Docteur de l'Université Louis Pasteur Strasbourg I

par

Sebastian M.C. CHARBONNIER

Étude structurale et cinétique des interactions entre l'oncoprotéine E6 du virus du papillome humain et les domaines PDZ
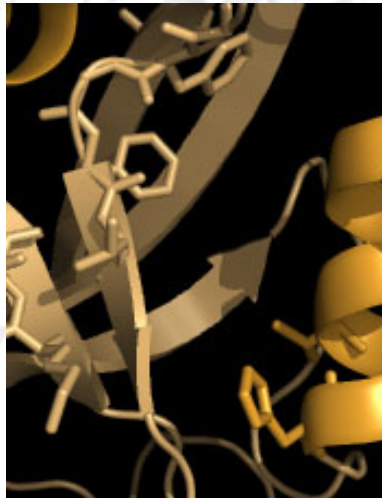
Soutenue publiquement le 28.11.2006

Membres du jury

Directeur de Thèse : **Dr Gilles Travé**

Rapporteur Interne : **Dr Arnaud Poterszman**

Rapporteur Externe : **Prof Dr Lawrence Banks**

Rapporteur Externe : **Prof Dr Arnt Raae**

Examinateur : **Prof Dr Bruno Kieffer**

REMERCIEMENTS

Je tiens à remercier les membres de mon jury de thèse : le Professeur Lawrence Banks, le Professeur Arnt Raae, le Dr. Arnaud Poterszman ainsi que le Professeur Bruno Kieffer d'avoir bien voulu me faire l'honneur de juger mon travail. Je remercie particulièrement les Professeurs Lawrence Banks et Anrt Raae pour avoir fait le long voyage d'Italie et de Norvège, respectivement, pour assister à ma soutenance.

Je remercie sincèrement le Professeur Claude Kédinger pour m'avoir accueilli au sein de l'UMR 7175-LC1, anciennement UMR 7100, pour l'intérêt qu'il a toujours porté à mon travail ainsi que pour son soutien au cours des presque 5 années de mon séjour.

Je souhaite dédier ce paragraphe au Dr. Gilles Travé mon directeur de thèse. Gilles, je te remercie du fond du cœur pour la confiance que tu m'as accordée durant ces 4 ans et demi de chemin commun. Je te serais toujours reconnaissant pour ton excellent encadrement et ta disponibilité continue. Je te dois la majorité de mes connaissances en matière de biochimie et de biophysique. C'est grâce à toi que j'ai découvert la fascination pour le détail atomique d'une structure de protéine, fascination qui ne m'a pas lâchée depuis. En outre, j'ai également bénéficié et appris énormément concernant la rédaction scientifique. Je n'oublierais jamais les nuits blanches ou tu m'as assisté en chambre froide ou les longues nuits de rédaction de publications. Mais je te remercie surtout de m'avoir fait découvrir ta façon parfois non-conventionelle de voir les choses, qui souvent a été source de créativité et d'inspiration. Merci également de m'avoir fait voyager et découvrir la façon de travailler dans des laboratoires allemands et norvégiens. Ces expériences à tes côtés m'ont certainement

ouvert l'esprit et m'ont beaucoup fait évoluer dans ma façon de travailler et de réfléchir. Je te suis aussi reconnaissant pour ton optimisme chronique qui m'a toujours motivé et qui m'a appris à ne pas baisser les bras après un échec, mais plutôt de le voir comme un défi. Bien que nos relations professionnelles étaient très proches, tu m'as également appris à prendre des initiatives et finalement à devenir indépendant. Je te remercie encore une fois pour l'énergie et la confiance que tu as investies en moi et comme tu me reprochais d'être trop modeste, je déclare que je pense ne pas t'avoir déçu à ce sujet. Je ne peux parler que pour moi, mais je ne regrette aucune minute que j'ai passé en tant que stagiaire ingénieur ainsi que doctorant à tes côtés et tiens à maintenir nos relations autant professionnelles que personnelles également après la fin du doctorat.

