, J.-Y., LIBERMANN, T. A., JIN, J., HARPER, J. W. et MUNGER, K. (2007). 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*, 81(18):9737–9747. [35](#)
- HUH, K.-W., DEMASI, J., OGAWA, H., NAKATANI, Y., HOWLEY, P. M. et MUNGER, K. (2005). 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*, 102(32):11492–11497. [35](#)

- HUIBREGTSE, J. M., SCHEFFNER, M. et HOWLEY, P. M. (1991). A cellular protein mediates association of p53 with the e6 oncoprotein of human papillomavirus types 16 or 18. *EMBO J*, 10(13):4129–4135. [38](#)
- HUIBREGTSE, J. M., SCHEFFNER, M. et HOWLEY, P. M. (1993). Localization of the e6-ap regions that direct human papillomavirus e6 binding, association with p53, and ubiquitination of associated proteins. *Mol Cell Biol*, 13(8):4918–4927. [38](#), [39](#), [43](#)
- HUR, E., PFAFF, S. J., PAYNE, E. S., GRON, H., BUEHRER, B. M. et FLETTERICK, R. J. (2004). Recognition and accommodation at the androgen receptor coactivator binding interface. *PLoS Biol*, 2(9):E274. [102](#)
- HWANG, S. G., LEE, D., KIM, J., SEO, T. et CHOE, J. (2002). Human papillomavirus type 16 e7 binds to e2f1 and activates e2f1-driven transcription in a retinoblastoma protein-independent manner. *J Biol Chem*, 277(4):2923–2930. [35](#)
- IFTNER, T., ELBEL, M., SCHOPP, B., HILLER, T., LOIZOU, J. I., CALDECOTT, K. W. et STUBENRAUCH, F. (2002). Interference of papillomavirus e6 protein with single-strand break repair by interaction with xrcc1. *EMBO J*, 21(17):4741–4748. [39](#), [45](#)
- IMAI, Y., TSUNOKAWA, Y., SUGIMURA, T. et TERADA, M. (1989). Purification and dna-binding properties of human papillomavirus type 16 e6 protein expressed in escherichia coli. *Biochem Biophys Res Commun*, 164(3):1402–1410. [47](#)
- JAMES, M. A., LEE, J. H. et KLINGELHUTZ, A. J. (2006). Human papillomavirus type 16 e6 activates nf-kappab, induces ciap-2 expression, and protects against apoptosis in a pdz binding motif-dependent manner. *J Virol*, 80(11):5301–5307. [43](#)
- JEON, J.-H., CHOI, K.-H., CHO, S.-Y., KIM, C.-W., SHIN, D.-M., KWON, J.-C., SONG, K.-Y., PARK, S.-C. et KIM, I.-G. (2003). Transglutaminase 2 inhibits rb binding of human papillomavirus e7 by incorporating polyamine. *EMBO J*, 22(19):5273–5282. [35](#)
- JEONG, K. W., KIM, H. Z., KIM, S., KIM, Y. S. et CHOE, J. (2007). Human papillomavirus type 16 e6 protein interacts with cystic fibrosis transmembrane regulator-associated ligand and promotes e6-associated protein-mediated ubiquitination and proteasomal degradation. *Oncogene*, 26(4):487–499. [39](#), [45](#)
- JHA, S., VANDE POL, S., BANERJEE, N. S., DUTTA, A. B., CHOW, L. T. et DUTTA, A. (2010). Destabilization of tip60 by human papillomavirus e6 results in attenuation of tip60-dependent transcriptional regulation and apoptotic pathway. *Mol Cell*, 38(5):700–711. [41](#)
- JING, M., BOHL, J., BRIMER, N., KINTER, M. et VANDE POL, S. B. (2007). Degradation of tyrosine phosphatase ptpn3 (ptph1) by association with oncogenic human papillomavirus e6 proteins. *J Virol*, 81(5):2231–2239. [13](#), [39](#), [43](#)
- JOH, J., JENSON, A. B., KING, W., PROCTOR, M., INGLE, A., SUNDBERG, J. P. et GHIM, S.-j. (2011). Genomic analysis of the first laboratory-mouse papillomavirus. *J Gen Virol*, 92(Pt 3):692–698. [117](#)
- KAMMER, C., TOMMASINO, M., SYRJANEN, S., DELIUS, H., HEBLING, U., WARTHORST, U., PFISTER, H. et ZEHBE, I. (2002). Variants of the long control region and the e6 oncogene in european human papillomavirus type 16 isolates: implications for cervical disease. *Br J Cancer*, 86(2):269–273. [4](#)
- KANAYA, T., KYO, S. et LAIMINS, L. A. (1997). 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*, 237(1):159–169. [9](#)
- KAPUST, R. B. et WAUGH, D. S. (1999). Escherichia coli maltose-binding protein is uncommonly effective at promoting the solubility of polypeptides to which it is fused. *Protein Sci*, 8(8):1668–1674. [116](#)
- KAYE, J. N., STARKEY, W. G., KELL, B., BISWAS, C., RAJU, K. S., BEST, J. M. et CASON, J. (1996). Human papillomavirus type 16 in infants: use of dna sequence analyses to determine the source of infection. *J Gen Virol*, 77 ( Pt 6):1139–1143. [11](#)
- KAZEMI, S., PAPADOPOULOU, S., LI, S., SU, Q., WANG, S., YOSHIMURA, A., MATLASHEWSKI, G., DEVER, T. E. et KOROMILAS, A. E. (2004). Control of alpha subunit of eukaryotic translation initiation factor 2 (eif2 alpha) phosphorylation by the human papillomavirus type 18 e6 oncoprotein: implications for eif2 alpha-dependent gene expression and cell death. *Mol Cell Biol*, 24(8):3415–3429. [39](#), [42](#)

## Références bibliographiques

---

- KIRNBAUER, R., BOOY, F., CHENG, N., LOWY, D. R. et SCHILLER, J. T. (1992). Papillomavirus l1 major capsid protein self-assembles into virus-like particles that are highly immunogenic. *Proc Natl Acad Sci U S A*, 89(24):12180–12184. [9](#)
- KISSELJOV, F. L. (2000). Virus-associated human tumors: cervical carcinomas and papilloma viruses. *Biochemistry (Mosc)*, 65(1):68–77. [9](#)
- KIYONO, T., FOSTER, S. A., KOOP, J. I., MCDUGALL, J. K., GALLOWAY, D. A. et KLINGELHUTZ, A. J. (1998). Both rb/p16ink4a inactivation and telomerase activity are required to immortalize human epithelial cells. *Nature*, 396(6706):84–88. [41](#)
- KLINGELHUTZ, A. J., FOSTER, S. A. et MCDUGALL, J. K. (1996). Telomerase activation by the e6 gene product of human papillomavirus type 16. *Nature*, 380(6569):79–82. [41](#)
- KLINGELHUTZ, A. J. et ROMAN, A. (2012). Cellular transformation by human papillomaviruses: lessons learned by comparing high- and low-risk viruses. *Virology*, 424(2):77–98. [32](#), [102](#)
- KNIGHT, G. L., GRAINGER, J. R., GALLIMORE, P. H. et ROBERTS, S. (2004). Cooperation between different forms of the human papillomavirus type 1 e4 protein to block cell cycle progression and cellular dna synthesis. *J Virol*, 78(24):13920–13933. [31](#)
- KOBE, B., CENTER, R. J., KEMP, B. E. et POUIMBOURIOS, P. (1999). Crystal structure of human t cell leukemia virus type 1 gp21 ectodomain crystallized as a maltose-binding protein chimera reveals structural evolution of retroviral transmembrane proteins. *Proc Natl Acad Sci U S A*, 96(8):4319–4324. [125](#)
- KOMOREK, J., KUPPUSWAMY, M., SUBRAMANIAN, T., VIJAYALINGAM, S., LOMONOSOVA, E., ZHAO, L. J., MYMRYK, J. S., SCHMITT, K. et CHINNADURAI, G. (2010). Adenovirus type 5 e1a and e6 proteins of low-risk cutaneous beta-human papillomaviruses suppress cell transformation through interaction with foxk1/k2 transcription factors. *J Virol*, 84(6):2719–2731. [39](#)
- KÜHNE, C., GARDIOL, D., GUARNACCIA, C., AMENITSCH, H. et BANKS, L. (2000). Differential regulation of human papillomavirus e6 by protein kinase a: conditional degradation of human discs large protein by oncogenic e6. *Oncogene*, 19(51):5884–5891. [39](#), [43](#)
- KUKIMOTO, I., AIHARA, S., YOSHIKE, K. et KANDA, T. (1998). Human papillomavirus oncoprotein e6 binds to the c-terminal region of human minichromosome maintenance 7 protein. *Biochem Biophys Res Commun*, 249(1):258–262. [39](#), [44](#)
- KUMAR, A., ZHAO, Y., MENG, G., ZENG, M., SRINIVASAN, S., DELMOLINO, L. M., GAO, Q., DIMRI, G., WEBER, G. F., WAZER, D. E., BAND, H. et BAND, V. (2002). Human papillomavirus oncoprotein e6 inactivates the transcriptional coactivator human ada3. *Mol Cell Biol*, 22(16):5801–5812. [41](#)
- KURG, R., TEKKEL, H., ABROI, A. et USTAV, M. (2006). Characterization of the functional activities of the bovine papillomavirus type 1 e2 protein single-chain heterodimers. *J Virol*, 80(22):11218–11225. [27](#), [105](#)
- KYTE, J. et DOOLITTLE, R. F. (1982). A simple method for displaying the hydropathic character of a protein. *J Mol Biol*, 157(1):105–132. [97](#)
- LAMBERT, P. F., SPALHOLZ, B. A. et HOWLEY, P. M. (1987). A transcriptional repressor encoded by bpv-1 shares a common carboxy-terminal domain with the e2 transactivator. *Cell*, 50(1):69–78. [8](#)
- LE ROUX, L. G. et MOROIANU, J. (2003). Nuclear entry of high-risk human papillomavirus type 16 e6 oncoprotein occurs via several pathways. *J Virol*, 77(4):2330–2337. [39](#), [45](#)
- LECHNER, M. S. et LAIMINS, L. A. (1994). Inhibition of p53 dna binding by human papillomavirus e6 proteins. *J Virol*, 68(7):4262–4273. [38](#)
- LEE, D. K., KIM, B.-C., KIM, I. Y., CHO, E.-a., SATTERWHITE, D. J. et KIM, S.-J. (2002). The human papilloma virus e7 oncoprotein inhibits transforming growth factor-beta signaling by blocking binding of the smad complex to its target sequence. *J Biol Chem*, 277(41):38557–38564. [35](#), [36](#)

- LEE, S. S., GLAUNSINGER, B., MANTOVANI, F., BANKS, L. et JAVIER, R. T. (2000). Multi-pdz domain protein mupp1 is a cellular target for both adenovirus e4-orf1 and high-risk papillomavirus type 18 e6 oncoproteins. *J Virol*, 74(20):9680–9693. [39](#), [42](#)
- LEHR, E., HOHL, D., HUBER, M. et BROWN, D. (2004). Infection with human papillomavirus alters expression of the small proline rich proteins 2 and 3. *J Med Virol*, 72(3):478–483. [14](#), [31](#)
- LI, B. et DOU, Q. P. (2000). Bax degradation by the ubiquitin/proteasome-dependent pathway: involvement in tumor survival and progression. *Proc Natl Acad Sci U S A*, 97(8):3850–3855. [13](#)
- LI, M., BEARD, P., ESTES, P. A., LYON, M. K. et GARCEA, R. L. (1998). Intercapsomeric disulfide bonds in papillomavirus assembly and disassembly. *J Virol*, 72(3):2160–2167. [11](#)
- LI, S., LABRECQUE, S., GAUZZI, M. C., CUDDIHY, A. R., WONG, A. H., PELLEGRINI, S., MATLASHEWSKI, G. J. et KOROMILAS, A. E. (1999). The human papilloma virus (hpv)-18 e6 oncoprotein physically associates with tyk2 and impairs jak-stat activation by interferon-alpha. *Oncogene*, 18(42):5727–5737. [39](#), [43](#)
- LI, S., LIU, P., XI, L., JIANG, X., WU, M., DENG, D., WEI, J., ZHU, T., ZHOU, L., WANG, S., XU, G., MENG, L., ZHOU, J. et MA, D. (2008). Expression of tmem87b interacting with the human papillomavirus type 18 e6 oncogene in the hela cDNA library by a yeast two-hybrid system. *Oncol Rep*, 20(2):421–427. [39](#), [43](#)
- LI, X. et COFFINO, P. (1996). High-risk human papillomavirus e6 protein has two distinct binding sites within p53, of which only one determines degradation. *J Virol*, 70(7):4509–4516. [38](#)
- LIANG, H., PETROS, A. M., MEADOWS, R. P., YOON, H. S., EGAN, D. A., WALTER, K., HOLZMAN, T. F., ROBINS, T. et FESIK, S. W. (1996). Solution structure of the dna-binding domain of a human papillomavirus e2 protein: evidence for flexible dna-binding regions. *Biochemistry*, 35(7):2095–2103. [27](#)
- LIN, B. Y., MAKHOV, A. M., GRIFFITH, J. D., BROKER, T. R. et CHOW, L. T. (2002). Chaperone proteins abrogate inhibition of the human papillomavirus (hpv) e1 replicative helicase by the hpv e2 protein. *Mol Cell Biol*, 22(18):6592–6604. [25](#)
- LIU, J. S., KUO, S. R., MAKHOV, A. M., CYR, D. M., GRIFFITH, J. D., BROKER, T. R. et CHOW, L. T. (1998). 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*, 273(46):30704–30712. [14](#)
- LIU, X., CLEMENTS, A., ZHAO, K. et MARMORSTEIN, R. (2006). Structure of the human papillomavirus e7 oncoprotein and its mechanism for inactivation of the retinoblastoma tumor suppressor. *J Biol Chem*, 281(1):578–586. [32](#), [33](#)
- LIU, X., DAKIC, A., ZHANG, Y., DAI, Y., CHEN, R. et SCHLEGEL, R. (2009). Hpv e6 protein interacts physically and functionally with the cellular telomerase complex. *Proc Natl Acad Sci U S A*, 106(44):18780–18785. [41](#)
- LIU, X., ROBERTS, J., DAKIC, A., ZHANG, Y. et SCHLEGEL, R. (2008). Hpv e7 contributes to the telomerase activity of immortalized and tumorigenic cells and augments e6-induced htert promoter function. *Virology*, 375(2):611–623. [32](#)
- LONGWORTH, M. S. et LAIMINS, L. A. (2004a). The binding of histone deacetylases and the integrity of zinc finger-like motifs of the e7 protein are essential for the life cycle of human papillomavirus type 31. *J Virol*, 78(7):3533–3541. [35](#)
- LONGWORTH, M. S. et LAIMINS, L. A. (2004b). Pathogenesis of human papillomaviruses in differentiating epithelia. *Microbiol Mol Biol Rev*, 68(2):362–372. [13](#), [31](#)
- LORENZ, S., VAKONAKIS, I., LOWE, E. D., CAMPBELL, I. D., NOBLE, M. E. M. et HOELLERER, M. K. (2008). Structural analysis of the interactions between paxillin ld motifs and alpha-parvin. *Structure*, 16(10):1521–1531. [75](#), [102](#), [117](#)
- LU, Z., HU, X., LI, Y., ZHENG, L., ZHOU, Y., JIANG, H., NING, T., BASANG, Z., ZHANG, C. et KE, Y. (2004). Human papillomavirus 16 e6 oncoprotein interferes with insulin signaling pathway by binding to tuberlin. *J Biol Chem*, 279(34):35664–35670. [39](#), [41](#)

## Références bibliographiques

---

- LUSCHER-FIRZLAFF, J. M., WESTENDORF, J. M., ZWICKER, J., BURKHARDT, H., HENRIKSSON, M., MULLER, R., PIROLLET, F. et LUSCHER, B. (1999). Interaction of the fork head domain transcription factor mpp2 with the human papilloma virus 16 e7 protein: enhancement of transformation and transactivation. *Oncogene*, 18(41):5620–5630. [35](#)
- MADISON, K. C. (2003). Barrier function of the skin: "la raison d'être" of the epidermis. *J Invest Dermatol*, 121(2):231–241. [13](#)
- MALILA, L. H., SHIVARI, K. M., MAKELA, M. J., JALONEN, J. E., LATIPAA, P. M., KUNAU, W. H. et HILTUNEN, J. K. (1993). Enzymes converting d-3-hydroxyacyl-coa to trans-2-enoyl-coa. microsomal and peroxisomal isoenzymes in rat liver. *J Biol Chem*, 268(29):21578–21585. [35](#)
- MANNHARDT, B., WEINZIMER, S. A., WAGNER, M., FIEDLER, M., COHEN, P., JANSEN-DURR, P. et ZWERSCHKE, W. (2000). Human papillomavirus type 16 e7 oncoprotein binds and inactivates growth-inhibitory insulin-like growth factor binding protein 3. *Mol Cell Biol*, 20(17):6483–6495. [35](#)
- MANTOVANI, F. et BANKS, L. (2001). The human papillomavirus e6 protein and its contribution to malignant progression. *Oncogene*, 20(54):7874–7887. [36](#)
- MARDH, P. A., TCHOUDOMIROVA, K., ELSHIBLY, S. et HELLBERG, D. (1998). Symptoms and signs in single and mixed genital infections. *Int J Gynaecol Obstet*, 63(2):145–152. [11](#)
- MASSIMI, P., GAMMOH, N., THOMAS, M. et BANKS, L. (2004). Hpv e6 specifically targets different cellular pools of its pdz domain-containing tumour suppressor substrates for proteasome-mediated degradation. *Oncogene*, 23(49):8033–8039. [39](#), [43](#)
- MASSIMI, P., PIM, D. et BANKS, L. (1997). 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*, 78 ( Pt 10):2607–2613. [35](#)
- MASSIMI, P., PIM, D., STOREY, A. et BANKS, L. (1996). Hpv-16 e7 and adenovirus e1a complex formation with tata box binding protein is enhanced by casein kinase ii phosphorylation. *Oncogene*, 12(11):2325–2330. [35](#)
- MAUFORT, J. P., SHAI, A., PITOT, H. C. et LAMBERT, P. F. (2010). A role for hpv16 e5 in cervical carcinogenesis. *Cancer Res*, 70(7):2924–2931. [31](#)
- MAZUREK, S., ZWERSCHKE, W., JANSEN-DURR, P. et EIGENBRODT, E. (2001). Metabolic cooperation between different oncogenes during cell transformation: interaction between activated ras and hpv-16 e7. *Oncogene*, 20(47):6891–6898. [35](#)
- MAZZARELLI, J. M., ATKINS, G. B., GEISBERG, J. V. et RICCIARDI, R. P. (1995). The viral oncoproteins ad5 e1a, hpv16 e7 and sv40 tag bind a common region of the tbp-associated factor-110. *Oncogene*, 11(9):1859–1864. [35](#)
- MCLAUGHLIN-DRUBIN, M. E., HUH, K.-W. et MUNGER, K. (2008). Human papillomavirus type 16 e7 oncoprotein associates with e2f6. *J Virol*, 82(17):8695–8705. [34](#), [35](#)
- MCLAUGHLIN-DRUBIN, M. E. et MUNGER, K. (2009). Oncogenic activities of human papillomaviruses. *Virus Res*, 143(2):195–208. [15](#)
- MCMURRAY, H. R., NGUYEN, D., WESTBROOK, T. F. et MCANCE, D. J. (2001). Biology of human papillomaviruses. *Int J Exp Pathol*, 82(1):15–33. [5](#)
- MELILLO, R. M., HELIN, K., LOWY, D. R. et SCHILLER, J. T. (1994). Positive and negative regulation of cell proliferation by e2f-1: influence of protein level and human papillomavirus oncoproteins. *Mol Cell Biol*, 14(12):8241–8249. [34](#)
- MENDOZA, J. A., JACOB, Y., CASSONNET, P. et FAVRE, M. (2006). Human papillomavirus type 5 e6 oncoprotein represses the transforming growth factor beta signaling pathway by binding to smad3. *J Virol*, 80(24):12420–12424. [39](#), [43](#)
- MODIS, Y., TRUS, B. L. et HARRISON, S. C. (2002). Atomic model of the papillomavirus capsid. *EMBO J*, 21(18):4754–4762. [9](#), [10](#)