J'exprime ma gratitude envers le Dr. Andrew Atkinson pour avoir toujours travaillé avec moi sur la structure de MAGI-1 PDZ1. Merci de m'avoir initié à l'utilisation du spectromètre RMN et des logiciels d'attribution. Merci également pour toutes les réponses à mes nombreuses questions et pour n'avoir jamais perdu patience.

Je remercie le Dr. Georges Orfanoudakis pour son soutien en tant que directeur de l'équipe Oncoprotéines et pour les discussions variées et très stimulantes que nous avons pu avoir au long des années que j'ai passé au laboratoire.

Je souhaite faire un commentaire global concernant mon environnement de travail à l'ESBS, qui était l'ancienne UMR 7100 et à présent, l'équipe Oncoprotéines. J'ai bénéficié d'un cadre collégial sympathique, chaleureux et dynamique, basé sur l'interaction et le dialogue constructif et non pas sur un esprit de compétition ou de concurrence, ce qui m'a permis de m'épanouir et de me donner à 100% de mes capacités. Mes visites dans d'autres laboratoires de recherche et les discussions que j'ai pu avoir avec de jeunes collègues m'ont conforté dans le fait que j'ai bénéficié d'un environnement de travail exceptionnel.

Je tiens à te dire merci, Dr. Murielle Masson, pour ta disponibilité et les interactions quotidiennes, en tant que voisine de paillasse et de bureau, qui ont souvent dépassé le cadre professionnel, ainsi que pour les conseils et corrections (last minute) de ce manuscrit. « Elle est adorable, notre Murielle… », citation souvent entendue au laboratoire, que je ne peux que confirmer.

Merci au Dr. Denis Dujardin et à Emmanuelle Briand pour les discussions variées et leurs explications concernant l'utilisation du Versadoc. Merci pour votre aide et sympathie.

Je remercie également mes jeunes collègues pour avoir partagé le quotidien avec moi et pour tous les moments agréables. Je dis merci à Charlotte (Cha-Cha) Boulade-Ladame pour son soutien moral, sa disponibilité et les nombreuses « electric barbecue parties » improvisées chez elle, à Jérôme (Jérrrrôme) Courtête, notre breton, pour les nombreuses sorties et les échanges de balles entre midi et deux, au Dr. Laurent (Nounou) Mailly pour ces conseils techniques conçernant la guitare électrique, à Xavier (Xavounnet) Bernard pour les discussions philosophiques non-conventionelles de tous les jours, à Anne-Sophie (AnSo) Rinaldi pour sa bonne humeur et sa joie de vivre contagieuse, au Dr Magali (Mag) Lagrange pour les discussions variées au quotidien et les nombreuses journées BioTechno organisées souvent suivies de restos sympas, le Dr. Tutik Ristriani-Fournel pour son énergie positive et sa place dans le bureau 304 dont j'ai hérité ainsi qu'à Audrey Stoessel. Merci à vous tous pour le temps partagé ensemble, ainsi que pour les moments de détente autour d'une tasse de café et d'une partie de tarot…

Je remercie également Gunter Stier de l'EMBL, Heidelberg, pour m'avoir « provided with the famous pETM vector series », pour m'avoir initié au « extreme high throughput cloning » et pour m'avoir appris « many tricks you can use in a wet lab ». Thank you Gunter.

Merci à mes amis pour leur amitié et leur soutien partout dans le monde: Felix, Phillip, Josip, Christoph, Frank, René, Steven + Carmen, Micha, Nicola, Rosi, Jule + Vincent, Nico, Ségo, Lena, Delph + Seb, Myriam, Gitti, Boris, Vero + Micha… Danke ! Merci ! Thank you ! Grazie ! Gracias !

Anissa, tout simplement : Merci d'être là. Merci de m'avoir soutenu et de m'avoir fait relativiser dans des instants un peu plus difficiles… Kanhebek…

Enfin je voudrais finir en exprimant mon affection à mes parents ainsi que ma sœur. Sans vous je ne serais pas ici et la personne que je suis. Merci du fond du cœur de m'avoir toujours soutenu dans tout ce que j'ai fait et pour l'amour inconditionnel que vous me portez.

Abbreviations

^{15}N	Nitrogen isotope, it has a spin quantum number of ½ and is "visible in NMR"
aa	amino acid
AMPA	α-Amino-5-hydroxy-3-Methyl-4-isoxazole Propionic Acid
APC	Adenomatous Polyposis Coli
ASCUS	Atypical Squamous Cells of Undetermined Significance
ATP	Adenosine Triphosphate
BMRB	Biological Magnetic Resonance Bank (database for NMR spectroscopy information)
CAL	CFTR Associated Ligand
CaMK	Calmodulin Kinase domain
CAP70	CFTR Associated Protein of 70 kDa
CDK	Cyclin Dependent Kinase
CDKI	Cyclin-Dependent Kinase Inhibitor
CFTR	Cystic Fibrosis Transmembrane conductance Regulator
CIN	Cervical Intraepithelial Neoplasia
CIPP	Channel Interacting PDZ domain Protein
CR	Conserved Region
CRB3	Crumbs homologue 3
CRIPT	Cysteine Rich Interactor of PDZ Three
DBD	DNA Binding Domain
DBL	Duffy Binding Like (domain)
DH	Dbl Homology (domain)
DHR	Discs large Homology Region
DLG	Discs Large protein (*Drosophila melanogaster*)
DNA	DesoxyriboNucleic Acid
E6AP	E6 Associated Protein
E6BP	E6 Binding Protein
EDC	1-ethyl-3-(3-dimethylamino-propyl)carbodiimide hydrochloride
EDTA	Ethylendiamintetraacetic acid
EGF	Epithelial Growth Factor
ELISA	Enzyme Linked Immunosorbent Assay
EMBL	European Molecular Biology Laboratory
EMT	Epithelial-to-Mesenchymal Transition
ESI MS	Electrospray Ionisation Mass Spectrometry
EV	Epidermodysplasia Verruciformis
EXAFS	Extended X-ray Absorption Fine Structure
FAP-1	Fas-Associated Phosphatase 1
GKAP	Guanylate Kinase Associated Protein

GRIP	Glutamate Receptor Interacting Protein
GST	Glutathione-S-Transferase
GuK	Guanylate Kinase homology domain
HC 2	Hybrid Capture 2
hDLG	Human homologue of Discs large protein
HDAC	Histone Deacetylase
HECT	Homology to E6AP C-Terminus
HPV	Human Papillomavirus
HSQC	Heteronuclear Single Quantum Coherence
HTLV-1	Human T-cell leukemia Virus type 1
IARC	International Agency for Research on Cancer
ICB	Intracytoplasmatic inclusion bodies
IL5Rα	Interleukin 5 Receptor α
INAD	Inactivation No Afterpotential D
IPTG	Isopropyl-1-β-D-thiogalactopyranoside
IRF-3	Interferon Regulatory Factor 3
JAM	Junctional Adhesion Molecule
k_{on}	Kinetic association rate $[M^{-1}s^{-1}]$
k_{off}	Kinetic dissociation rate $[s^{-1}]$
K_D	Equilibrium dissociation constant $[M^{-1}]$
kDa	kilo-Dalton
Kir2.3	Inward rectifier K^+ channel 2.3
LAP	Leucine rich repeat And PDZ
LCR	Long Control Region
Lgl	Lethal giant larvae
LIM	Protein-protein interaction domain named from the Lin-11, Isl-1 and Mec-3 genes
LRR	Leucine Rich Repeat
M	"molar" = 1 mol/L
MAGI	Membrane Associated Guanylate kinase with Inverted domains
MAGUK	Membrane Associated Guanylate Kinase
MBP	Maltose Binding Protein
MCM7	MultiCopy Maintenance protein 7 (MiniChromosome Maintenance 7)
MDCK	Madin-Darby Canine Kidney
MDM2	Murine Double Minute 2
mRNA	messenger RNA
MUPP1	Multi PDZ domain Protein
NHERF	Na^+/H^+ Exchange Regulatory Factor
NHS	N-hydroxysuccinimide
NIH	National Institute of Health