- MOODY, C. A. et LAIMINS, L. A. (2010). Human papillomavirus oncoproteins: pathways to transformation. *Nat Rev Cancer*, 10(8):550–560. [12](#)
- MORANDELL, D., KAISER, A., HEROLD, S., ROSTEK, U., LECHNER, S., MITTERBERGER, M. C., JANSEN-DURR, P., EILERS, M. et ZWERSCHKE, W. (2012). The human papillomavirus type 16 e7 oncoprotein targets myc-interacting zinc-finger protein-1. *Virology*, 422(2):242–253. [35](#), [36](#)
- MORRIS, J. D., CROOK, T., BANDARA, L. R., DAVIES, R., LATHANGUE, N. B. et VOUSDEN, K. H. (1993a). Human papillomavirus type 16 e7 regulates e2f and contributes to mitogenic signalling. *Oncogene*, 8(4):893–898. [34](#)
- MORRIS, P. J., DENT, C. L., RING, C. J. et LATCHMAN, D. S. (1993b). The octamer binding site in the hpv16 regulatory region produces opposite effects on gene expression in cervical and non-cervical cells. *Nucleic Acids Res*, 21(4):1019–1023. [9](#)
- MUNGER, K., BASILE, J. R., DUENSING, S., EICHTEN, A., GONZALEZ, S. L., GRACE, M. et ZACNY, V. L. (2001). Biological activities and molecular targets of the human papillomavirus e7 oncoprotein. *Oncogene*, 20(54):7888–7898. [13](#), [34](#)
- MUNOZ, N., CASTELLSAGUE, X., de GONZALEZ, A. B. et GISSMANN, L. (2006). Chapter 1: Hpv in the etiology of human cancer. *Vaccine*, 24 Suppl 3:S3/1–10. [5](#), [15](#), [17](#)
- MUSCHIK, D., BRASPENNING-WESCH, I., STOCKFLETH, E., RÖSL, F., HOFMANN, T. G. et NINDL, I. (2011). Cutaneous hpv23 e6 prevents p53 phosphorylation through interaction with hipp2. *PLoS One*, 6(11). [39](#), [41](#)
- NAKAGAWA, S. et HUIBREGTSE, J. M. (2000). 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*, 20(21):8244–8253. [13](#), [39](#), [42](#)
- NAKAHARA, T., PEH, W. L., DOORBAR, J., LEE, D. et LAMBERT, P. F. (2005). Human papillomavirus type 16 e1circumflexe4 contributes to multiple facets of the papillomavirus life cycle. *J Virol*, 79(20):13150–13165. [31](#)
- NEARY, K., HORWITZ, B. H. et DiMAIO, D. (1987). Mutational analysis of open reading frame e4 of bovine papillomavirus type 1. *J Virol*, 61(4):1248–1252. [31](#)
- NGUYEN, C. L., EICHWALD, C., NIBERT, M. L. et MUNGER, K. (2007). Human papillomavirus type 16 e7 oncoprotein associates with the centrosomal component gamma-tubulin. *J Virol*, 81(24):13533–13543. [35](#)
- NGUYEN, C. L. et MUNGER, K. (2008). Direct association of the hpv16 e7 oncoprotein with cyclin a/cdk2 and cyclin e/cdk2 complexes. *Virology*, 380(1):21–25. [35](#)
- NGUYEN, C. L. et MUNGER, K. (2009). Human papillomavirus e7 protein deregulates mitosis via an association with nuclear mitotic apparatus protein 1. *J Virol*, 83(4):1700–1707. [35](#)
- NISHIMURA, A., ONO, T., ISHIMOTO, A., DOWHANICK, J. J., FRIZZELL, M. A., HOWLEY, P. M. et SAKAI, H. (2000). Mechanisms of human papillomavirus e2-mediated repression of viral oncogene expression and cervical cancer cell growth inhibition. *J Virol*, 74(8):3752–3760. [27](#)
- NOBBENHUIS, M. A., WALBOOMERS, J. M., HELMERHORST, T. J., ROZENDAAL, L., REMMINK, A. J., RISSE, E. K., van der LINDEN, H. C., VOORHORST, F. J., KENEMANS, P. et MEIJER, C. J. (1999). Relation of human papillomavirus status to cervical lesions and consequences for cervical-cancer screening: a prospective study. *Lancet*, 354(9172):20–25. [17](#)
- NOMINE, Y., CHARBONNIER, S., RISTRANI, T., STIER, G., MASSON, M., CAVUSOGLU, N., VAN DORSSELAER, A., WEISS, E., KIEFFER, B. et TRAVE, G. (2003). Domain substructure of hpv e6 oncoprotein: biophysical characterization of the e6 c-terminal dna-binding domain. *Biochemistry*, 42(17):4909–4917. [48](#)
- NOMINE, Y., MASSON, M., CHARBONNIER, S., ZANIER, K., RISTRANI, T., DERYCKERE, F., SIBLER, A.-P., DESPLANCQ, D., ATKINSON, R. A., WEISS, E., ORFANOUDAKIS, G., KIEFFER, B. et TRAVE, G. (2006). Structural and functional analysis of e6 oncoprotein: insights in the molecular pathways of human papillomavirus-mediated pathogenesis. *Mol Cell*, 21(5):665–678. [36](#), [38](#), [39](#), [45](#), [47](#), [48](#), [49](#), [71](#), [103](#), [117](#)

## Références bibliographiques

---

- NOMINE, Y., RISTRANI, T., LAURENT, C., LEFEVRE, J. F. et TRAVE, G. (2001). Formation of soluble inclusion bodies by hpv e6 oncoprotein fused to maltose-binding protein. *Protein Expr Purif*, 23(1):22–32. [48](#)
- OH, S. T., LONGWORTH, M. S. et LAIMINS, L. A. (2004). Roles of the e6 and e7 proteins in the life cycle of low-risk human papillomavirus type 11. *J Virol*, 78(5):2620–2626. [11](#), [38](#)
- ORTH, G., FAVRE, M. et CROISSANT, O. (1977). Characterization of a new type of human papillomavirus that causes skin warts. *J Virol*, 24(1):108–120. [3](#)
- ORTH, G., JABLONSKA, S., FAVRE, M., CROISSANT, O., JARZABEK-CHORZELSKA, M. et RZESA, G. (1978). Characterization of two types of human papillomaviruses in lesions of epidermodysplasia verruciformis. *Proc Natl Acad Sci U S A*, 75(3):1537–1541. [3](#)
- ORTH, G., JABLONSKA, S., JARZABEK-CHORZELSKA, M., OBALEK, S., RZESA, G., FAVRE, M. et CROISSANT, O. (1979). Characteristics of the lesions and risk of malignant conversion associated with the type of human papillomavirus involved in epidermodysplasia verruciformis. *Cancer Res*, 39(3):1074–1082. [3](#)
- OZBUN, M. A. et MEYERS, C. (1998). Human papillomavirus type 31b e1 and e2 transcript expression correlates with vegetative viral genome amplification. *Virology*, 248(2):218–230. [14](#)
- PAN, H. et GRIEP, A. E. (1995). Temporally distinct patterns of p53-dependent and p53-independent apoptosis during mouse lens development. *Genes Dev*, 9(17):2157–2169. [38](#)
- PARK, J. S., KIM, E. J., KWON, H. J., HWANG, E. S., NAMKOONG, S. E. et UM, S. J. (2000). 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*, 275(10):6764–6769. [35](#), [36](#)
- PARK, R. B. et ANDROPHY, E. J. (2002). Genetic analysis of high-risk e6 in episomal maintenance of human papillomavirus genomes in primary human keratinocytes. *J Virol*, 76(22):11359–11364. [38](#)
- PATEL, D., HUANG, S. M., BAGLIA, L. A. et MCCANCE, D. J. (1999). The e6 protein of human papillomavirus type 16 binds to and inhibits co-activation by cbp and p300. *EMBO J*, 18(18):5061–5072. [39](#)
- PEH, W. L., BRANDSMA, J. L., CHRISTENSEN, N. D., CLADEL, N. M., WU, X. et DOORBAR, J. (2004). The viral e4 protein is required for the completion of the cottontail rabbit papillomavirus productive cycle in vivo. *J Virol*, 78(4):2142–2151. [13](#)
- PENG, Y. C., BREIDING, D. E., SVERDRUP, F., RICHARD, J. et ANDROPHY, E. J. (2000). Amf-1/gps2 binds p300 and enhances its interaction with papillomavirus e2 proteins. *J Virol*, 74(13):5872–5879. [40](#)
- PENROSE, K. J., GARCIA-ALAI, M., de PRAT-GAY, G. et MCBRIDE, A. A. (2004). Casein kinase ii phosphorylation-induced conformational switch triggers degradation of the papillomavirus e2 protein. *J Biol Chem*, 279(21):22430–22439. [28](#)
- PHELPS, W. C., MUNGER, K., YEE, C. L., BARNES, J. A. et HOWLEY, P. M. (1992). Structure-function analysis of the human papillomavirus type 16 e7 oncoprotein. *J Virol*, 66(4):2418–2427. [34](#)
- PIM, D. et BANKS, L. (2010). Interaction of viral oncoproteins with cellular target molecules: infection with high-risk vs low-risk human papillomaviruses. *APMIS*, 118(6-7):471–493. [38](#)
- PIM, D., MASSIMI, P., DILWORTH, S. M. et BANKS, L. (2005). Activation of the protein kinase b pathway by the hpv-16 e7 oncoprotein occurs through a mechanism involving interaction with pp2a. *Oncogene*, 24(53):7830–7838. [35](#)
- PRASAD, S., RAVINDRAN, J. et AGGARWAL, B. B. (2010). Nf-kappab and cancer: how intimate is this relationship. *Mol Cell Biochem*, 336(1-2):25–37. [43](#)
- PRATHAPAM, T., KUHNE, C. et BANKS, L. (2001). The hpv-16 e7 oncoprotein binds skip and suppresses its transcriptional activity. *Oncogene*, 20(52):7677–7685. [35](#)

- RAYMENT, I., BAKER, T. S., CASPAR, D. L. et MURAKAMI, W. T. (1982). Polyoma virus capsid structure at 22.5 Å resolution. *Nature*, 295(5845):110–115. [10](#)
- RISTRIANI, T., NOMINE, Y., MASSON, M., WEISS, E. et TRAVE, G. (2001). Specific recognition of four-way dna junctions by the c-terminal zinc-binding domain of hpv oncoprotein e6. *J Mol Biol*, 305(4):729–739. [39](#), [45](#)
- RODEN, R. B., LOWY, D. R. et SCHILLER, J. T. (1997). Papillomavirus is resistant to desiccation. *J Infect Dis*, 176(4):1076–1079. [14](#)
- ROMANCZUK, H. et HOWLEY, P. M. (1992). 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*, 89(7):3159–3163. [14](#), [31](#)
- ROMMEL, O., DILLNER, J., FLIGGE, C., BERGSDORF, C., WANG, X., SELINKA, H.-C. et SAPP, M. (2005). Heparan sulfate proteoglycans interact exclusively with conformationally intact hpv 11 assemblies: basis for a virus-like particle elisa. *J Med Virol*, 75(1):114–121. [24](#)
- RONCO, L. V., KARPOVA, A. Y., VIDAL, M. et HOWLEY, P. M. (1998). Human papillomavirus 16 e6 oncoprotein binds to interferon regulatory factor-3 and inhibits its transcriptional activity. *Genes Dev*, 12(13):2061–2072. [39](#), [43](#)
- SANTER, F. R., MOSER, B., SPODEN, G. A., JANSEN-DURR, P. et ZWERSCHKE, W. (2007). Human papillomavirus type 16 e7 oncoprotein inhibits apoptosis mediated by nuclear insulin-like growth factor-binding protein-3 by enhancing its ubiquitin/proteasome-dependent degradation. *Carcinogenesis*, 28(12):2511–2520. [35](#)
- SAVKUR, R. S. et BURRIS, T. P. (2004). The coactivator lxxll nuclear receptor recognition motif. *J Pept Res*, 63(3):207–212. [102](#)
- SCHIEFFNER, M., HUIBREGTSE, J. M., VIERSTRA, R. D. et 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(3):495–505. [38](#)
- SCHIEFFNER, M., WERNES, B. A., HUIBREGTSE, J. M., LEVINE, A. J. et HOWLEY, P. M. (1990). The e6 oncoprotein encoded by human papillomavirus types 16 and 18 promotes the degradation of p53. *Cell*, 63(6):1129–1136. [38](#)
- SCHILLING, B., DE-MEDINA, T., SYKEN, J., VIDAL, M. et MUNGER, K. (1998). A novel human dnaj protein, htid-1, a homolog of the drosophila tumor suppressor protein tid56, can interact with the human papillomavirus type 16 e7 oncoprotein. *Virology*, 247(1):74–85. [35](#)
- SCHLEGEL, R., WADE-GLASS, M., RABSON, M. S. et YANG, Y. C. (1986). The e5 transforming gene of bovine papillomavirus encodes a small, hydrophobic polypeptide. *Science*, 233(4762):464–467. [31](#)
- SCHUCK, S. et STENLUND, A. (2005). Role of papillomavirus e1 initiator dimerization in dna replication. *J Virol*, 79(13):8661–8664. [25](#)
- SCHWARZ, E., FREESE, U. K., GISSMANN, L., MAYER, W., ROGGENBUCK, B., STREMLAU, A. et zur HAUSEN, H. (1985). Structure and transcription of human papillomavirus sequences in cervical carcinoma cells. *Nature*, 314(6006):111–114. [31](#)
- SELINKA, H.-C., GIROGLOU, T. et SAPP, M. (2002). Analysis of the infectious entry pathway of human papillomavirus type 33 pseudovirions. *Virology*, 299(2):279–287. [11](#)
- SEO, Y. S., MULLER, F., LUSKY, M., GIBBS, E., KIM, H. Y., PHILLIPS, B. et HURWITZ, J. (1993). Bovine papilloma virus (bpv)-encoded e2 protein enhances binding of e1 protein to the bpv replication origin. *Proc Natl Acad Sci U S A*, 90(7):2865–2869. [25](#)
- SEVERINO, A., ABBRUZZESE, C., MANENTE, L., VALDERAS, A. A., MATTAROCCHI, S., FEDERICO, A., STARACE, G., CHERSI, A., MILEO, A. M. et PAGGI, M. G. (2007). Human papillomavirus-16 e7 interacts with siva-1 and modulates apoptosis in hacat human immortalized keratinocytes. *J Cell Physiol*, 212(1):118–125. [35](#)



## Références bibliographiques

---

- SHAFTI-KERAMAT, S., HANDISURYA, A., KRIEHLER, E., MENEGUZZI, G., SLUPETZKY, K. et KIRNBAUER, R. (2003). Different heparan sulfate proteoglycans serve as cellular receptors for human papillomaviruses. *J Virol*, 77(24):13125–13135. [24](#)
- SHERMAN, L., JACKMAN, A., ITZHAKI, H., STOPPLER, M. C., KOVAL, D. et SCHLEGEL, R. (1997). Inhibition of serum- and calcium-induced differentiation of human keratinocytes by hpv16 e6 oncoprotein: role of p53 inactivation. *Virology*, 237(2):296–306. [13](#)
- SIDI, A. O. M. O., BABAH, K. O., BRIMER, N., NOMINE, Y., ROMIER, C., KIEFFER, B., POL, S. V., TRAVE, G. et ZANIER, K. (2011). Strategies for bacterial expression of protein-peptide complexes: application to solubilization of papillomavirus e6. *Protein Expr Purif*, 80(1):8–16. [77](#), [87](#), [102](#), [103](#)
- SMITH, J. S., MUNOZ, N., HERRERO, R., ELUF-NETO, J., NGELANGEL, C., FRANCESCHI, S., BOSCH, F. X., WALBOOMERS, J. M. M. et PEELING, R. W. (2002). Evidence for chlamydia trachomatis as a human papillomavirus cofactor in the etiology of invasive cervical cancer in brazil and the philippines. *J Infect Dis*, 185(3):324–331. [11](#)
- SOLIVAN, G. A., SMITH, K. J. et JAMES, W. D. (1990). Cutaneous horn of the penis: its association with squamous cell carcinoma and hpv-16 infection. *J Am Acad Dermatol*, 23(5 Pt 2):969–972. [17](#)
- SPARKOWSKI, J., MENSE, M., ANDERS, J. et SCHLEGEL, R. (1996). E5 oncoprotein transmembrane mutants dissociate fibroblast transforming activity from 16-kilodalton protein binding and platelet-derived growth factor receptor binding and phosphorylation. *J Virol*, 70(4):2420–2430. [31](#)
- SPITKOVSKY, D., HEHNER, S. P., HOFMANN, T. G., MOLLER, A. et SCHMITZ, M. L. (2002). The human papillomavirus oncoprotein e7 attenuates nf-kappa b activation by targeting the ikappa b kinase complex. *J Biol Chem*, 277(28):25576–25582. [35](#)
- SRIVENUGOPAL, K. S. et ALI-OSMAN, F. (2002). The dna repair protein, o(6)-methylguanine-dna methyltransferase is a proteolytic target for the e6 human papillomavirus oncoprotein. *Oncogene*, 21(38):5940–5945. [39](#), [45](#)
- STACEY, S. N., JORDAN, D., WILLIAMSON, A. J., BROWN, M., COOTE, J. H. et ARRAND, J. R. (2000). Leaky scanning is the predominant mechanism for translation of human papillomavirus type 16 e7 oncoprotein from e6/e7 bicistronic mrna. *J Virol*, 74(16):7284–7297. [13](#)
- STANLEY, M. A., BROWNE, H. M., APPLEBY, M. et MINSON, A. C. (1989). Properties of a non-tumorigenic human cervical keratinocyte cell line. *Int J Cancer*, 43(4):672–676. [12](#)
- STAUFFER, Y., RAJ, K., MASTERNAK, K. et BEARD, P. (1998). Infectious human papillomavirus type 18 pseudovirions. *J Mol Biol*, 283(3):529–536. [14](#)
- STEGER, G. et CORBACH, S. (1997). Dose-dependent regulation of the early promoter of human papillomavirus type 18 by the viral e2 protein. *J Virol*, 71(1):50–58. [27](#)
- STORRS, C. H. et SILVERSTEIN, S. J. (2007). Patj, a tight junction-associated pdz protein, is a novel degradation target of high-risk human papillomavirus e6 and the alternatively spliced isoform 18 e6. *J Virol*, 81(8):4080–4090. [39](#), [42](#)
- STRAIGHT, S. W., HINKLE, P. M., JEWERS, R. J. et MCCANCE, D. J. (1993). The e5 oncoprotein of human papillomavirus type 16 transforms fibroblasts and effects the downregulation of the epidermal growth factor receptor in keratinocytes. *J Virol*, 67(8):4521–4532. [32](#)
- STRAUSS, M. J., BUNTING, H. et MELNICK, J. L. (1950). Virus-like particles and inclusion bodies in skin papillomas. *J Invest Dermatol*, 15(6):433–444. [3](#)
- STUBENRAUCH, F., HUMMEL, M., IFTNER, T. et LAIMINS, L. A. (2000). The e8e2c protein, a negative regulator of viral transcription and replication, is required for extrachromosomal maintenance of human papillomavirus type 31 in keratinocytes. *J Virol*, 74(3):1178–1186. [8](#)