NiNTA	Nickel chelating agarose resin (QIAGEN)
NLS	Nuclear Localisation Signal
NMDA	N-methyl-D-aspartate
NMR	Nuclear Magnetic Resonance
nNOS	neuronal Nitric Oxide Synthase
NusA	N-Utilisation Substance protein A
ORF	Open Reading Frame
ORI	Origin of replication
PAGE	Polyacrylamide Gel Electrophoresis
PATJ	PALS-1 Associated Tight Junction protein
PCR	Polymerase Chain Reaction
PDGF	Platelet Derived Growth Factor
PH	Pleckstrin Homology (domain)
PKA	Protein Kinase A
PKC	Protein Kinase C
PKN	Protein Kinase N
PLC-β	Phospholipase C - β
PSD 95	PostSynapic Density protein 95
PTB	PhosphoTyrosine Binding (domain)
PTEN	Phosphatase and TENsin homologue
PTP1E	Protein-Tyrosine Phosphatase 1E
Rb	Retinoblastoma protein
RLU	Relative Light Unit
rTEV	recombinant Tobacco Etch Virus protease
SAP97	Synapse Associated Protein 97
SDS	Sodium dodecyl sulphate
SIB	Soluble Inclusion Body
siRNA	small interfering RNA
shRNA	small hairpin RNA
SH2	Src Homology 2 (domain)
SH3	Src Homology 3 (domain)
SPR	Surface Plasmon Resonance
STD	Sexually Transmitted Disease
TAF	TBP-Associated Factor
TBP	TATA-box Binding Protein
TEV	Tobacco Etch Virus
TGF	Transforming Growth Factor
Tip	Tax interacting protein
TNF	Tumour Necrosis Factor

TRX   Thioredoxin
TyK2   Tyrosine Kinase 2
URR   Upstream Regulatory Region
WHO   World Health Organisation
ZO-1   Zonula occludens protein 1

Amino acids

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 OF CONTENTS

RÉSUMÉ EN FRANÇAIS

Introduction

Le cancer du col de l'utérus représente le septième cancer le plus fréquent au monde (Parkin et al., 2005). Le niveau d'incidence le plus élevé est observé dans les pays en voie de développement où le dépistage systématique fait défaut.

Le cancer du col de l'utérus est la deuxième cause de mortalité chez la femme (Parkin et al., 2005). Le facteur étiologique majeur pour le développement d'un cancer du col de l'utérus est l'infection avec un virus du papillome humain (HPV) de souche « haut risque » (les principaux étant : HPVs 16, 18, 31, 33 et 51, par ordre décroissant d'incidence) (Walboomers et al., 1999). Par ailleurs, il semblerait que des cofacteurs additionnels tels que des prédispositions génétiques, l'immunosuppression, des traitements hormonaux (œstrogène et progestérone), le tabagisme ou la multiparité soient impliqués dans le développement du cancer utérin.

Les techniques de dépistage les plus courantes comportent le frottis de Papanicolaou (test Pap) ou la colposcopie. Ces deux méthodes cytologiques permettent une détection précoce de lésions cellulaires anormales, réduisant ainsi considérablement la mortalité dans les pays développés. Récemment, des tests basés sur des méthodes d'hybridation d'ARN ou basés sur des méthodes PCR ont été développés et permettront d'augmenter le niveau de détection de lésions précoces (Denny & Wright, 2005). Récemment, un vaccin prophylactique à été mis sur le marché par Merck & Co (Munoz, 2006). Sur le plan thérapeutique, la méthode à choisir dépendra du stade auquel se situe le cancer. En général, les patients sont soumis à des techniques de chirurgie (micro-) invasive ou à de la chimiothérapie ainsi qu'à de la radiothérapie.