- STUBENRAUCH, F., ZOBEL, T. et IFTNER, T. (2001). The e8 domain confers a novel long-distance transcriptional repression activity on the e8e2c protein of high-risk human papillomavirus type 31. *J Virol*, 75(9):4139–4149. [8](#)
- SWINDLE, C. S., ZOU, N., VAN TINE, B. A., SHAW, G. M., ENGLER, J. A. et CHOW, L. T. (1999). Human papillomavirus dna replication compartments in a transient dna replication system. *J Virol*, 73(2):1001–1009. [14](#)
- THOMAS, M. et BANKS, L. (1998). Inhibition of bak-induced apoptosis by hpv-18 e6. *Oncogene*, 17(23):2943–2954. [13](#)
- THOMAS, M. et BANKS, L. (1999). Human papillomavirus (hpv) e6 interactions with bak are conserved amongst e6 proteins from high and low risk hpv types. *J Gen Virol*, 80 ( Pt 6):1513–1517. [38](#), [39](#)
- THOMAS, M., GLAUNSINGER, B., PIM, D., JAVIER, R. et BANKS, L. (2001). Hpv e6 and maguk protein interactions: determination of the molecular basis for specific protein recognition and degradation. *Oncogene*, 20(39):5431–5439. [39](#), [43](#)
- THOMAS, M. C. et CHIANG, C.-M. (2005). E6 oncoprotein represses p53-dependent gene activation via inhibition of protein acetylation independently of inducing p53 degradation. *Mol Cell*, 17(2):251–264. [40](#)
- TOMAIC, V., GARDIOL, D., MASSIMI, P., OZBUN, M., MYERS, M. et BANKS, L. (2009). Human and primate tumour viruses use pdz binding as an evolutionarily conserved mechanism of targeting cell polarity regulators. *Oncogene*, 28(1):1–8. [117](#)
- TOMMASINO, M., ADAMCZEWSKI, J. P., CARLOTTI, F., BARTH, C. F., MANETTI, R., CONTORNI, M., CAVALIERI, F., HUNT, T. et CRAWFORD, L. (1993). Hpv16 e7 protein associates with the protein kinase p33cdk2 and cyclin a. *Oncogene*, 8(1):195–202. [35](#)
- TONG, X. et 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(9):4412–4417. [39](#), [44](#)
- TONG, X., SALGIA, R., LI, J. L., GRIFFIN, J. D. et HOWLEY, P. M. (1997). 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*, 272(52):33373–33376. [39](#), [44](#)
- TRUS, B. L., RODEN, R. B., GREENSTONE, H. L., VRHEL, M., SCHILLER, J. T. et BOOY, F. P. (1997). Novel structural features of bovine papillomavirus capsid revealed by a three-dimensional reconstruction to 9 a resolution. *Nat Struct Biol*, 4(5):413–420. [9](#)
- TSENG, C. J., LIN, C. Y., WANG, R. L., CHEN, L. J., CHANG, Y. L., HSIEH, T. T. et PAO, C. C. (1992). Possible transplacental transmission of human papillomaviruses. *Am J Obstet Gynecol*, 166(1 Pt 1):35–40. [11](#)
- TUNGTEAKKHUN, S. S., FILIPPOVA, M., NEIDIGH, J. W., FODOR, N. et DUERKSEN-HUGHES, P. J. (2008). The interaction between human papillomavirus type 16 and fadd is mediated by a novel e6 binding domain. *J Virol*, 82(19):9600–9614. [39](#)
- ULLAH, H., SCAPPINI, E. L., MOON, A. F., WILLIAMS, L. V., ARMSTRONG, D. L. et PEDERSEN, L. C. (2008). Structure of a signal transduction regulator, rack1, from arabidopsis thaliana. *Protein Sci*, 17(10):1771–1780. [125](#)
- UNDERBRINK, M. P., HOWIE, H. L., BEDARD, K. M., KOOP, J. I. et GALLOWAY, D. A. (2008). E6 proteins from multiple human betapapillomavirus types degrade bak and protect keratinocytes from apoptosis after uvb irradiation. *J Virol*, 82(21):10408–10417. [40](#)
- VALLE, G. F. et BANKS, L. (1995). 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*, 76 ( Pt 5):1239–1245. [31](#)
- VAMBUTAS, A., DEVOTI, J., PINN, W., STEINBERG, B. M. et BONAGURA, V. R. (2001). Interaction of human papillomavirus type 11 e7 protein with tap-1 results in the reduction of atp-dependent peptide transport. *Clin Immunol*, 101(1):94–99. [35](#)
- VANDE POL, S. B., BROWN, M. C. et TURNER, C. E. (1998). Association of bovine papillomavirus type 1 e6 oncoprotein with the focal adhesion protein paxillin through a conserved protein interaction motif. *Oncogene*, 16(1):43–52. [39](#), [44](#), [53](#)

## Références bibliographiques

---

- VELDMAN, T., LIU, X., YUAN, H. et SCHLEGEL, R. (2003). Human papillomavirus e6 and myc proteins associate in vivo and bind to and cooperatively activate the telomerase reverse transcriptase promoter. *Proc Natl Acad Sci U S A*, 100(14):8211–8216. [41](#)
- VILLA, L. L., COSTA, R. L. R., PETTA, C. A., ANDRADE, R. P., AULT, K. A., GIULIANO, A. R., WHEELER, C. M., KOUTSKY, L. A., MALM, C., LEHTINEN, M., SKJELDESTAD, F. E., OLSSON, S.-E., STEINWALL, M., BROWN, D. R., KURMAN, R. J., RONNETT, B. M., STOLER, M. H., FERENCZY, A., HARPER, D. M., TAMMS, G. M., YU, J., LUPINACCI, L., RAILKAR, R., TADDEO, F. J., JANSEN, K. U., ESSER, M. T., SINGS, H. L., SAAH, A. J. et BARR, E. (2005). Prophylactic quadrivalent human papillomavirus (types 6, 11, 16, and 18) li virus-like particle vaccine in young women: a randomised double-blind placebo-controlled multicentre phase ii efficacy trial. *Lancet Oncol*, 6(5):271–278. [20](#)
- VOS, R. M., ALTREUTER, J., WHITE, E. A. et HOWLEY, P. M. (2009). The ubiquitin-specific peptidase usp15 regulates human papillomavirus type 16 e6 protein stability. *J Virol*, 83(17):8885–8892. [39](#), [43](#)
- WANG, H., MO, P., REN, S. et YAN, C. (2010a). Activating transcription factor 3 activates p53 by preventing e6-associated protein from binding to e6. *J Biol Chem*, 285(17):13201–13210. [39](#)
- WANG, H., MO, P., REN, S. et YAN, C. (2010b). Activating transcription factor 3 activates p53 by preventing e6-associated protein from binding to e6. *J Biol Chem*, 285(17):13201–13210. [42](#)
- WANG, J., ZHOU, D., PRABHU, A., SCHLEGEL, R. et YUAN, H. (2010c). The canine papillomavirus and gamma hpv e7 proteins use an alternative domain to bind and destabilize the retinoblastoma protein. *PLoS Pathog*, 6(9):e1001089. [35](#)
- WANG, Q., GRIFFIN, H., SOUTHERN, S., JACKSON, D., MARTIN, A., MCINTOSH, P., DAVY, C., MASTERSON, P. J., WALKER, P. A., LASKEY, P., OMARY, M. B. et DOORBAR, J. (2004). 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*, 78(2):821–833. [14](#), [31](#)
- WERNESS, B. A., LEVINE, A. J. et HOWLEY, P. M. (1990). Association of human papillomavirus types 16 and 18 e6 proteins with p53. *Science*, 248(4951):76–79. [39](#)
- WHITE, P. W., PELLETIER, A., BRAULT, K., TITOLO, S., WELCHNER, E., THAUVETTE, L., FAZEKAS, M., CORDINGLEY, M. G. et ARCHAMBAULT, J. (2001). 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*, 276(25):22426–22438. [25](#), [114](#)
- WILSON, V. G., WEST, M., WOYTEK, K. et RANGASAMY, D. (2002). Papillomavirus e1 proteins: form, function, and features. *Virus Genes*, 24(3):275–290. [12](#)
- XU, M., LUO, W., ELZI, D. J., GRANDORI, C. et GALLOWAY, D. A. (2008). Nfx1 interacts with msin3a/histone deacetylase to repress htert transcription in keratinocytes. *Mol Cell Biol*, 28(15):4819–4828. [39](#)
- YAGINUMA, Y., YODA, K. et OGAWA, K. (2010). Characterization of physical binding between human papillomavirus 18 protein e7 and centromere protein c. *Oncology*, 79(3-4):219–228. [35](#)
- YANG, R., YUTZY, W. H. t., VISCIDI, R. P. et RODEN, R. B. S. (2003). Interaction of l2 with beta-actin directs intracellular transport of papillomavirus and infection. *J Biol Chem*, 278(14):12546–12553. [24](#)
- YEE, C., KRISHNAN-HEWLETT, I., BAKER, C. C., SCHLEGEL, R. et HOWLEY, P. M. (1985). Presence and expression of human papillomavirus sequences in human cervical carcinoma cell lines. *Am J Pathol*, 119(3):361–366. [31](#)
- YIM, E.-K., LEE, K.-H., MYEONG, J., TONG, S.-Y., UM, S.-J. et PARK, J.-S. (2007). Novel interaction between hpv e6 and bard1 (brca1-associated ring domain 1) and its biologic roles. *DNA Cell Biol*, 26(10):753–761. [39](#), [42](#)
- YOU, J., CROYLE, J. L., NISHIMURA, A., OZATO, K. et HOWLEY, P. M. (2004). Interaction of the bovine papillomavirus e2 protein with brd4 tethers the viral dna to host mitotic chromosomes. *Cell*, 117(3):349–360. [12](#), [27](#), [106](#)

- ZANIER, K., CHARBONNIER, S., BALTZINGER, M., NOMINE, Y., ALTSCHUH, D. et TRAVE, G. (2005a). Kinetic analysis of the interactions of human papillomavirus e6 oncoproteins with the ubiquitin ligase e6ap using surface plasmon resonance. *J Mol Biol*, 349(2):401–412. [53](#)
- ZANIER, K., CHARBONNIER, S., BALTZINGER, M., NOMINE, Y., ALTSCHUH, D. et TRAVE, G. (2005b). Kinetic analysis of the interactions of human papillomavirus e6 oncoproteins with the ubiquitin ligase e6ap using surface plasmon resonance. *J Mol Biol*, 349(2):401–412. [102](#)
- ZANIER, K., OULD M'HAMED OULD SIDI, A., BOULADE-LADAME, C., RYBIN, V., CHAPPELLE, A., ATKINSON, A., KIEFFER, B. et TRAVE, G. (2012). Solution structure analysis of the hpv16 e6 oncoprotein reveals a self-association mechanism required for e6-mediated degradation of p53. *Structure*, 20(4):604–617. [36](#), [48](#), [50](#), [65](#), [71](#), [95](#), [116](#), [117](#)
- ZANIER, K., RUHLMANN, C., MELIN, F., MASSON, M., OULD M'HAMED OULD SIDI, A., BERNARD, X., FISCHER, B., BRINO, L., RISTRIANI, T., RYBIN, V., BALTZINGER, M., VANDE POL, S., HELLWIG, P., SCHULTZ, P. et TRAVE, G. (2010). E6 proteins from diverse papillomaviruses self-associate both in vitro and in vivo. *J Mol Biol*, 396(1):90–104. [48](#), [74](#)
- ZENG, M., KUMAR, A., MENG, G., GAO, Q., DIMRI, G., WAZER, D., BAND, H. et BAND, V. (2002). Human papilloma virus 16 e6 oncoprotein inhibits retinoic x receptor-mediated transactivation by targeting human ada3 coactivator. *J Biol Chem*, 277(47):45611–45618. [39](#)
- ZERFASS-THOME, K., ZWERSCHKE, W., MANNHARDT, B., TINDLE, R., BOTZ, J. W. et JANSEN-DURR, P. (1996). Inactivation of the cdk inhibitor p27kip1 by the human papillomavirus type 16 e7 oncoprotein. *Oncogene*, 13(11):2323–2330. [34](#), [35](#)
- ZHANG, B., CHEN, W. et ROMAN, A. (2006). The e7 proteins of low- and high-risk human papillomaviruses share the ability to target the prb family member p130 for degradation. *Proc Natl Acad Sci U S A*, 103(2):437–442. [35](#)
- ZHANG, P., NOURI, M., BRANDSMA, J. L., IFTNER, T. et STEINBERG, B. M. (1999). Induction of e6/e7 expression in cottontail rabbit papillomavirus latency following uv activation. *Virology*, 263(2):388–394. [12](#)
- ZHANG, Y., FAN, S., MENG, Q., MA, Y., KATIYAR, P., SCHLEGEL, R. et ROSEN, E. M. (2005). Brcal interaction with human papillomavirus oncoproteins. *J Biol Chem*, 280(39):33165–33177. [39](#), [42](#)
- ZHENG, Z.-M. et BAKER, C. C. (2006). Papillomavirus genome structure, expression, and post-transcriptional regulation. *Front Biosci*, 11:2286–2302. [8](#), [9](#)
- ZHOU, J., DOORBAR, J., SUN, X. Y., CRAWFORD, L. V., MCLEAN, C. S. et FRAZER, I. H. (1991). Identification of the nuclear localization signal of human papillomavirus type 16 l1 protein. *Virology*, 185(2):625–632. [21](#)
- ZHOU, J., STENZEL, D. J., SUN, X. Y. et FRAZER, I. H. (1993). Synthesis and assembly of infectious bovine papillomavirus particles in vitro. *J Gen Virol*, 74 ( Pt 4):763–768. [14](#)
- ZHOU, Y., WEI, Y., ZHU, J., WANG, Q., BAO, L., MA, Y., CHEN, Y., FENG, D., ZHANG, A., SUN, J., NALLAR, S. C., SHEN, K., KALVAKOLANU, D. V., XIAO, W. et LING, B. (2011). Grim-19 disrupts e6/e6ap complex to rescue p53 and induce apoptosis in cervical cancers. *PLoS One*, 6(7). [39](#), [40](#)
- ZHU, J., KOKEN, M. H., QUIGNON, F., CHELBI-ALIX, M. K., DEGOS, L., WANG, Z. Y., CHEN, Z. et de THE, H. (1997). Arsenic-induced pml targeting onto nuclear bodies: implications for the treatment of acute promyelocytic leukemia. *Proc Natl Acad Sci U S A*, 94(8):3978–3983. [115](#)
- ZIMMERMANN, H., DEGENKOLBE, R., BERNARD, H. U. et O'CONNOR, M. J. (1999). The human papillomavirus type 16 e6 oncoprotein can down-regulate p53 activity by targeting the transcriptional coactivator cbp/p300. *J Virol*, 73(8):6209–6219. [40](#)
- ZIMMERMANN, H., KOH, C. H., DEGENKOLBE, R., O'CONNOR, M. J., MÜLLER, A., STEGER, G., CHEN, J. J., LUI, Y., ANDROPHY, E. et BERNARD, H. U. (2000). Interaction with cbp/p300 enables the bovine papillomavirus type 1 e6 oncoprotein to downregulate cbp/p300-mediated transactivation by p53. *J Gen Virol*, 81(Pt 11):2617–2623. [39](#)

## Références bibliographiques

---

- zur HAUSEN, H. (2000). Papillomaviruses causing cancer: evasion from host-cell control in early events in carcinogenesis. *J Natl Cancer Inst*, 92(9):690–698. [17](#)
- zur HAUSEN, H. (2002). Papillomaviruses and cancer: from basic studies to clinical application. *Nat Rev Cancer*, 2(5):342–350. [3](#)
- ZWERSCHKE, W., MANNHARDT, B., MASSIMI, P., NAUENBURG, S., PIM, D., NICKEL, W., BANKS, L., REUSER, A. J. et JANSEN-DURR, P. (2000). Allosteric activation of acid alpha-glucosidase by the human papillomavirus e7 protein. *J Biol Chem*, 275(13):9534–9541. [35](#)

---

## Annexe

- 1 **Manuscrit en cours de rédaction : Structural basis for hijacking of cellular LxxLL motifs by a papillomavirus E6 oncoprotein**

**Structural basis for hijacking of cellular LxxLL motifs by a papillomavirus E6 oncoprotein.**

Sebastian Charbonnier, Nicole Brimer, Abdellahi Ould M'hamed Ould Sidi, Katja Luck, Katia Zanier, Khaled Ould Babah, Tina Ansari, Isabelle Muller, Pierre Poussin, Vincent Cura, Charles Lyons, Jean Cavarelli, Scott Vande Pol, Gilles Travé.

Affiliations

Sebastian Charbonnier, Abdellahi Ould M'hamed Ould Sidi, Katja Luck, Katia Zanier, Khaled Ould Babah, Isabelle Muller, Gilles Travé:

Oncoprotein Team, FRE CNRS 3211, Ecole Supérieure de Biotechnologie de Strasbourg, FRE 3211 CNRS-UDS, Boulevard Sébastien Brant, BP 10413, F-67412 Illkirch, FRANCE.

Nicole Brimer, Tina Ansari, Charles Lyons, Scott Vande Pol:

Department of Pathology, University of Virginia, P.O. Box 800904, Charlottesville, Virginia 22908-0904, USA.

Pierre Poussin, Vincent Cura, Jean Cavarelli:

Institut de Génétique et de Biologie Moléculaire et Cellulaire (IGBMC), Institut National de Santé et de Recherche Médicale (INSERM) U964/Centre National de Recherche Scientifique (CNRS) UMR 7104/Université de Strasbourg, 1 rue Laurent Fries, BP 10142, F-67404 Illkirch, France.

## **Abstract**

Papillomaviruses (PV) induce epithelial tumours in vertebrates, including cervical cancer in humans. PV E6 oncoproteins alter cellular processes such as proliferation, transcription, adhesion and apoptosis by capturing host cell proteins via acidic LxxLL helical motifs. Here, we solved the first structure of a full-length E6 protein bound to a cellular partner. Bovine Papillomavirus Type 1 (BPV-1) E6 was crystallised in complex with a "LD" protein interaction motif derived from the focal adhesion protein paxillin and containing a LxxLL sequence. BPV-1 E6 recognises its target peptide in a novel and tighter fashion as compared to cellular domains known to bind LD motifs or other LxxLL sequences. We show that critical sequence features of the bound peptide are conserved within a number of potential host target proteins of BPV-1 E6. Thus, E6 proteins appear to have evolved a binding pocket for hijacking a particular class of protein interaction motifs, thereby specifically perturbing the host cell functions mediated by these motifs.



## Introduction

Papillomaviruses (PV) infect the cutaneous or mucosal epithelia of vertebrates, with more than 200 PV types so far identified and sequenced<sup>1</sup>. Whereas most PVs produce epithelial hyperplasias, infections by a subset of types known as "high-risk" PVs may eventually lead to cancer. In particular, cervical cancers are caused by High-risk mucosal Human PVs (Hrm-HPVs)<sup>2</sup> and some skin cancers have been associated with high risk cutaneous HPV<sup>s</sup><sup>3</sup>. Bovine Papillomavirus Type 1 (BPV-1) is a model system for papillomavirus transcription, transformation and replication<sup>4</sup> and induces tumours in its natural host (cattle) and in a heterologous host (equids).

PV carcinogenesis is primarily linked to two PV oncoproteins, E6 and E7. Hrm-HPV E6 recruits the ubiquitin ligase E6AP and the tumour suppressor p53, leading to ubiquitin-mediated degradation of p53<sup>5</sup>. Hrm-HPV E6 also interacts with many other cellular proteins, sometimes resulting in their proteasome-dependent degradation<sup>6</sup>. Hrm-HPV E6 recognises several of its target proteins (including E6AP<sup>5</sup> and IRF-3<sup>7</sup>) *via* acidic Leucine-rich motifs containing a LxxLL consensus sequence<sup>8,9</sup>. Low-risk mucosal HPV-11 also interacts with the Leucine-rich motif of E6AP<sup>10</sup>. Finally, BPV-1 E6 recognises within the focal adhesion protein paxillin several acidic sequences, containing the LxxLL consensus plus additional conserved features, known as "LD motifs"<sup>11</sup>. E6 binding to LD motifs is required for cellular transformation by E6<sup>12,9,13</sup>. Within the host cell, LD motifs regulate cell motility, cell adhesion and gene expression by mediating the interaction of paxillin and related proteins with partner proteins including Focal Adhesion Kinase (FAK), Vinculin, GIT1<sup>11</sup> and alpha Parvin<sup>14</sup>.

Several structures of cellular domains bound to LD motifs or other LxxLL-containing sequences have been solved so far. They include the FAT domains of FAK, Vinculin and GIT1 bound to LD motifs<sup>15,16,17,18</sup>, the Calponin-Homology (CH) domain of alpha Parvin also bound to a LD motif<sup>14,19</sup>, and several LBD domains of nuclear receptors bound to the LxxLL motifs found within the "NR boxes" of transcription coactivators<sup>20</sup>. In all cases, the bound LxxLL motif adopts a helical conformation, occupies a peripheral position relative to the target domain and establishes mainly helix-helix contacts with the surface of the target domain.

Most mammalian PV E6 proteins are small cysteine-rich proteins consisting of two zinc-binding domains, E6N and E6C. Whereas the solution structure of a soluble mutant of HPV-16 E6C domain has been determined<sup>21</sup>, full-length mammalian E6

proteins, including BPV-1 E6, undergo self-oligomerisation processes, which have precluded their structural analysis<sup>22</sup>.

### **Crystal structure of an MBP - LD motif - E6 fusion construct**

In an attempt to prevent E6 oligomerisation, the LxxLL sequence present in the LD1 motif of paxillin (sequence [<sub>1</sub>MDDLADALLAD<sub>10</sub>]) was fused, via a short linker (sequence [KEGGG]), to the N-terminus of the E6 protein. This fusion construct, called LD-E6, was highly soluble (Vande Pol S., unpublished observations) and yielded NMR spectra representative of a monomeric folded protein (Ould Sidi M'hamed et al., manuscript in preparation). Furthermore, the same LD-E6 construct was selectively disrupted for *in vivo* interactions with cellular target proteins of E6 and transformation by E6<sup>12</sup> (supplemental Fig. 1). This strongly suggested that the fused LD motif efficiently docked into its cognate binding site within the E6 protein. The LD-E6 construct was then fused via a three-alanine linker to the C-terminus of a crystallisation-prone mutant of Maltose-Binding Protein (MBP)<sup>23,24</sup>. The resulting MBP-LD-E6 triple fusion construct crystallised in the space group C222<sub>1</sub>, yielding diffraction data at a resolution better than 2.3 Å using synchrotron radiation. The structure was solved by molecular replacement using the known structure of MBP as a template (Supplemental Fig. 2). The MBP and LD-E6 moieties are assembled as a single globular unit whose compactness may have facilitated crystallisation. The 10 first N-terminal and the 7 last C-terminal residues of E6 are not visible in the electron density map. However, residues 1-10 (<sup>1</sup>MDLKPFARTN<sup>10</sup>) of BPV-1 E6 do not seem essential for target motif recognition, since their replacement by the EE epitope tag sequence (<sup>1</sup>EEEEYYMPME<sup>10</sup>) did not ablate *in vitro* binding of E6 to the LD motif (supplemental Fig. 3).

### **Structural features of the E6 protein**

The structure of E6 bound to paxillin comprises two zinc-binding domains joined by a linker helix (Fig. 1A, B and C). The C-terminal domain (E6C, residues 58-137) adopts a zinc-binding fold similar to that of the isolated HPV16 E6C domain in solution<sup>21</sup>. The resolved region of the N-terminal domain (E6N, residues 11-57) shares common structural features with the corresponding region of E6C, onto which it can be well superimposed (supplemental Fig. 4).

The interdomain linker helix comprises four turns spanning residues 51-64 (Fig. 1B and C), and is anchored by domain interactions at both ends. The first two turns of the helix are anchored to the E6N domain via zinc-coordinating residues as well as polar and hydrophobic contacts (Fig. 1D). The two last turns of the helix are anchored to the E6C domain via hydrophobic and polar contacts (Fig. 1E). The E6N and E6C domains also directly interact with each other through additional contacts, mainly electrostatic in nature (Fig. 1F).

The sequences of E6 proteins of various papillomaviruses are well aligned with that of BPV-1 E6 (Fig. 2), indicating the conservation of the overall structure composed of two zinc-binding domains connected via a linker helix.

### **E6 displays a novel mode of LxxLL motif recognition**

The E6-bound peptide (sequence M<sub>1</sub>D<sub>2</sub>D<sub>3</sub>L<sub>4</sub>D<sub>5</sub>A<sub>6</sub>L<sub>7</sub>L<sub>8</sub>A<sub>9</sub>D<sub>10</sub>) adopts a helical conformation from residue D<sub>2</sub> to residue L<sub>8</sub> (Fig. 1C) and inserts inside a groove composed by the two zinc-binding domains and the linker helix of E6 (Fig. 1C, 3A and 3C). The three leucine residues L<sub>4</sub>, L<sub>7</sub> and L<sub>8</sub> defining the LxxLL motif are plugged into a pocket of hydrophobic residues (W19, F37, V40, A49, L54, C57, L58) exclusively contributed by the E6N domain (Fig. 3, A and E). The E6N domain also contributes an electrostatic component to the complex, via R42 whose side chain forms a salt bridge with D<sub>3</sub> of the peptide and is also proximal to D<sub>2</sub> (Fig. 3B, bottom right). However, most E6-peptide charge interactions are contributed by E6C. The surface of E6C oriented towards the peptide (Fig. 3C) bears seven basic residues (H79, K81, R85, R89, K96, R116 and R121), resulting in a large positively charged surface, which provides a favorable electrostatic environment likely to attract negatively charged peptides into the binding pocket. The side-chain of R116 establishes a critical salt bridge with the side chain of residue D<sub>5</sub> of the peptide. Other positive residues of the E6C surface participate to a network of interactions mediated by water molecules connecting the peptide and the binding pocket (Fig. 3B). Thus, the amphipatic acidic LxxLL helical motif resides within a complementary hydrophobic and positively charged pocket, with further anchoring being provided by specific salt bridges and H-bonds (Fig. 3E).

The cellular FAT and CH domains are helical domains that recognise the LD motif by means of surface interactions, with the bound motif occupying a peripheral position<sup>16,19</sup> (Fig. 3D). Helix-shaped LxxLL motifs are also found within the "NR boxes" that mediate the interaction of transcription coactivators with the LBD domains

of nuclear receptors<sup>20</sup>. Like the FAT and CH domains, the LBD domains are fully helical and capture the LxxLL motif in a rather peripheral position (Fig. 3D). Comparatively, the LD peptide of paxillin is significantly more buried when clamped into the hydrophobic pocket of viral E6 (Fig. 3C). Accordingly, E6 binds tightly to the paxillin LD motif, with an equilibrium dissociation constant  $K_D \sim 50$  nM (Ould Sidi M'hamed et al., manuscript in preparation). By contrast, cellular LD-binding or NR box-binding domains bind more weakly to their cognate LD motifs, with  $K_D \sim 100$   $\mu$ M for FAT domains<sup>16</sup>,  $K_D$  values ranging between 100  $\mu$ M and 1 mM for CH domains<sup>19</sup>, and  $K_D \sim 1$   $\mu$ M for LBDs<sup>25</sup>, respectively. This indicates that E6 should be able to compete very efficiently with host proteins for binding to paxillin and other proteins displaying suitable LD or LxxLL motifs.

In E6 protein sequence alignments (Fig. 2) most hydrophobic and polar positions critical for peptide binding are well conserved in nature, yet not identical. This suggests that HPV E6 proteins will recognise comparable acidic / hydrophobic helical motifs, albeit with fine variations in their sequence recognition specificities.

#### **Engineering E6 mutations that either suppress or alter motif recognition.**

The E6-peptide structure reveals 21 BPV-1 E6 residues (indicated on top of the alignment in Fig. 2) that are involved in atomic contacts with the peptide motif. 12 single point mutations altering residues of the motif binding site of E6 were generated, among which 8 mutants were defective for LD motif recognition (Fig. 4). Remarkably, E6 mutants disrupted for LD-motif binding systematically lost the transformation phenotype in living cells (Fig. 4A).

Two ion bridges involving basic residues of E6 and acidic residues of the LD motif (R42-D3 and R116-D5) constitute electrostatic clamps (Fig. 3, B and C), which may be critical for maintaining the peptide in place, thereby contributing to ligand selectivity. We mutated three aspartic residues of the paxillin LD1 motif into arginine (mutants D3R, D5R, D3R/D5R, D10R) and two arginine residues of E6 into aspartate: R42D, R116D and R42D/R116D. The pattern of recognition of each E6 mutant against the various peptide sequences was then investigated both by yeast two-hybrid (Fig.4B) and pulldown assay (Fig. 4C). Wild-type E6 retained strong binding to all sequences with the exception of the double mutant peptide (Fig. 4, B and C). Therefore, no single aspartic position in the peptide is essential for recognition by E6 wt. The E6 R42D

mutant retained a recognition pattern quite similar to that observed for E6 wt, indicating that the R42-D3 ion bridge contributes only moderately to recognition. By contrast, the charge inversion R116D (mutants E6 R116D and E6 R42D/R116D) created a more selective E6 protein, displaying a strong preference for the mutated D5R motif bearing the inverse charge mutation. This indicates that R116 plays a critical role in recognition and that a swapped R5/D116 ion bridge has successfully replaced the original D5/R116 ion bridge in the mutated complex, resulting in novel peptide specificity.

### **Refining the definition of the BPV-1 E6-binding motif.**

To further define the critical determinants of LD-E6 interaction, 12-meric peptides binding selectively to BPV-1 E6 were isolated out of a random peptide library using phage display (supplemental Fig. 5). Selected sequences were strikingly similar to the sequences of the E6-binding paxillin repeats LD1, LD2 and LD4, allowing us to propose a 10-residue long regular expression for BPV-1 E6-binding sequences:  $\Phi_1 X_2 D_3 L_4 D_5 [-]_6 L_7 (F/L)_8 X_9 [-]_{10}$  (Fig. 3F and supplemental Fig. 5). The strongest conservation was observed for the hydrophobic positions L4, L7 and (F/L)8, probably due to high steric constraints within the hydrophobic pocket of E6N. The preferred hydrophobic residue at position 1 is coherent with the structural data showing hydrophobic contacts between residues M1 of paxillin and W65 of BPV-1 E6. Aspartate residue D5 in the motif was more conserved than aspartate residue D3, confirming the mutagenesis data (Fig. 4), which showed that R116 (which contacts D5 via an ion bridge) contributes more to binding strength than R42 (which contacts D3 via an ion bridge). In total, four positions in the motif (3, 5, 6 and 10) showed preferences for acidic residues (Fig. 3F). The net charge of the E6-binding peptides was always negative; yet no position strictly required a negatively charged residue. This relaxed sequence requirement for negatively charged residues is in agreement with the structural data, which showed that the E6C domain displayed a strongly positive surface prone to attract negatively charged ligands, yet did not engage many specific interactions with the bound ligand (Fig. 3 B, C and E). Whereas the phage display revealed residues that can be accepted at each position of E6-binding peptides (Supplemental Fig. 5C), mutagenesis by Bohl et al.<sup>12</sup> revealed residues that strongly disrupted the binding of paxillin peptides to BPV-1 E6 (Supplemental Fig. 5B). Interestingly, most of the

disruptive residues described in Bohl et al. were not selected by phage display, indicating a very good agreement between the two types of approaches.

### **Structure-driven proteome-wide predictions of novel putative E6 binders**

By combining the information gained from structural analysis, phage display and mutagenesis data, we built a position specific scoring matrix (PSSM) profile<sup>26</sup> defining precisely the LxxLL sequence motif recognised by BPV-1 E6 (supplemental table 1). This profile was used to perform a sequence similarity search on the whole human proteome. Because short protein interaction motifs are mainly found in non-folded areas<sup>27</sup>, we restricted our search to regions of the proteome predicted to be non-folded. This procedure allowed us to generate and rank a short list of host cell proteins potentially targeted by BPV-1 E6 (Table 1 and supplemental Table 2). Remarkably, 9 human proteins (Fig. 5A) predicted by our approach to be highly likely binders of E6, were independently identified by co-precipitation experiments coupled with mass spectroscopy analysis (Fig. 5B). Furthermore, all the short acidic LxxLL motifs detected within these proteins by our *in silico* search were found to bind to BPV-1 E6 *in vitro* (Fig. 5C). This indicates that a thorough structural analysis of a viral-host protein complex may foster accurate proteome-wide bioinformatic predictions of the cellular targets of a viral protein. We propose that, when possible, such structure-based predictions should accompany high-throughput experimental interactomics approaches in order to increase the significance and/or decrease the number of the putative targets deserving to be explored for further biological validation.

Statistical tests reveal that the list of potential targets of E6 is highly enriched in proteins, which, like paxillin, participate in cell adhesion and cytoskeletal dynamics and organisation, as well as numerous regulators and coactivators of transcription (Fig. 5A and supplemental Table1). Interestingly, some of the best ranked putative E6-binding proteins display the consensus sequence LDxLLxxL representative of the full LD motif<sup>11</sup>. However, most sequences in the list do not match the full LD motif, while presenting acidic leucine-rich patterns compatible with the E6 peptide binding pocket. Therefore, the LD motif may represent only a sub-class of a larger family of acidic leucine rich motifs potentially targeted by the E6 protein.

### **Discussion**

**The first structural data on a full-length E6 protein bound to a cellular target.**

The structure of a full-length E6 protein has been awaited for more than twenty years, mainly due to limitations in the solubility of E6 proteins. We turned around this problem by expressing a triple fusion construct combining the crystallisation prone MBP, the E6-binding motif of paxillin and the BPV-1 E6 protein. As yet, the same strategy did not work for HPV E6 proteins, which may require exploration of other approaches.

The structure of the complex fully explains the strong binding preference of E6 proteins for acidic LxxLL motifs<sup>12,8</sup>, with the E6N domain providing a hydrophobic pocket for the three leucine residues, and the E6C domain presenting a strongly positive surface to attract and interact with the negative charges of the peptide. Sequence conservation indicates that the overall structure of E6 proteins and their mode of binding to acidic hydrophobic motifs are conserved across mammalian PV species.

It seems unlikely that E6 would present the same structure in the absence of its target peptide. The few electrostatic interactions observed between E6N and E6C in the complex (see Fig. 1F) should be insufficient to maintain the two domains in their relative positions. In addition, solvent exposure of the large hydrophobic pocket hidden by the peptide (see Fig. 3A) should be energetically unfavorable. An alternative structure, in which E6N and E6C domains faced each other symmetrically in a pseudodimeric fashion, was previously proposed for unbound HPV16 E6<sup>21</sup>. Unfortunately, in our hands all samples of unbound wild-type E6 proteins (either from BPV or HPV) were prone to self-association and therefore unsuitable for structural analysis<sup>22</sup>. Indeed, it seems likely that most E6 molecules exist as target-bound complexes in infected cells.

**A viral strategy for hijacking of motif-mediated interaction networks**

Selective interactions between short linear interaction motifs and their cognate target domains are known to mediate protein-protein interaction networks involved in particular cellular functions<sup>28</sup>. We suggest that a number of viral proteins, such as E6, have evolved to capture interaction motifs, such as the LD motif, which participate in biological functions whose perturbation is required for a productive viral life cycle. Such a strategy is in principle extremely efficient, because a single binding pocket recognising the key conserved residues of a target motif is able to capture a large

number of instances of the motif present among numerous cellular proteins and therefore strongly disrupt the entire functional pathway mediated by the motif. This strategy may be further potentiated when, as shown here for E6, the viral protein evolves a pocket which binds tighter to the motif than its natural cellular partners. Indeed, it was previously proposed that BPV-1 E6 might transform cells through such competitive interactions at the LD motifs of paxillin<sup>29</sup>.

The E6 proteins of bovine and human papillomaviruses are phylogenetically linked and both contain a pocket which recognises acidic LxxLL sequences. However, they distinguish between sub-categories of such sequences. For instance, BPV-1 E6 preferentially recognises paxillin LD motifs, whereas HPV16 E6 preferentially recognises the E6AP LxxLL sequence. It is tempting to speculate that there may be a common functional and phylogenetic link between the subsets of acidic LxxLL sequences recognised by the different mammalian papillomavirus E6 proteins. These LxxLL sequences might represent members of a large family of protein interaction motifs generally involved in biological functions targeted by papillomavirus E6 proteins, such as transcription, cell adhesion and dynamics, cell proliferation or cell death.

In this regard, it will be interesting to analyse the structures of human homologues of BPV-1 E6, together with their fine variations in target peptide sequence recognition and to use this information to perform bioinformatics searches of their potential cellular target proteins. It will be also exciting to investigate whether other LxxLL-containing sequences targeted by E6 proteins constitute, similarly to LD motifs, protein-protein interaction motifs mediating host functional networks hijacked by papillomaviruses.

### **Concluding remarks**

The structure-function analysis of the E6-paxillin complex has revealed the general rule for E6-LxxLL motif recognition, illuminated a viral strategy for hijacking of cellular protein-protein interaction motifs, and has allowed us to predict, out of the entire host proteome, short lists of likely targets of the E6 oncoprotein. Comparable structural and functional studies on the cancer relevant complexes engaged by human HPV proteins should help us to identify and/or further validate lists of cellular target proteins of oncogenic HPVs and to generate accurate structural models of viral-cellular complexes, which could facilitate the efficient design and improvement of inhibitory compounds.



## Methods Summary

A triple MBP-LxxLL-E6 fusion, using a crystallisation prone mutant of MBP<sup>24,23</sup>, was bacterially expressed, purified as described<sup>22</sup> and crystallised. X-ray diffraction data were processed<sup>33</sup> and the structure was solved at a 2.3 Å resolution by molecular replacement using MBP structure (pdb code 4MPB) and standard protein crystallography programs<sup>34,35,37,36</sup>. Proteomes from Ensembl v58<sup>38</sup> were searched for potential E6-binding LxxLL motifs within putatively disordered regions<sup>39</sup>, returning 387 motifs which we ranked using a PSSM<sup>26</sup> describing the peptide binding preferences of BPV-1 E6, and annotated for various structural, functional and localisation informations<sup>40,41</sup>. CBD-TEVsite-FLAG-E6 and CBD-TEVsite-FLAG-LxxLL-E6 fusions expressed in mammalian cells using vaccinia virus were incubated with cell lysates for tandem affinity capture of target proteins that were subsequently identified by Mass Spectrometry<sup>43,44,45</sup>. Proteins associated with E6 but not with LxxLL-E6 were retained as potential binders of the peptide-binding pocket of E6. For selecting E6-binding peptides, four phage display rounds were performed with alternation of GST-E6 and CBD-E6-coated selection beads. Protein interactions were analysed by yeast two-hybrid, pulldown and chromatographic holdup assays as previously described<sup>46,9,32</sup>. Anchorage independent colony formation assays of BPV-1 E6 and E6 mutants expressed using retroviral vectors<sup>47</sup> were performed as previously described<sup>9</sup>.

## Acknowledgments

This work was supported by institutional support from CNRS, Université de Strasbourg, INSERM, the European Commission SPINE2-Complexes project (contract n° LSHG-CT-2006-031220) and grants from ARC (n° 3171) and ANR (ANR-MIME-2007 EPI-HPV-3D). S.C. was supported by ANR, A.o.M.o.S. by ARC and K.L. by Région Alsace and College Doctoral Européen. S.V.P, N.B, and T.A. were supported by NIH grants (CA120352 and CA08093) to S.V.P, and institutional support from the University of Virginia and the Department of Pathology Mass Spectrometry Facility. The authors thank all members of the ESRF-EMBL joint structural biology groups, all members of SOLEIL for the use of their synchrotron beamline facilities and for help during data collection, all members of the Structural Genomics Platform of IGBMC for setting up automated procedures, members of the IGBMC's common services for

assistance, and all members of the Oncoprotein group for helpful discussions and advice.

#### **Author Contributions**

S.C., N.B., A.o.M.o.S., K.Z., T.A., K.o.B., I.M., P.P., V.C., C.L. and S.V.P. performed experiments; K.L. performed bioinformatics predictions; J.C. performed structure determination. S.C., K.L., J.C., S.V.P. and G.T. analysed the data. G.T., S.V.P., S.C. and J.C. prepared the manuscript. G.T., S.V.P. and J.C. supervised the work.

corresponding authors: G.T., S.V.P. and J.C.

gilles.trave@unistra.fr

vandepol@virginia.edu

cava@igbmc.fr

The authors declare no competing financial interests.