Les HPV sont des virus non enveloppés à ADN circulaire double brin appartenant à la famille des papovavirus. Plus de 100 types d'HPV ont été identifiés, parmi lesquels on trouve une vingtaine de types dits à « haut risque ». Ceux-ci peuvent induire le développement de lésions intraépithéliales qui ont le potentiel de se développer en tumeur (zur Hausen, 2002). La réplication de ces HPV a lieu

exclusivement dans les kératinocytes épithéliaux de la muqueuse vaginale. La réplication des HPV n'est possible qu'au sein de cellules en voie de différentiation.

Le génome viral code pour 6 protéines dites « précoces » (E1, E2, E4, E5, E6, E7) ainsi que pour les deux protéines tardives d'encapsidation (L1, L2). L'expression de ces protéines dépend de deux promoteurs qui sont activés différemment selon l'avancement du cycle viral et de l'état de différentiation de la cellule hôte.

Deux oncoprotéines du HPV, E6 et E7, induisent en synergie des microproliférations cellulaires permettant la réplication du génome viral et la formation de virions, aboutissant parfois à l'immortalisation et à la transformation de ces cellules.

E7 interagit notamment avec des protéines de la famille Rb et induit leur dégradation, supprimant un point de contrôle négatif du cycle cellulaire.

E6 est une protéine contenant deux domaines à zinc homologues, nommés E6-N et E6-C. La protéine E6 des HPV du type haut risque induit la dégradation du suppresseur de tumeur P53 en formant un complexe trimérique avec l'ubiquitine ligase E6AP. Cet évènement est nécessaire pour le déclenchement d'un cancer. Mais l'oncogénicité de E6 ne se réduit pas à son action sur P53. En effet, E6 active la transcription de promoteurs viraux ou cellulaires (Dey et al., 1997 ; Morosov et al., 1994 ; Ronco et al., 1998 ; Sedman et al., 1991) tel que le promoteur de la protéine hTERT du complexe de la télomérase (Klingelhutz et al., 1996 ; Veldman et al., 2001). En outre, E6 interagit avec des acides nucléiques structurés (Ristriani et al., 2000 ; Ristriani et al., 2001). Enfin, E6 interagit avec plus de 30 protéines cellulaires autres que E6AP et p53. Parmi ces protéines ont été caractérisées : E6BP (Chen et al., 1995), paxillin (Tong & Howley, 1997 ; Tong et al., 1997 ; Vande Pol et al., 1998), IRF3 (Ronco et al., 1998), clathrin-adaptor complex AP-1 (Tong et al., 1998), hMCM7 (Kukimoto et al., 1998), CBP/p300 (Patel et al., 1999 ; Zimmermann et al., 1999), Bak (Thomas & Banks, 1998), E6TP1 (Gao *et al.*, 1999), Tyk2 (Li *et al.*, 1999), MUPP1 (Lee et al. 2000), hDLG (Kiyono et al. 1997), hSCRIB (Nakagawa et al. 2000), MAGI1 (Glaunsinger et al 2000) etc. Parmi ces protéines se trouve une famille de protéines multidomaines possédant une activité "Membrane-Associated Guanylate Kinase" (MAGUK) (Craven & Bredt, 1998). Ces protéines se localisent au niveau de jonctions cellulaires, principalement dans les cellules épithéliales ou neuronales