## References

- 1 H. U. Bernard, R. D. Burk, Z. Chen et al., *Virology* **401** (1), 70 (2010).  
2 H. zur Hausen, *Semin Cancer Biol* **9** (6), 405 (1999).  
3 H. Pfister, P. G. Fuchs, S. Majewski et al., *Arch Dermatol Res* **295** (7), 273  
(2003).  
4 L. Nasir and M. S. Campo, *Vet Dermatol* **19** (5), 243 (2008).  
5 J. M. Huibregtse, M. Scheffner, and P. M. Howley, *Mol Cell Biol* **13** (8), 4918  
(1993).  
6 S.S. Tungteakkhun and P.J. Duerksen-Hughes, *Arch Virol.* **153** (3), 397 (2008).  
7 L. V. Ronco, A. Y. Karpova, M. Vidal et al., *Genes Dev* **12** (13), 2061 (1998).  
8 X. Be, Y. Hong, J. Wei et al., *Biochemistry* **40** (5), 1293 (2001).  
9 S. B. Vande Pol, M. C. Brown, and C. E. Turner, *Oncogene* **16** (1), 43 (1998).  
10 N. Brimer, C. Lyons, and S. B. Vande Pol, *Virology* **358** (2), 303 (2007).  
11 D. A. Tumbarello, M. C. Brown, and C. E. Turner, *FEBS Lett* **513** (1), 114  
(2002).  
12 J. Bohl, K. Das, B. Dasgupta et al., *Virology* **271** (1), 163 (2000).  
13 R. Wade, N. Brimer, and S. Vande Pol, *J Virol.* **82** (12), 5962 (2008).  
14 X. Wang, K. Fukuda, I. J. Byeon et al., *J Biol Chem* **283** (30), 21113 (2008).  
15 C. Bakolitsa, J. M. de Pereda, C. R. Bagshaw et al., *Cell* **99** (6), 603 (1999).  
16 M. K. Hoellerer, M. E. Noble, G. Labesse et al., *Structure* **11** (10), 1207 (2003).  
17 G. Gao, K. C. Prutzman, M. L. King et al., *J Biol Chem* **279** (9), 8441 (2004).  
18 Z. M. Zhang, J. A. Simmerman, C. D. Guibao et al., *J Biol Chem* **283** (27),  
18685 (2008).  
19 S. Lorenz, I. Vakonakis, E. D. Lowe et al., *Structure* **16** (10), 1521 (2008).  
20 R. S. Savkur and T. P. Burris, *J Pept Res* **63** (3), 207 (2004).  
21 Y. Nomine, M. Masson, S. Charbonnier et al., *Mol Cell* **21** (5), 665 (2006).  
22 K. Zanier, C. Ruhlmann, F. Melin et al., *J Mol Biol* **396** (1), 90 (2010).  
23 H. Ullah, E. L. Scappini, A. F. Moon et al., *Protein Sci* **17** (10), 1771 (2008).  
24 B. Kobe, R. J. Center, B. E. Kemp et al., *Proc Natl Acad Sci U S A* **96** (8), 4319  
(1999).  
25 E. Hur, S. J. Pfaff, E. S. Payne et al., *PLoS Biol* **2** (9), E274 (2004).  
26 D.W. Mount, *Sequence and Genome Analysis* 2nd Edition ed. (Cold Spring  
Harbor Laboratory Press, New York City, 2004).  
27 F. Diella, S. Chabanis, K. Luck et al., *Bioinformatics* **25** (1), 1 (2009).  
28 F. Diella, N. Haslam, C. Chica et al., *Front Biosci* **13**, 6580 (2008).  
29 X. Tong, R. Salgia, J. L. Li et al., *J Biol Chem* **272** (52), 33373 (1997).  
30 K. Das, J. Bohl, and S. B. Vande Pol, *J Virol* **74** (2), 812 (2000).  
31 R. Ned, S. Allen, and S. Vande Pol, *J Virol* **71** (6), 4866 (1997).  
32 S. Charbonnier, K. Zanier, M. Masson et al., *Protein Expr Purif* **50** (1), 89  
(2006).  
33 Z. Otwinowski and W. Minor, *Methods in Enzymology* **276**, 307 (1997).  
34 *Acta Crystallogr D Biol Crystallogr* **50** (Pt 5), 760 (1994).  
35 P. D. Adams, P. V. Afonine, G. Bunkoczi et al., *Acta Crystallogr D Biol  
Crystallogr* **66** (Pt 2), 213 (2010).  
36 P. Emsley and K. Cowtan, *Acta Crystallogr D Biol Crystallogr* **60** (Pt 12 Pt 1),  
2126 (2004).  
37 V.B. Chen, W.B. 3rd Arendall, J.J. Headd et al., *Acta Crystallogr D Biol  
Crystallogr* **66**, 12 (2010).

- 38 T. J. Hubbard, B. L. Aken, S. Ayling et al., *Nucleic Acids Res* **37** (Database  
issue), D690 (2009).
- 39 Z. Dosztanyi, V. Csizmok, P. Tompa et al., *Bioinformatics* **21** (16), 3433 (2005).
- 40 G. Dennis, Jr., B. T. Sherman, D. A. Hosack et al., *Genome Biol* **4** (5), P3  
(2003).
- 41 A. Lupas, M. Van Dyke, and J. Stock, *Science* **252** (5010), 1162 (1991).
- 42 A. Shevchenko, M. Wilm, O. Vorm et al., *Anal Chem* **68** (5), 850 (1996).
- 43 L. Y. Geer, S. P. Markey, J. A. Kowalak et al., *J Proteome Res* **3** (5), 958  
(2004).
- 44 K. G. Victor, M. Murgai, C. E. Lyons et al., *J Am Soc Mass Spectrom* **21** (1), 80  
(2010).
- 45 L. Kall, J. D. Storey, M. J. MacCoss et al., *J Proteome Res* **7** (1), 29 (2008).
- 46 B. Cooper, D. Hutchison, S. Park et al., *Plant Mol Biol* **53** (3), 273 (2003).
- 47 J. P. Morgenstern and H. Land, *Nucleic Acids Res* **18** (12), 3587 (1990).

## Figure legends

### Figure 1.

**A.** Schematics of the crystallised fusion protein. Pink: optimized MBP mutant, red: AAA linker, green: LxxLL peptide derived from paxillin LD1, black: KEGGG linker, blue: E6N, grey: linker helix, gold: E6C. Residues 1-10 of paxillin LD1 are numbered 372-381 in the pdb file.

**B.** Topology of E6 structure.  $Zn_1$  and  $Zn_2$  are zinc atoms. 1' and 2' are short 3/10 helices.

**C.** Overall structure of E6 bound to paxillin LD1 peptide.

**D.** The first two turns of the linker helix are anchored within E6N domain via zinc chelators C50 and C53, electrostatic interactions involving T51, R31 and D28, and hydrophobic interactions involving L54, F47, A49 and W19.

**E.** The two last turns of the linker helix are anchored within the E6C domain via electrostatic interactions between E61, R121, R116 and between R62 and C91; and stacking of the ring of W65 against the side chain of R116.

**F.** E6N and E6C also directly interact with each other via electrostatic contacts. D36 of E6N is proximal to K96 and R89 of E6C. R89 of E6C forms two H-bonds with D36 and K35 of E6N. K35 of E6N also forms a H-bond with C93.

### Figure 2.

Alignment of BPV-1 E6 sequence versus E6 sequences of BPV8 and HPV 1, 3, 5, 6, 9, 16 and 18. BPV-1 residue labels are indicated below the alignment. Residue solvent accessibilities were calculated using PYMOL in presence (white bars) or absence (black bars) of the bound peptide. Residues becoming more exposed upon peptide removal are labelled as peptide binders on top of the figure. The red line serves to identify residues that are more than 80% buried (solvent exposure below 20%). Peptide binding residues that remain more than 80% buried upon peptide removal are boxed. Above the bar graph, secondary elements are plotted with labels indicating the starting residue of each element.

### Figure 3.

**A.** The hydrophobic pocket responsible for recognition of the conserved LxxLL motif.

**B.** Charged, polar and water-mediated interactions between E6 and peptide. Note the electrostatic charge clamps R42 - D3 and R116 - D5.

**C.** E6 surface charge potential, expressed as a spectrum ranging from  $-12$  kT/e to  $+12$  kT/e. Blue: positive charge, red: negative charge. Note the strong positive surfaces facing the peptide acidic residues (coloured red).

**D.** The LxxLL motif displays superficial binding to cellular domains in contrast to its tight clamped binding to viral E6. From left to right are shown known structures of LxxLL motifs bound to a FAT domain (PDB: 1OW7), a CH domain (PDB: 2VZI) and a Nuclear Receptor LBD (PDB: 3ERD).

**E.** Key contacts between E6 residues and bound peptide. Magenta: hydrophobic residues; blue: positive charges; pink dashed lines: hydrophobic van der Waals contacts; black lines: polar contacts involving the side chain (plain black lines) or the main chain (interrupted black lines) of an E6 residue.

**F.** The consensus motif recognised by BPV-1 E6, derived from aligned sequences of paxillin motifs LD1, LD2 and LD4 and of E6-binding phage peptides (Supplemental Fig. 5).

#### **Figure 4.**

**A.** E6 mutants impaired for LxxLL recognition are defective for transformation, whereas mutants retaining LxxLL recognition remain proficient for transformation. Upper part: yeast two-hybrid interactions (indicated by dark spots) between *lexA*-fused wild-type or mutant E6 and B42-fused full-length paxillin (PXN) or LD1 motif (paxillin residues M<sub>1</sub>DDLDALLADL<sub>11</sub>). Lower part: transformation phenotype of E6 mutants, quantified using numbers of anchorage independent colonies<sup>9</sup> normalised to results from E6 wt. '+' and '#' indicate previously published results<sup>30,31,9</sup> and temperature sensitive mutants<sup>30</sup>, respectively. Boxed labels indicate residues that remain more than 80% buried upon peptide removal (see Fig. 2). Bar lines denote Standard Error of the Mean (SEM). '\*' indicate E6 mutants significantly impaired for transformation ( $P < 0.05$ ).

**B.** Engineering E6 mutants with altered LxxLL-binding specificities (yeast two-hybrid assay). R42 and R116 of E6 are both involved in electrostatic clamps with D3 and D5 of the peptide (see Fig. 3B). We generated charge inversion mutants of E6 (R42D and R116D) and of the paxillin LD1 motif (D3R, D5R and D3R-D5R). E6 wt displays binding to most peptides except the D3R-D5R double mutant; E6 R42D bound

only to LD1 wt and D3R mutant, whereas mutant E6 R116D bound only to the D5R mutant.

**C.** Engineering E6 mutants with altered LxxLL-binding specificities (GST-pulldown assay). Names of GST-fused LD1 peptides and of <sup>35</sup>S-labeled E6 constructs are indicated on top of the figure and at the right of each panel, respectively. For each panel, from top to bottom: GST-peptide inputs, E6 inputs, pulled-down E6, and bar graph quantitation of binding from three separate assays. '\*' indicate E6 mutants significantly impaired for binding (P<0.05). E6 wt, E6 R42D and E6 R116D show peptide binding patterns comparable to those observed in B. Double mutant E6 R42D/R116D is comparable to E6 R116D, except that it is further disrupted for binding to LD1 wt motif.

**Figure 5.**

**A.** Proteome-wide prediction of putative cellular targets of BPV-1 E6. We screened *in silico* all presumably unfolded regions of the human proteome to identify and rank a list of 387 putative BPV-1 E6 binding motifs (full list provided in Supplemental Table 3). Remarkably, 6 out of the 8 best ranked predicted motifs belong to proteins (coloured red or blue in the list) that we independently identified as BPV-1 E6 binders using Tandem Affinity Purification and Mass Spectrometry (TAP-MS) (see panel B). The rectangular hit distribution plot above the list also shows that most experimentally validated predictions are well ranked in the list. Red and blue colours indicate the best scored and the lower scored motifs found in each E6-coprecipitated protein, respectively. Columns indicate, from left to right: prediction ranking number; name of protein; motif sequence; motif localisation in the protein; net charge of the motif; coincidence ('yes' or 'no') of the LxxLL motif with a LD motif; and occurrence (marked by a 'x') of keywords that are highly enriched in annotation terms of the predicted binders of E6.

**B.** Representative SDS-PAGE of potential cellular binders of E6 identified by TAP-MS. Lane 1 : TAP using FLAG-E6. Lane 2 : TAP using FLAG-LxxLL-E6. Note that presence in Cis of the LxxLL-peptide inhibits binding of most cellular interaction partners. Proteins interacting with E6 and not LxxLL-E6 are labeled left of lane 1. **C.** BPV-1 E6 interacts with all predicted LxxLL motifs belonging to proteins independently identified by TAP-MS. Biotinylated LxxLL motifs were assayed for interaction with MBP-fused E6 using a modification of the "holdup" comparative

chromatographic retention assay<sup>32</sup> followed by SDS-PAGE analysis. For each assay, left band and right band represent the MBP-E6 flow-through from biotin-saturated resin and from peptide-saturated resin, respectively. Intensities from densitometric analysis of three independent experiments were normalized against biotine and mean values plotted with SEM error bars. The amount of peptide-bound MBP-E6 is proportional to the difference in intensity between the left and the right band. All LxxLL motifs bound significantly to MBP-E6 as compared to the two negative control peptides, CTRLR-1 and CTRLR-2.



## **Full Material and Methods**

### **Protein production :**

A fusion protein coding for the 10 first amino acids of the protein Paxillin (MDDLADALLAD ; AccNb. : P49023) connected via a short linker (KEGGG) to BPV-1 E6 protein (AccNb. : P06931), was PCR amplified and cloned via a three alanine linker at the C-terminus of a modified Maltose Binding carrier protein (MBP). D83, K84, K240, E360, K363 and D364 of MBP were mutated to Alanine for increasing MBPs propensity to crystallise, according to the works of Kobe et al., 1999<sup>24</sup> and Ullah et al., 2008<sup>23</sup>. BPV-1 E6 was also cloned into the pETM-41 vector generating a MBP-BPV-1 E6 fusion construct as previously described<sup>22</sup>. Both the MBP-paxillin-BPV-1 E6 and MBP-BPV-1 E6 fusion proteins were expressed in *Escherichia coli* BL21 DE3 cells in LB medium adjusted with 100  $\mu$ M ZnSO<sub>4</sub>. Cells were grown at 37°C until an OD<sub>600</sub> of 0.8. Recombinant protein expression was induced with 0.5 mM IPTG and expression was performed for 16 hrs at 16°C. Cells were harvested by centrifugation, lysed by sonication in buffer A (50mM Tris (pH 6.8), 400mM NaCl, 2mM DTT) containing 2.5  $\mu$ g/ml DNaseI ; 2.5  $\mu$ g/ml RNaseA and Roche Complete antiprotease cocktail (EDTA free), centrifuged at 18000 x g for 30 min and 4°C and filtered through 0.22  $\mu$ m filters. Cleared lysates were affinity purified on a 30 ml Amylose column (Amylose Agarose, New England Biolabs) in buffer A were centrifuged at 160000 x g and 4°C for 14 hrs and further purified on a HiLoad 16/60 Superdex 200pg Gel filtration using buffer A. Both fusion proteins eluted as a monomeric peak. MBP-paxillin-BPV-1 E6 was concentrated to 60 mg/ml in buffer A adjusted to 5 mM Maltose prior to crystallogenesis. MBP-BPV-1 E6 was stored in a diluted form since this construct is prone to self-oligomerisation and fiber formation<sup>22</sup> in the absence of the paxillin peptide.

### **Crystallisation, data collection and structure determination**

Crystallisation conditions were screened using commercially available kits (Qiagen) by the sitting-drop vapor-diffusion method in 96-well Innovaplate plates, employing a Tecan robot. After optimisation, crystals (200 x 250 x 100  $\mu$ m) suitable for X-ray diffraction grew in a few days in sitting drops made by mixing 200 nl of protein

solution at 30 mg/ml and 200 nl of reservoir solution consisting of 1.6 M AmSO<sub>4</sub>, 2% PEG 2000MME, 0.1 M HEPES Na Salt (pH 7.5). Drops were equilibrated against 50  $\mu$ l of reservoir solution at 290K. Crystals were flash-cooled in liquid nitrogen after a brief transfer in 5 ml of the reservoir solution supplemented with 20% (v/v) glycerol as a cryoprotectant and stored in liquid nitrogen.

X-ray diffraction data were collected using a single cryo-cooled (100K) crystal on beam line ID23 at the European Synchrotron Radiation Facility (ESRF). Best crystals diffracted beyond 2.3 Å and a complete data set was obtained between 40 Å and 2.29 Å from 360 frames of 0.5° oscillation (exposure time of 0.8s per oscillation with an attenuated beam, crystal-detector distance of 277.1mm, Q315 (ADSC) CCD detector). The data were indexed, processed and scaled with HKL2000<sup>33</sup>. The crystals belong to the space group C222<sub>1</sub> with unit cell parameters a=105.0Å, b=123.1Å c=94.6Å. The asymmetric unit contains one MBP-paxillin-BPV-1 E6 molecule, with a corresponding crystal volume per protein weight of 2.6 Å<sup>3</sup>/Da and a solvent content of 53% (assuming a partial specific volume of 0.74 ml/g). Data collection statistics are summarised in supplementary Table 1. The structure of chimeric MBP-paxillin-BPV-1 E6 molecule was solved by molecular replacement methods using the corresponding known MBP structure (pdb code 4MPB) as a probe with PHASER<sup>34</sup>.

The final model contains one copy of the fusion protein. Several regions of the polypeptide are not visible in the electron density map and are assumed to be disordered, namely residues : (i) the first three residues of MPB, (ii) the last four residues of the linker between paxillin and E6, (iii) the first ten residues and the last 7 residues of the E6 protein. Therefore, we have used the following nomenclature in the pdb file: chain A for MPB (residues 1 to 368)-linker (A<sub>369</sub>A<sub>370</sub>A<sub>371</sub>)-paxillin (M<sub>372</sub>D<sub>373</sub>D<sub>374</sub>L<sub>375</sub>D<sub>376</sub>A<sub>377</sub>L<sub>378</sub>L<sub>379</sub>A<sub>380</sub>D<sub>381</sub>)-linker (K<sub>382</sub>EGGG<sub>386</sub>), chain B (residues 1 to 137) for E6 protein. Using this convention, 8 residues of chain A (1 to 3, 382 to 386) and 17 residues from chain B (1 to 10, 131 to 137) are disordered. The current refined model has a crystallographic R-factor of 18.3% (R<sub>free</sub>=22.4%) at 2.29 Å resolution (see supplementary Table 1 for detailed statistics). Refinements were done with REFMAC<sup>34</sup> and PHENIX<sup>35</sup> and iterative models building with Coot<sup>36</sup>. All other crystallographic calculations were carried out with the CCP4 package<sup>34</sup>. The quality of the refined structures was assessed using MOLPROBITY<sup>37</sup>. The refinement statistics are

summarized in Table I. Surface charge potentials and molecular graphics figures were generated using PyMOL Version 1.3 (Schroedinger LCC).

### **Proteome-wide prediction of protein binders to E6 :**

The human and bovine proteome together with annotations were downloaded from Ensembl v58<sup>38</sup>. The proteomes were pre-screened with a regular expression representing a raw definition of the leucine-rich E6-binding motif. The regular expression was defined as follows:

$$\dots\text{LD.L}[\text{LFM}].. | ..[\text{DE}]\text{L}[\text{^E}]\text{L}[\text{LFM}]..$$

A dot represents any amino acid, amino acids in brackets are allowed at this position, ^ represents “not” and | represents the logical “or”. Matches of the regular expression were analysed for their propensities of being in disordered regions of the proteins using IUPred<sup>39</sup>. The mean IUPred scores of the motifs and, if possible of the 10 amino acids upstream and downstream were calculated. Motifs with at least one of these 3 mean scores < 0.4 were rejected.

The above procedure retained 387 potential BPV-1 E6-binding motifs, which were subsequently ranked using a score calculated with a Position Specific Scoring Matrix (PSSM) based on the work of Mount et al.<sup>26</sup>. The PSSM contains values for the likeliness of a residue to be part of the motif for each amino acid at each position of the motif. These values were determined by taking the logarithm of the frequency of occurrence of an amino acid at a particular position divided by the background frequency of this amino acid in the proteome. Frequencies of amino acids at motif positions were based on the mutagenesis experiments, the phage display data and observations based on the BPV-1 E6-LxxLL structure presented in this manuscript. Amino acids that do not appear at specific motif positions were given a very small pseudo count of  $10^{-10}$ . The scored motifs were further annotated for functional, structural and cellular location information using the Database for Annotation, Visualisation and Integrated Discovery (DAVID) v6.7<sup>40</sup>. A motif was marked for its occurrence in coiled-coil regions, if at least one of its residues was predicted to be part of coiled-coil<sup>41</sup>. Note that several proteins, such as PXN, AP1G1 or TGFB1I1, were found to contain several instances of potential BPV1 E6-binding motifs (Figure 5A).

## **Tandem Affinity and Mass Spectrometry Methods**

A triple tag composed of a Chitin Binding Domain (CBD) followed by a TEV protease cleavage site and a FLAG epitope was fused either to BPV-1 E6 or the LxxLL-E6 fusion that was used for crystallisation in this study. Both constructs were expressed in mammalian cells by vaccinia virus pTM1 transfections and used as tandem affinity reagents for capture of cellular proteins from C33A cell lysate. 25 confluent 15 cm plates of C33A cells (ATCC, Manassas, Virginia) per sample were washed three times with PBS, and lysed on ice with 0.5X NP40 lysis buffer (1X NP40 lysis buffer is 150mM NaCl; 50mM Tris pH 7.5; 50mM NaF; 5mM NaPPi; 1% IPEGAL CA-630; 0.01% phenylmethylsulfonyl flouride; 1mM sodium vanadate; 1ug/ml leupeptin/aprotinin). Lysates were centrifuged at 15000 X G for 10 minutes. Clarified lysates were pre-cleared on 3.0 ml columns of chitin-agarose (NEB). The flow-throughs were then incubated with 1-2 ug of CBD-TEV-FLAG-E6 or CBD-TEV-FLAG-LxxLL-E6 fusion proteins immobilised on chitin agarose beads with slow rotation for 1 hr at 4° C. The beads were washed three times with NP40 lysis buffer, twice with NP40 buffer without protease inhibitors, and bound proteins were released by cleavage with TEV protease (Invitrogen) in the manufacturer-supplied buffer supplemented with 0.1% IPEGAL CA-630. Eluted proteins were applied to FLAG antibody beads (Sigma) for one hour on ice, washed 3x with 1x NP40 Lysis buffer and eluted from the beads with 1% SDS, combined with an equal volume of 2x SDS PAGE sample buffer, and applied to 4-20% Novex polyacrylamide gels (Invitrogen), which were stained with silver to visualise protein bands<sup>42</sup>. The gel region from 30 kDa to 300 kDa from both lanes was divided into 5 segments each, and the resulting 10 gel segments processed for mass spectrometry: proteins were reduced, alkylated and digested with trypsin in the gel. The peptides formed in the digestion were extracted, concentrated, and characterised by capillary column LC-tandem MS. The mass spectrometer was operated in the data dependent mode throughout the HPLC gradient with an acquisition duty cycle consisting of an initial MS scan with a mass range of 300-2000 m/z<sup>43</sup> followed by an MS2 acquisition of the 5 most abundant ions<sup>44</sup>. All data were converted from the original Thermo RAW format to the mzXML format using the program ReAdW.exe (<http://tools.proteomecenter.org/ReAdW.php>), provided by the Institute for Systems

Biology. A concatenated DTA file suitable for searching with the OMSSA engine<sup>43</sup> was generated by MAZIE<sup>44</sup> an in-house Perl script that utilises a modified version of readmzXML.exe (<http://tools.proteomecenter.org/readmzXML.php>) to read the header and scan data from mzXML files. The MS data was searched with OMSSA as a tryptic digest against a composite database containing the human refseq database, acquired from NCBI at <ftp://ftp.ncbi.nih.gov/refseq> and the reversed sequence was generated by an in-house Perl script. The carbamidomethylation of cysteine (+57 Da) and oxidation of methionine (+16 Da) were treated as variable modifications. The tryptic search results for each data file were filtered by a 2% false discovery rate to determine significance<sup>45</sup>. Keratin peptides were excluded. All proteins that were identified via at least one significant peptide unique to the protein, and were associated with E6 but not with LxxLL-E6, were retained as potential binders of the peptide-binding pocket of E6. All names of putative E6 binders identified both by tandem affinity experiments and by bioinformatics predictions are coloured red or blue on Figure 5A, B and C.

#### **Phage Display :**

BPV-1 E6 fused to either glutathione S transferase (GST) or chitin binding domain (CBD) was expressed in monkey CV1 cells by vaccinia virus. A 6 cm plate of transduced cells was lysed on ice 18 hours after transduction in 200 ul 0.1X NP40 lysis buffer (1X NP40 lysis buffer contains 150mM NaCl; 50mM Tris pH 7.5; 50mM NaF; 5mM NaPPi; 1% IPEGAL; 0.01% phenylmethylsulfonyl fluoride; 1mM Sodium Vanadate; 1µg/ml Leupeptin/Aprotinin), 1% Bovine Serum Albumin (BSA) and 3 mM dithiothreitol (DTT), centrifuged at 14000 rpm in a microfuge, and the soluble fraction bound to either glutathione agarose or chitin-agarose beads that had been pre-blocked in lysis buffer overnight. Beads were washed 3 times with 1 ml lysis buffer and then twice with 1 ml sterile Luria Broth supplemented with 1% BSA 0.01% IPEGAL and 3 mM DTT (binding buffer). 10<sup>-11</sup> phages from a commercial phage display library (PhD.-12, New England Biolabs), which were blocked overnight in binding buffer were added to the beads, rocked for 4 hrs at 4 °C, washed 6 times with 1 ml binding buffer, once with saline and then phage were eluted with 100 ul 0.2M Glycine pH 2.2, neutralised with 10 ul Tris-Glycine pH 9. Eluted phage were amplified by infection of XL1-Blue-MRF<sup>7</sup> bacteria (Stratagene). Four rounds of selection and amplification were performed with

alternation of GST-E6 and CBD-E6-coated selection beads. Plaques were isolated and DNA sequencing of the gene III-peptide fusion DNA was performed.

**Yeast two-hybrid interaction assays :**

Modified LexA-based yeast two-hybrid assays were performed as previously described<sup>46</sup>.

**Pull down interaction assays :**

BPV-1 E6 and E6 mutants were expressed *in vitro* in rabbit reticulocyte lysate, incubated in 1x NP40 lysis buffer together with 0.1 to 1 ug GST or GST fused to amino acids 1-11 of paxillin or mutated paxillin peptides fused to GST as previously described<sup>9</sup>. Beads with bound E6 were washed 3 times with 1 ml 1x NP40 binding buffer and E6 then eluted by incubation in SDS-PAGE sample buffer, resolved on 15% SDS poly-acrylamide gels, and electrotransferred onto PVDF membranes, stained with Ponceau red. Radiolabeled E6 was visualised and quantitated with a Molecular Dynamics Imaging station.

**Holdup interaction assays:**

To monitor *in vitro* the interaction of BPV-1 E6 with various LxxLL sequences and evaluate their relative affinities, we used a modified version of the " holdup" comparative chromatographic retention assay<sup>32</sup>, based on avidin-biotin affinity.

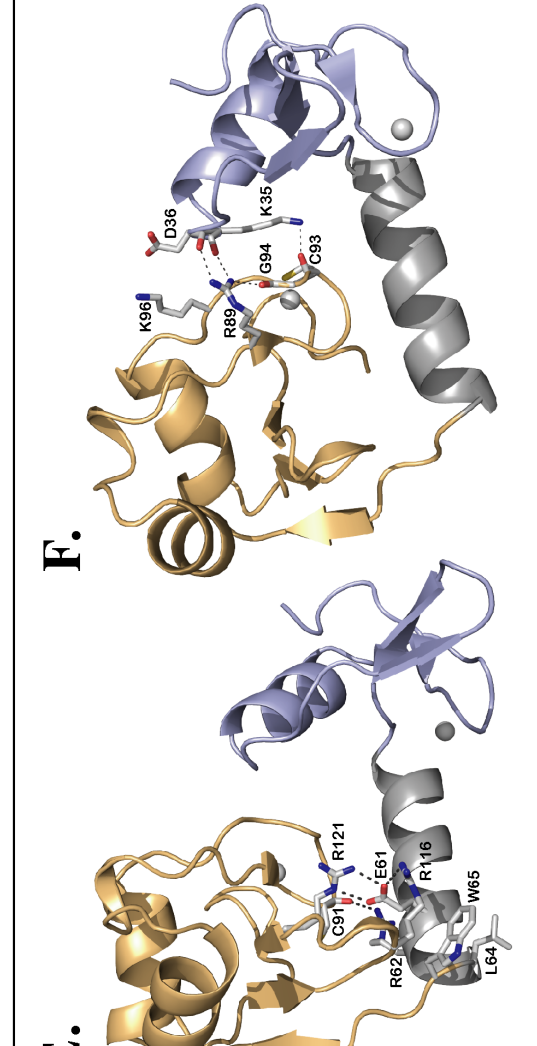
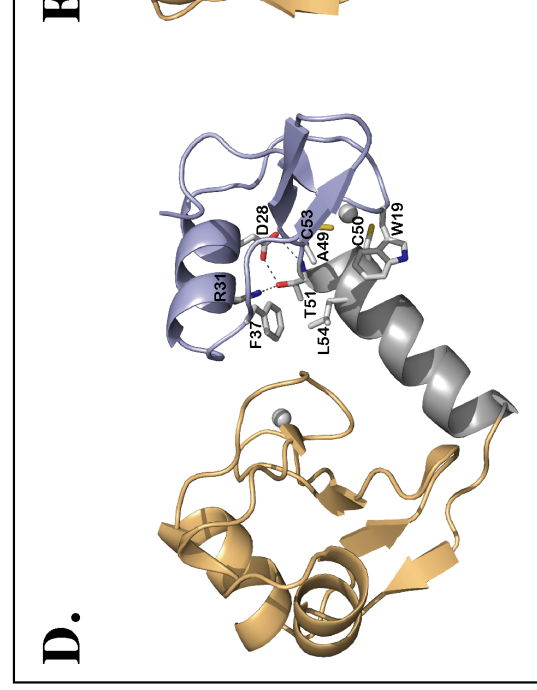
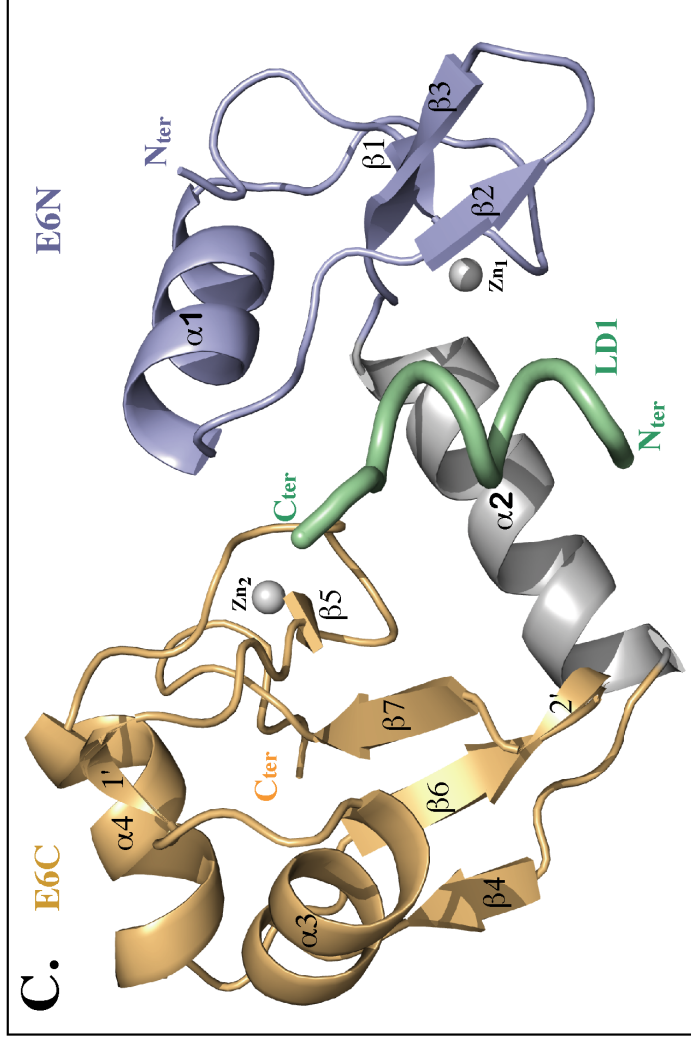
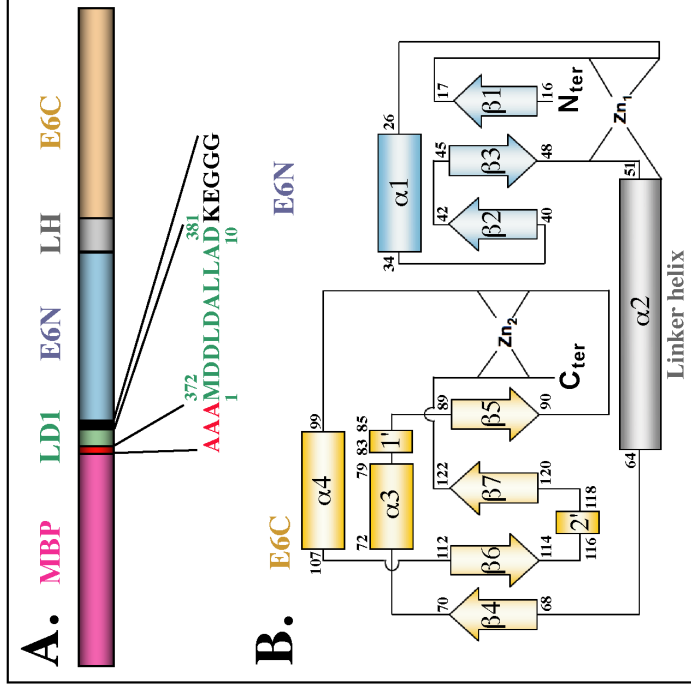
14 biotinylated 12-meric LxxLL peptides derived from human proteins, which were predicted and experimentally validated as BPV-1 E6-binders (Fig. 5) were chemically synthesised (JPT Peptide Technologies, Berlin). These peptides present the following names and sequences, respectively: MED13 (YTDLDNLFNSDE); PXN\_1 (MDDL DALLADLE); TGFB1\_1 (MEDLDALLSDLE); MAML1 (MSDLDDL LGSQ); p53 (FSDLWKLLPENN); TGFB1\_2 (LCELDRL LQELN); AP1B1 (GGGLDSL MGDEP); IRF3 (EDILDELLGNMV); PXN\_2 (LSELDRL LLELN); PXN\_4 (TRELDEL MASLS); TGFB1\_3 (TLELDRL MASLS); AP1G1\_1 (GGELLDLLGDIN); EBP1 (DAELKALLQSSA); AP1G1\_2 (ANDLLDLLGGND). Two additional biotin peptides were also synthesised as negative controls: 16E6-CT (SSRTRRETQL) corresponding to the PDZ-binding C-terminus of HPV16 E6; and CTRL (GSNSGNGNS), which represents an irrelevant non-charged polar sequence. Each biotin peptide was resuspended in PBS buffer at a 1 mM concentration then

incubated with 20  $\mu$ l of high-performance streptavidin sepharose (GE-Healthcare). The peptide-saturated resin was further incubated with 2 mM biotin in order to saturate all streptavidin sites. We also prepared an additional batch of control resin containing only biotin. The resulting concentration of biotin peptide (estimated from the manufacturer's resin specifications) within the wet bedded resin approximates 150-200 mM.

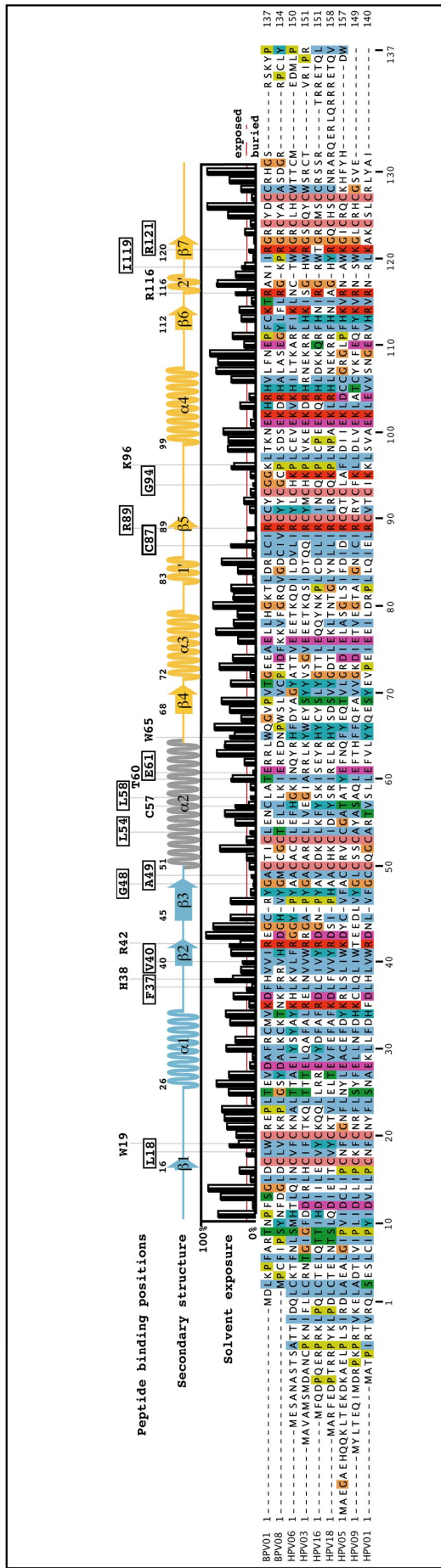
15  $\mu$ l of peptide-saturated or biotin-saturated bedded resin were then mixed and incubated for 15 min with 30  $\mu$ l of freshly purified monomeric MBP-BPV-1 E6 (devoid of the LxxLL linker). Each resin-protein mixture contained an approximate concentration of 50  $\mu$ M of peptide and 4  $\mu$ M of MBP-E6 protein. The resin-protein mixtures were finally transferred into a centrifugeable filtrating device and the liquid phases were rapidly extracted. An equal volume of each flow-through was run on SDS PAGE. The amount of MBP-E6 bound to each peptide can be quantified by subtracting the amount of MBP-E6 found in the flow-through of the peptide-saturated resin from the amount of MBP-E6 found in the flow-through of the biotin-saturated resin.

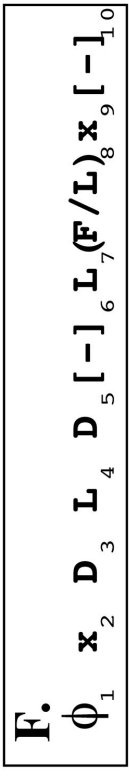
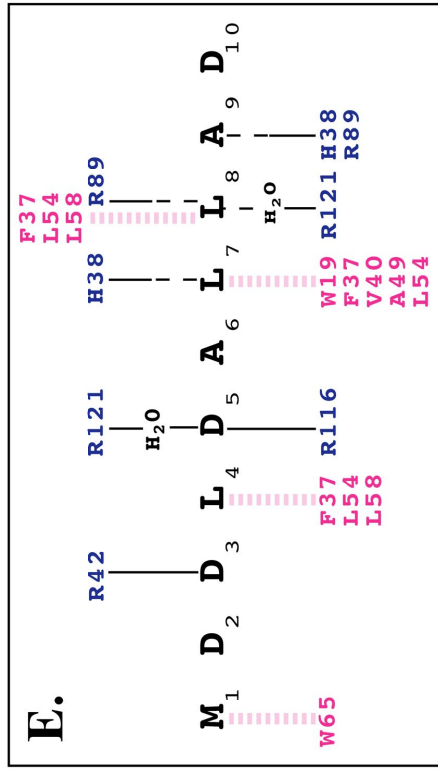
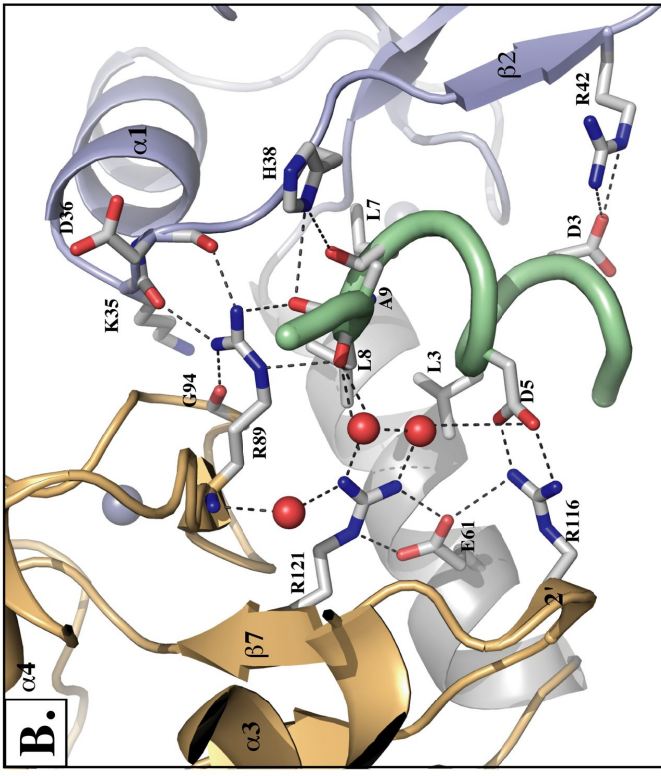
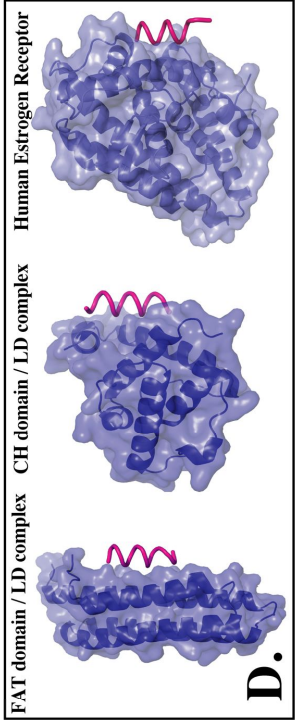
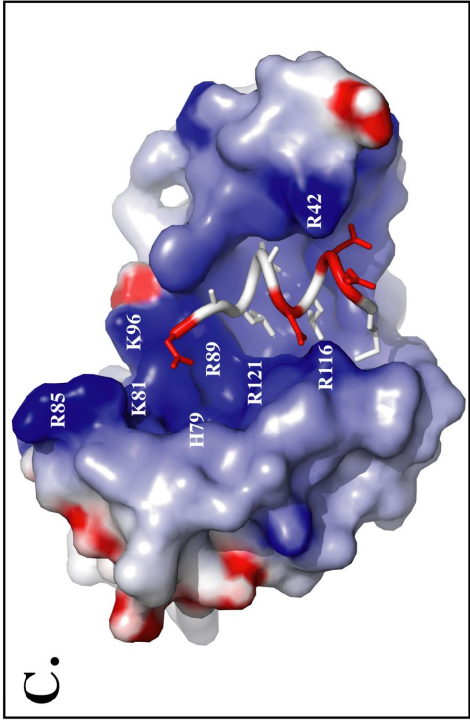
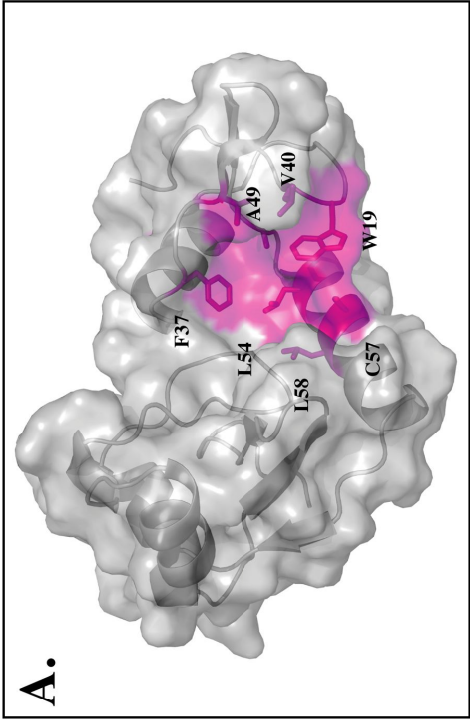
#### **Cell culture :**

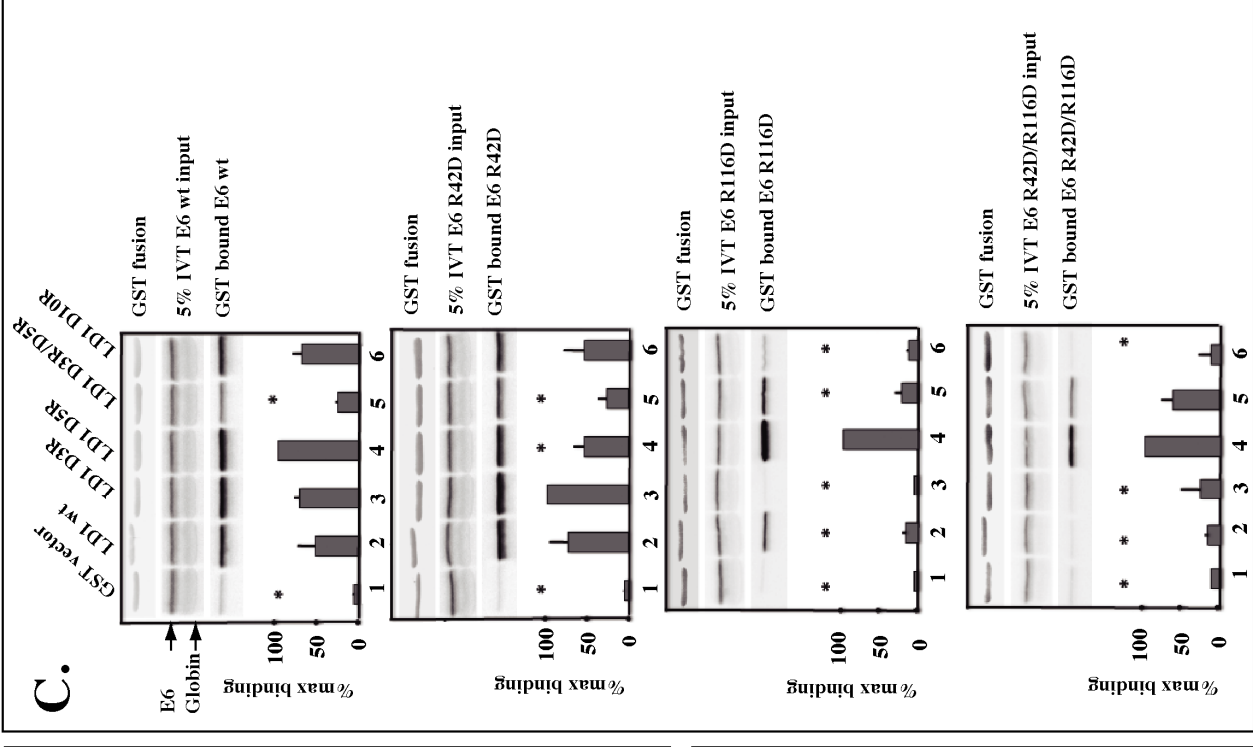
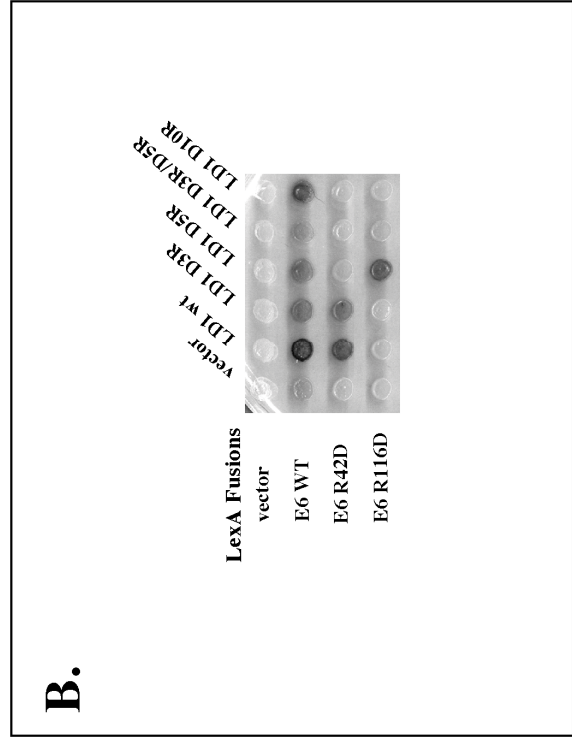
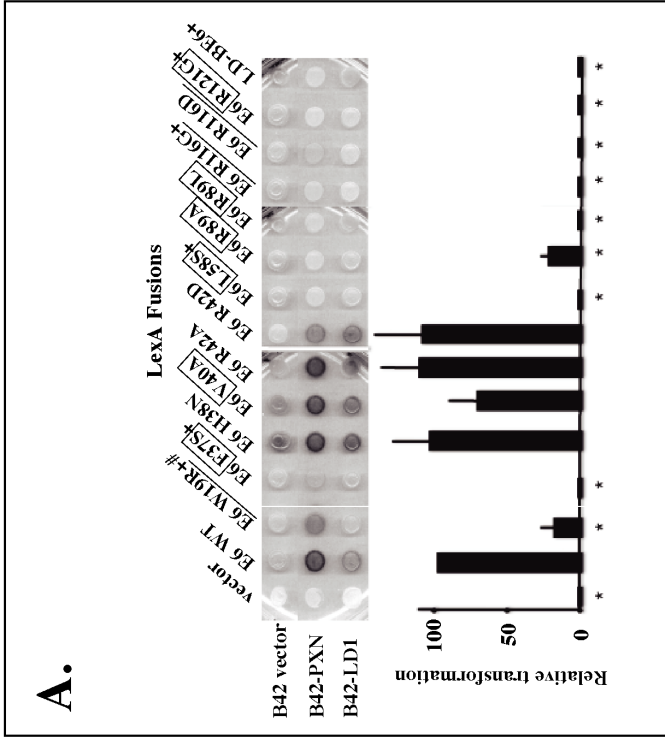
Murine C127 cells (from ATCC, Manassas, Virginia) cultured in DMEM supplemented with 10% fetal bovine serum were transduced with BPV-1 E6 and E6 mutants packaged as replication defective murine retroviruses using the vector pBabepuro<sup>47</sup>. Puromycin resistant cells were passed four times in drug containing media before seeding into agarose to assay for anchorage independent colony formation as previously described<sup>9</sup>.







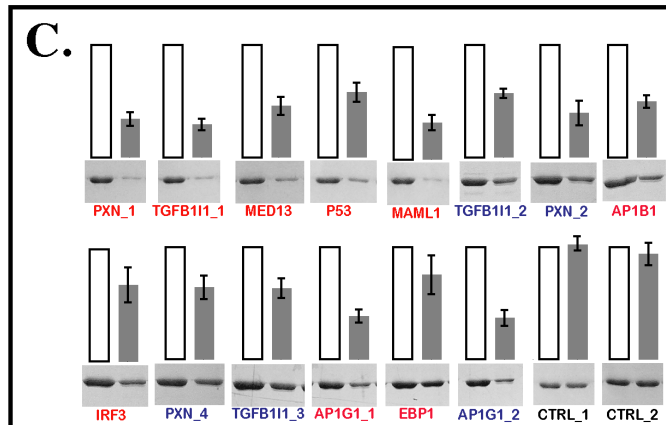
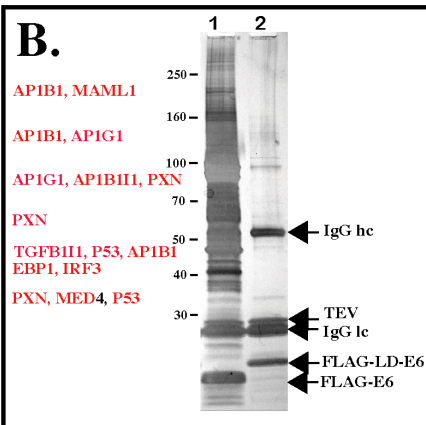




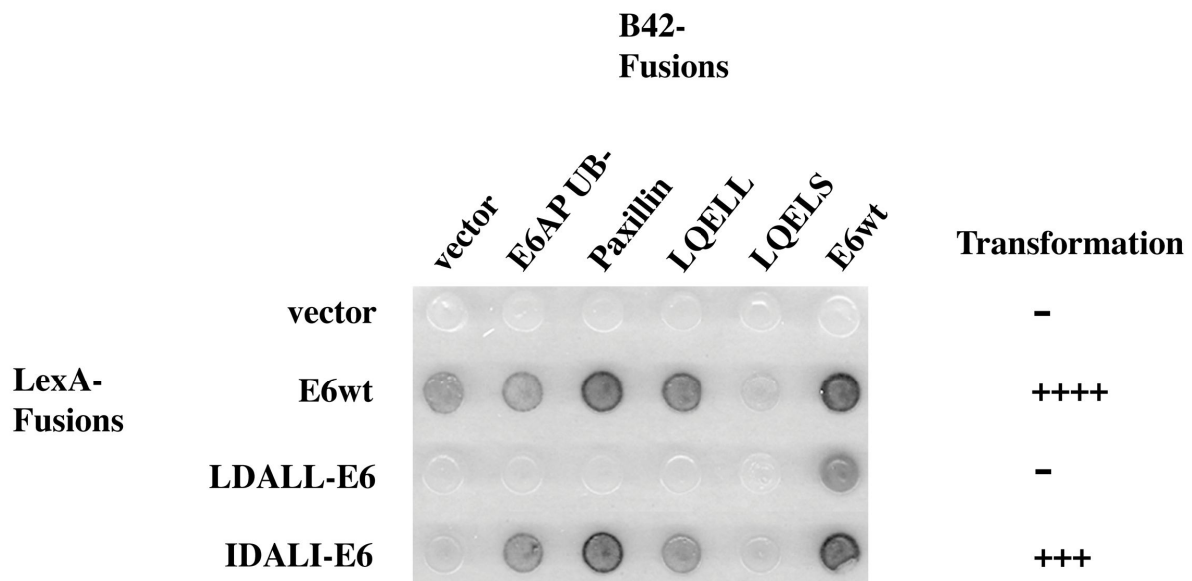
**A.**

Distribution of hits from proteomic screen in list of predicted binders

n° prediction (out of 387)	target protein	motif-containing sequence	localization in sequence	net charge	LD-motif	regulation of transcription	Epithelium	cytoskeleton
1	<b>paxillin (PXN_1)</b>	<b>MDDL</b> DALLAD LESTT	1-11	-4	yes		x	x
2	ninein	QKRLS <b>WDKLDHLMNE</b> EQQLL	1838-1848	-1	no			x
3	<b>TGFβ1 induced transcript 1 (TGFB11_1)</b>	<b>MEDLDALLSD</b> LETTT	1-11	-4	yes	x		x
4	<b>mediator complex subunit 13 (MED13)</b>	DLAVS <b>YTDLDNLFNS</b> DEDEL	778-788	-2	no	x	x	
5	<b>tumor protein p53 (P53)</b>	LSQET <b>FSDLWKLLPE</b> NNVLS	19-29	-1	no	x		
7	<b>mastermind-like 1 (MAML1)</b>	TSEEW <b>MSDLDLLGS</b> Q	1006-1016	-3	no	x		
8	<b>TGFβ1 induced transcript 1 (TGFB11_2)</b>	VLGTG <b>LCELDRLLOE</b> LNATQ	73-83	-2	yes	x		x
9	amyloid beta (A4) precursor protein-binding family B member 1 interacting protein	LNAL E <b>DQDLDALMAD</b> LVADI	61-71	-4	yes			x
14	meningioma 1	KSAMS <b>TIDDSLMAE</b> HSAAW	1213-1223	-3	no		x	
19	Decidual protein induced by progesterone	PMADT <b>VDPLDWLFGE</b> SQEKQ	102-112	-3	no			
20	golgi-associated gamma adaptin ear containing ARF binding protein 3	SALHH <b>LDALDQLLEE</b> AKVTS	516-526	-4	no			
29	<b>paxillin (PXN_2)</b>	SLGSN <b>LSELDRLLE</b> LNAVQ	142-152	-2	yes		x	x
43	kinesin light chain 2	KGDVP <b>KDTELDLFPN</b> EDEQS	161-171	-2	no		x	x
44	spindlin family member 2B	ITQWK <b>GTVLDQLLDD</b> YKEGD	70-80	-3	no			
58	consortin connexin sorting protein	CGNNQ <b>ISDLGILLPE</b> VCMAP	516-526	-2	no			
64	protocadherin 10	HSTLE <b>RKELDGLLTN</b> TRAPY	1015-1025	0	no			
70	<b>adaptor-related protein complex 1. beta 1 subunit (AP1B1)</b>	AVDLL <b>GGGLDSLMD</b> EPEGI	660-670	-2	no			
79	<b>interferon regulatory factor 3 (IRF3)</b>	TSDTQ <b>EDILDELLGN</b> MVLAP	137-147	-4	no	x		
90	<b>paxillin (PXN_4)</b>	SASSA <b>TRELDLMS</b> LSDFK	263-273	-2	yes		x	x
93	<b>TGFβ1 induced transcript 1 (TGFB11_3)</b>	SATSA <b>TLELDRLMS</b> LSDFR	136-146	-1	yes	x		x
95	clathrin interactor 1	PAASN <b>SSDLFDLMS</b> SQATM	402-412	-2	no		x	
97	eukaryotic translation initiation factor 4E nuclear import factor 1	QKA <b>KVDLKPFLSS</b> LSANK	4-14	1	no			
105	centrobin centrosomal BRCA2 interacting protein	QQVAE <b>DYELRLLLLD</b> PPAPG	520-530	-2	no			x
113	<b>adaptor-related protein complex 1. gamma 1 subunit (AP1G1_1)</b>	KPSSA <b>GGELLDLLGD</b> INLTG	651-661	-3	no			
135	<b>proliferation-associated 2G4 (EBP1)</b>	EMEVQ <b>DAELKALLQS</b> SASRK	349-359	-1	no	x	x	
262	<b>adaptor-related protein complex 1. gamma 1 subunit (AP1G1_2)</b>	QPTSQ <b>ANDLLDLLGG</b> NDITP	626-636	-2	no			



## Supplemental Fig. 1

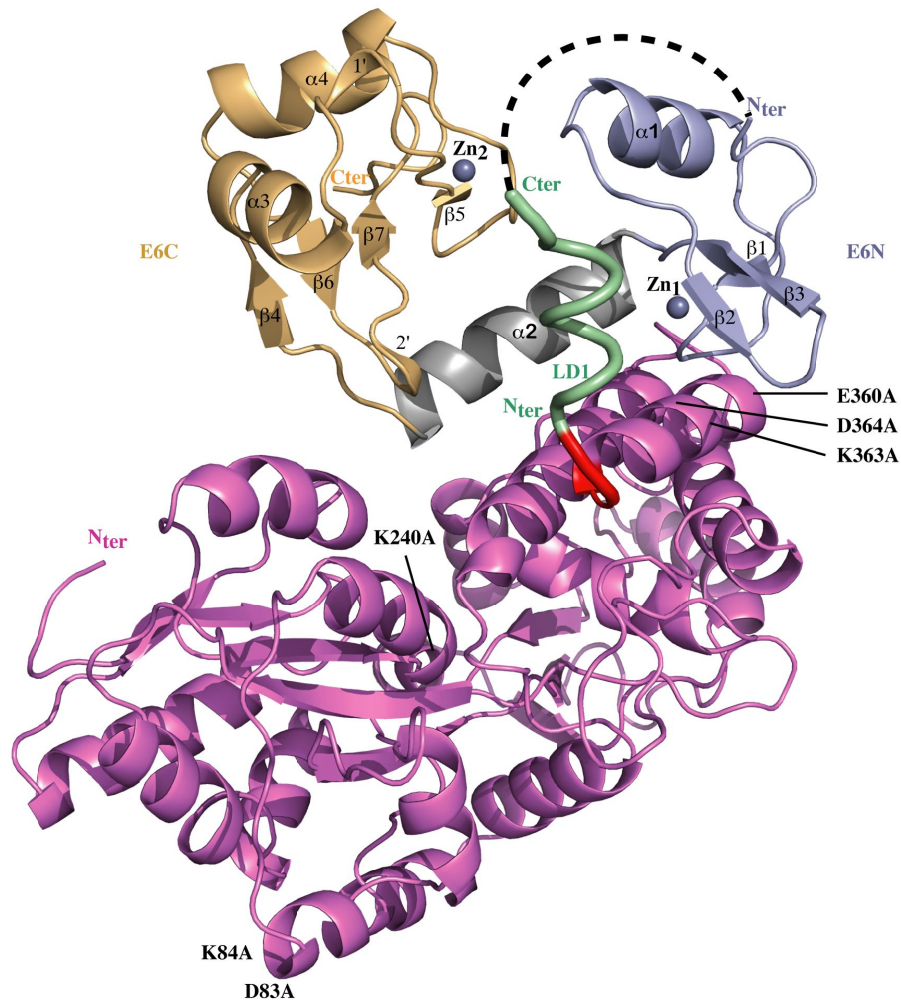


LDALL-E6: LexA-----LXXLL-----linker---E6

LDALL-E6: ...DWLEF-MDDL DALLAD-KEGGG-MDLK...

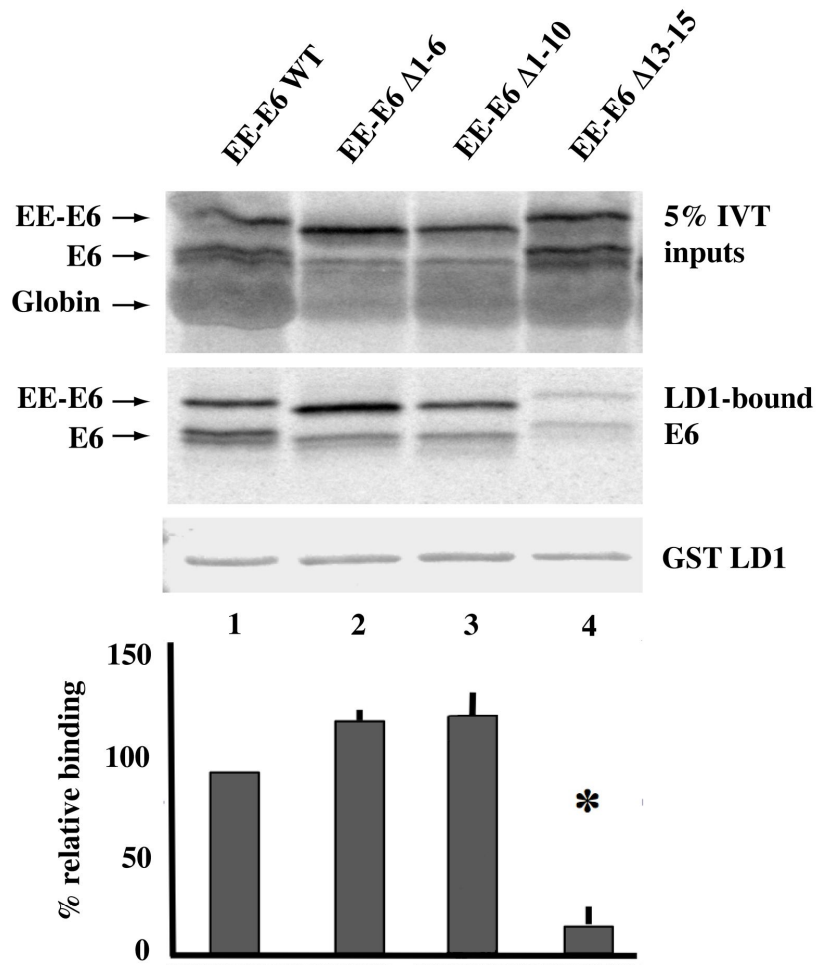
**Fusion of a LXXLL peptide to the amino-terminus of E6 blocks the interaction of E6 with paxillin and E6AP.** A LxxLL motif from the first 10 amino acids of paxillin was fused to the amino-terminus of BPV-1 E6 with a short glycine linker and expressed as a fusion to the *lexA* DNA binding domain as illustrated at the bottom, and is denoted as the *lexA* fusion LDALL-BE6. Mutations within the LD1 motif (shown as IDALI) were also expressed as fusions to *lexA* and E6 as illustrated in the figure. *LexA* alone, *lexA* fused to wild-type E6 and the indicated LxxLL fusions were expressed in the yeast reporter strain EGY48 containing a *lexA* responsive *lacZ* reporter gene. The B42 synthetic transactivator domain fused to full length E6AP mutated in ubiquitin ligase activity (E6AP\_Ub-), full length paxillin or the isolated LxxLL peptide of E6AP that interacts with E6 (LQELL) or a mutated E6AP LxxLL peptide (LQELS) were introduced to the reporter strains by mating. Interactions between *lexA* fusions and B42 fusions are visualised 18 hours after patching of equal numbers of mated and selected diploid yeast onto galactose plates containing XGAL. While fusion of LDALL peptide to E6 blocks the interaction of E6 in vivo with paxillin, E6AP, or LxxLL peptides in trans, mutation of the peptide to IDALI restores interactions in vivo. Transformation of murine fibroblasts by E6, LDALL-E6, and IDALI-E6 are previously published results. The white horizontal line indicates the deletion of irrelevant samples from the same plate.

## Supplemental Fig. 2



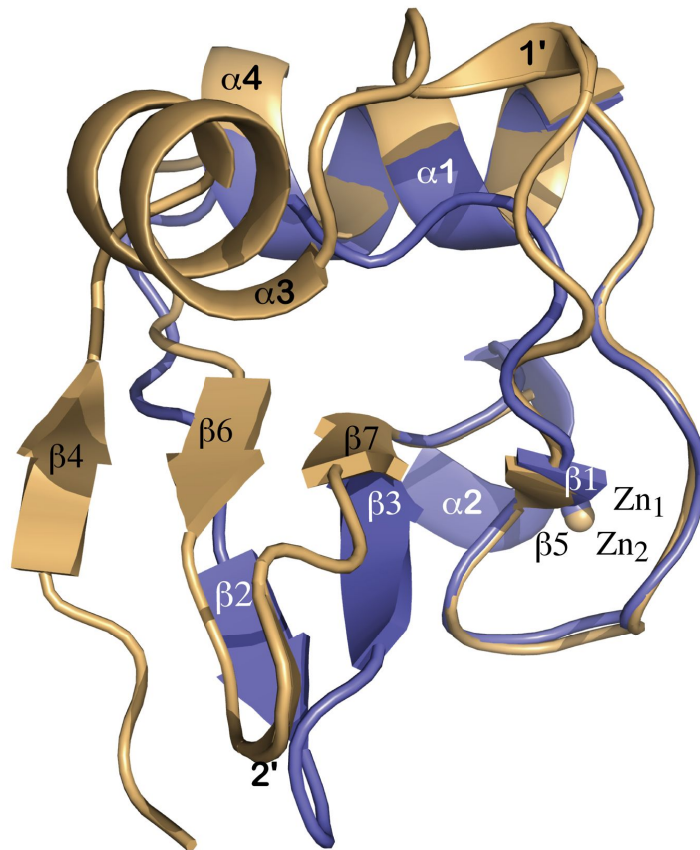
**The crystal structure of full length MBP-LxxLL-BPV-1 E6 construct .** Pink : Maltose Binding Protein (MBP) carrier protein. Crystallisation prone mutations D83A, K84A, K240A, E360A, K363A, D364A are indicated in black. Red : short three alanine linker, which adopts half of a helical turn at the end of the C-terminal alpha-helix of MBP. Green : E6 binding peptide, bearing the 10 N-terminal amino acids of paxillin. This peptide contains an LxxLL sequence and is part of the LD1 motif in full-length paxillin. Black dotted lines : Residues 1-10 of BPV-1 E6. These residues could not be placed into the electron density map, probably due to flexibility in this particular region. Light blue : N-terminal zinc-binding domain of E6 (E6N). The zinc atom is labeled Zn<sub>1</sub>. Grey : Linker helix between the two zinc binding domains of E6. Gold : C-terminal zinc-binding domain of E6 (E6C). The zinc atom is labeled Zn<sub>2</sub>.

Supplemental Fig. 3



**Replacement of ten amino-terminal amino acids of BPV-1 E6 by an EE-tag do not ablate interaction with the paxillin LD1 LXXLL motif in vitro.** A GST pull-down was performed using GST-fused wild-type paxillin LD1 motif and <sup>35</sup>S-labeled in vitro translated BPV-1 E6 constructs. All E6 constructs for this experiment were fused to a N-terminal EE-tag (EEEEYMPME). Residues 1-6 (<sub>1</sub>MDLKPF<sub>6</sub>), 1-10 (<sub>1</sub>MDLKPFARTN<sub>10</sub>) and 13-15 (<sub>13</sub>SGL<sub>15</sub>) of E6 were deleted, resulting in the mutants EE-E6Δ1-6, EE-E6Δ1-10 and EE-E6Δ13-15, respectively. In the EE-E6Δ1-6 and EE-E6Δ1-10 constructs, the deleted N-terminal residues of E6 may be partially or fully replaced by residues of the EE tag. In vitro transcribed and translated <sup>35</sup>S-labeled and EE-tagged E6 and E6 N-terminal deletion mutants were assayed for association with the first 10 amino acids of paxillin (the LD1 LxxLL motif) fused to GST. Histogram : Quantification of three separate binding assays was performed and the mean results with SEM error bars were plotted. Asterisk denotes binding differences from wild-type BE6 with P < 0.05.

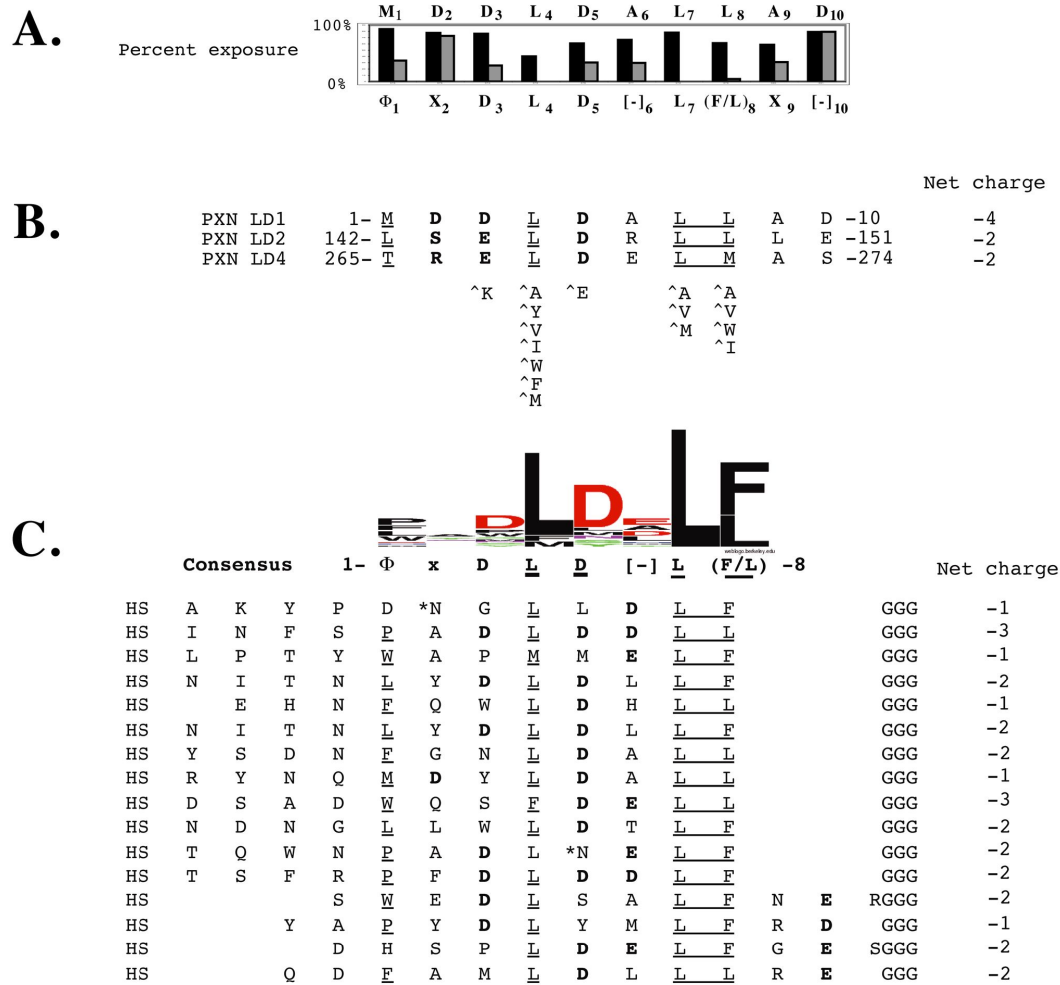
## Supplemental Fig. 4



**Comparison of the E6N and the E6C fold.** The BPV-1 E6N and E6C domain were aligned on the residues 13-35 on E6N and 86-108 and the structure coordinates of the two domains superimposed. The overall topology is conserved and the structures are comparable with a low RMSD except for the missing N-terminal residues corresponding to residues 65-83 on the E6C (beta strand 4 and helix 3). Lightblue : E6N domain ; Gold : E6C domain. Secondary structure elements are labeled white in case of E6N and black in case of E6C.



## Supplemental Fig. 5



### Identifying key positions of the LxxLL motif that are critical for recognition of BPV-1 E6.

A. Exposure plot of the LxxLL peptide : The percentage of solvent exposure of each peptide residue was calculated using the PyMol program in the presence (grey bar) or the absence (black bar) of the bound BPV-1 E6 protein, and plotted against the peptide sequence (above) and the consensus sequence derived from the phage display (bottom). The residues exhibiting a higher increase in exposure in the absence of bound E6 are the residues most involved in binding. B. Sequence alignment of the three BPV-1 E6 binding LD1 motifs of paxillin. The alignment reveals the preference for a hydrophobic residue at position 1, a negative charge at position 3 and 5, strongly conserved Leucine residues at position 4, 7 and 8 and a negative charge at position 10. Residues listed on the bottom of this figure labeled with a « ^ » are residues that can be excluded according to this alignment. C. Phage display results. 16 phage peptides selected after several rounds of phage display vs chitin resin-bound BPV-1 E6 are listed and aligned. The net charge of each peptide is plotted at the right and varies between -1 and -3. The resulting consensus sequence is plotted on top of the alignment. The coloured consensus plot shown above the consensus was created with WebLogo (<http://weblogo.berkeley.edu/logo.cgi>). The consensus features of the selected phage peptides are extremely similar to those of the E6-binding LD motifs of paxillin.

---

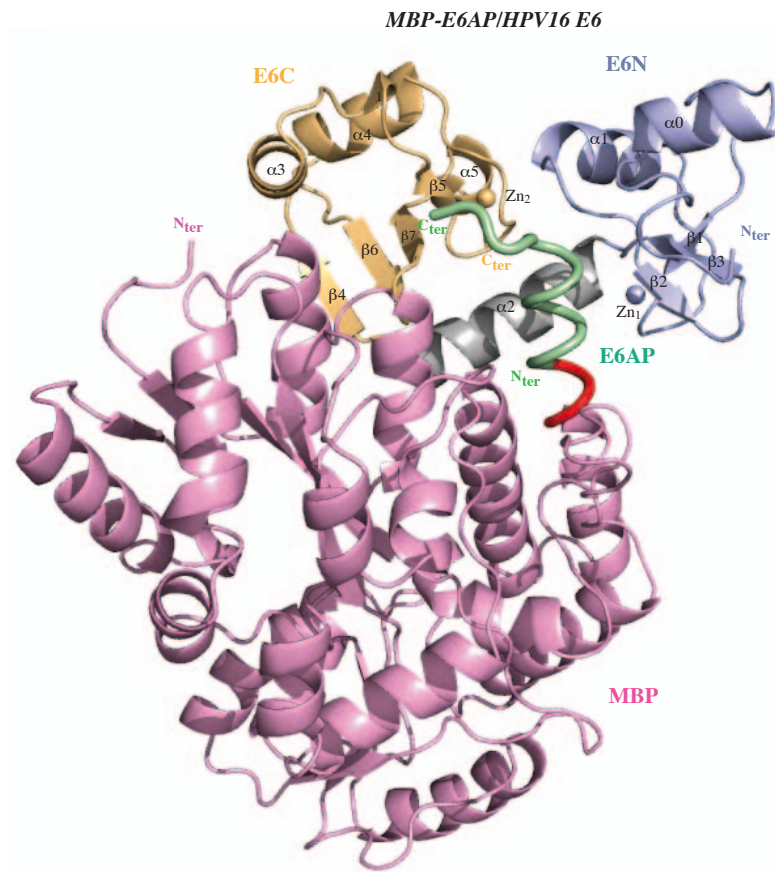
# Annexe

## 2 Figures supplémentaires

Tableau 1 : récapitulatif de données du cristal HPV16 E6/E6AP

	HPV16 E6/E6AP
<b>Data collection</b>	
Space group	P 2 <sub>1</sub> 2 <sub>1</sub> 2 <sub>1</sub>
Cell dimensions	
<i>a, b, c</i> (Å)	106.28 134.94 138.80
<i>α, β, γ</i> (°)	90 90 90
Resolution (Å)	50 – 2.55 (2.59 – 2.55)
<i>R</i> <sub>sym</sub>	8.2 (53.7)
<i>I</i> / <i>σI</i>	31.0 (2.9)
Completeness (%)	99.8 (99.1)
Redundancy	7.2 (6.3)
<b>Refinement</b>	
Resolution (Å)	37 – 2.55
No. reflections	65602 (4811)
<i>R</i> <sub>work</sub> / <i>R</i> <sub>free</sub>	16.47/19.58 (20.38/24.48)
No. atoms	9036
Protein	8326
Ligand/ion	134/4
Water	572
<i>B</i> -factors (isotropic)	
Protein	52.4
Ligand/ion	59.1/40.5
Water	53.1
<b>R.m.s. deviations</b>	
Bond lengths (Å)	0.01
Bond angles (°)	1.08

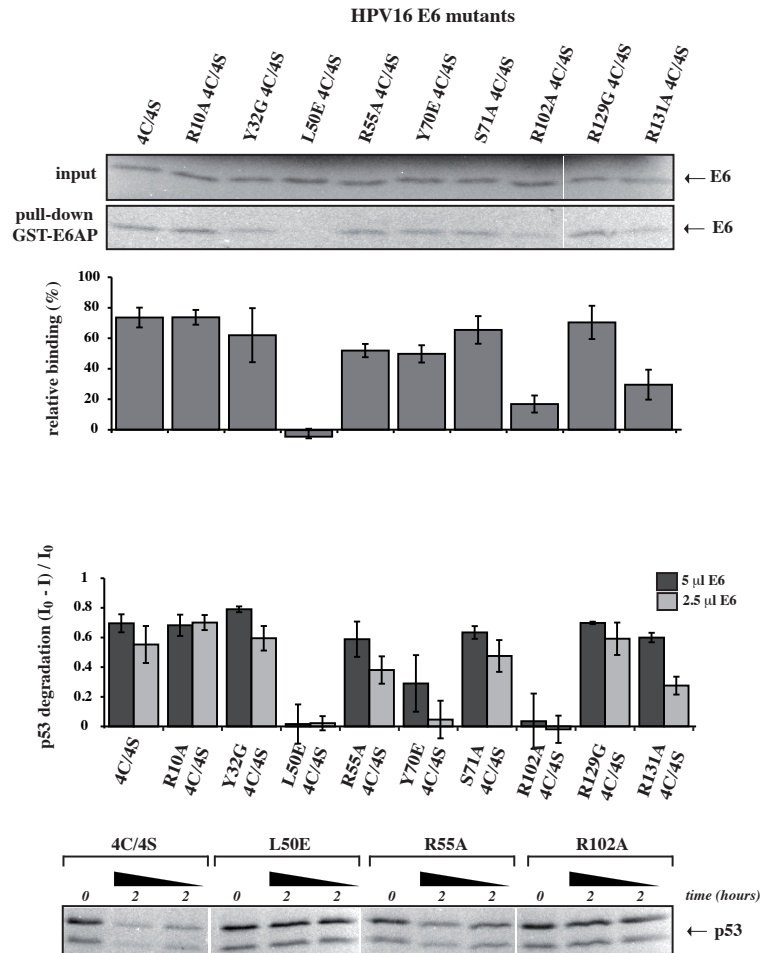
**Figure 1**



Structure cristallographique du complexe triple entre la protéine solubilisante et cristallisante MBP (rose), HPV16 E6 (Violet [E6N] , Jaune [E6C] et gris [région inter-donmajnes]) et le peptide E6AP (Vert) lié par un court "linker" de 3 alanines (rouge) au C-terminale de MBP.

**Figure 1 : Structure 3D du complexe MBP-E6AP/ HPV16 E6**

Figure 2



En haut : Interaction des mutants de E6 avec le peptide de fusion GST-E6AP (E6AP : E401SSELTLQELLGEEER418). En bas : Capacités des mutants à dégrader p53 (incubation de 2 µl de 35S p53 avec 2 ou 5 µl de E6) (Manuscrit en cours de rédaction : Katia zanier et al.).

Figure 2 : Activités biologiques des mutants de HPV16 E6 (Dégradation de p53 et liaison de E6AP)



**Résumé :**

L' oncoprotéine E6 - qui joue un rôle crucial dans le processus d'oncogenèse induit par les papillomavirus - a longtemps résisté à toute analyse. Depuis 1995 l'équipe Oncoprotéines a concentré ses efforts sur cette problématique. Ce qui a permis la résolution par RMN de la structure du domaine C-terminal de E6 en 2006. C'est dans ce cadre que j'ai commencé ce Doctorat en 2008, avec objectif de continuer la quête de données structurales sur E6 tout en acquérant des informations sur ses modes d'interaction avec ses cibles cellulaires. Les travaux de cette thèse ont permis l'obtention de la structure cristallographique de E6 (HPV16) en complexe avec un peptide de E6AP, en utilisant une approche originale capable de produire des protéines E6 stables et solubles. Cette structure constitue la première information structurale publiée sur des protéines E6 entières, attendue depuis plus de 20 ans par la communauté scientifique. J'ai effectué également durant cette thèse une analyse du système d'interaction de la protéine E6 basée sur une large étude d'interaction entre les protéines E6 (7 types) et 93 peptides porteurs de motif LxxLL.

**mots-clés :** papillomavirus, oncoprotéines E6, cancer du col de l'utérus, cristallographie, optimisation de production de protéines, interaction protéique, protéine E2, protéine Brd4.

**Abstract :**

The oncoprotein E6 which plays a crucial role in the process of carcinogenesis induced by HPV, has withstood all tests for a long time. Since 1995 the team « oncoproteins » has focused on this issue. This allowed the resolution of structure of C-terminal domain of E6 by NMR analysis, in 2006. In this context, I started my PhD in 2008 with aim to continue the pursuit of structural data on E6 while also acquiring information about its modes of interaction with its cellular targets. The work of this thesis has enabled us to obtain the crystal structure of E6 (HPV16) in complex with a peptide of E6AP, using an original approach capable of producing stable and soluble proteins E6. This structure represents the first structural information on full-length E6, awaited for over 20 years by the scientific community. I also performed during this thesis an analysis of the interaction system of the E6 protein based on a large study of interaction between proteins E6 (7 types) and 93 peptides bearing LxxLL motif.

**keywords :** papillomavirus, E6 oncoproteines, cervical cancer, crystallography, optimization of protein production, protein interactions, E2 protein, Brd4 protein.