(Dimitratos et al. 1999). Les MAGUKs recrutent diverses protéines cellulaires pour former des complexes de signalisation intercellulaires (Fanning et al., 1999 ; Kotelevets et al., 2005 ; Wu et al. 2000a; Wu et al., 2000b). Des perturbations dans ces complexes multiprotéiques peuvent induire des néoplasies ou la perte de polarisation apico-basale de cellules (Latorre 2005 ; Nguyen 2003). Les MAGUKs possèdent plusieurs domaines d'interaction protéine-protéine parmi lesquels plusieurs copies de domaines dits "PDZ" (Gonzalez-Mariscal et al., 2000). Les études structurales de domaines PDZ (Harrisson, 1996 ; Nourry et al., 2003 ; Songyang et al., 1997) ont montré que ces domaines présentent un mode très conservé d'interaction avec leurs peptides cibles. Le peptide s'insère entre un feuillet β et une hélice α du domaine PDZ, en formant un nouveau brin β qui vient agrandir un feuillet β préexistant. La présence de la fonction carboxyle COO- libre à l'extrémité du peptide est essentielle à la reconnaissance. Des études comparatives ont permis d'identifier des motifs consensus, à partir desquels un classement des multiples peptides C-terminaux a été mis en place. Ces peptides sont divisés en quatre familles en fonction de leur homologie de séquence. Ainsi, les peptides C-terminaux dits de classe I présentent une séquence consensus X[S/T]X[V/L]-*COOH*, les peptides de classe II un motif XφXφ-*COOH* et enfin, les peptides de classe III un motif [D/E]XW[C/S]-*COOH*, où C, D, E, S, T, W correspondent aux résidus en code d'acides aminés en une lettre, X à un résidu quelconque, φ à un résidu hydrophobe et *COOH* au groupement carboxylate libre. Récemment, une nouvelle classe a été mise en place pour les peptides qui présentent le motif XXΨ[D/E]-COOH dit de classe IV (Vaccaro & Dente, 2002).

Toutefois, des données récentes démontrent que le mode d'interaction n'est pas exclusivement gouverné par des séquences consensus C-terminales. En effet, plusieurs exemples ont été décrits où un domaine PDZ est capable de lier des peptides C-terminaux de classes différentes. Ceci a permis l'évolution du modèle des classes de peptides C-terminaux vers un modèle de poches d'interaction qui sont situées au niveau du sillon de fixation du peptide au sein du domaine PDZ. Ce mode d'interaction est plus flexible et permet une meilleure description des interactions peptide-PDZ qui ont été étudiées jusqu'à présent (Kang et al., 2003). En outre, il a été démontré que, parallèlement aux acides aminés de la séquence consensus C-terminale, d'autres résidus en amont peuvent être impliqués dans l'interaction et

ainsi participer à la médiation de l'affinité et à la spécificité d'interaction (Cai et al., 2002; Dobrosotskaya et al., 2001; Niethammer et al., 1998).

Les interactions E6/MAGUK se mettent en place *via* des domaines PDZ, faisant intervenir un motif consensus de 4 résidus X[T/S]X[V/L] situé à l'extrême C-terminus de E6 (Thomas et al., 2001).

MAGI-1 est une MAGUK comprenant 5 domaines PDZ. Elle interagit avec E6 par son domaine PDZ1 et est dégradée par la voie d'ubiquitination du protéasome (Thomas et al., 2001). Des expériences sur des souris transgéniques exprimant un mutant de HPV16-E6 incapable de lier des domaines PDZ ont montré que celles-ci ne développent plus d'hyperplasies épithéliales ainsi que de carcinomes squameux. La dégradation de protéines à domaine PDZ représente une des activités essentielles de E6 pour l'induction de cancers (Nguyen et al., 2003 ; Simonson et al., 2005). Il s'agirait donc d'une cible thérapeutique potentielle. Cependant, plus de 400 domaines PDZ dans plus de 250 protéines ont été prédits dans le génome humain et la simple classification en fonction de motifs consensus ne constitue pas un élément suffisant pour expliquer la spécificité des différents couples PDZ/protéine (Sheng and Sala, 2001). Ainsi, des résidus additionnels en amont du consensus ont été décrits comme essentiels pour une spécificité de liaison (Cai et al., 2002 ; Kang et al., 2003; Skelton et al., 2003 ; Wiedemann et al., 2004). Si plusieurs études très poussées ont permis d'établir les bases moléculaires de la spécificité d'interaction de plusieurs couples PDZ/peptide (Doyle et al., 1996; Kang et al., 2003; Skelton et al., 2003), peu d'études ont été effectuées concernant les interactions E6/PDZ.

Projet de thèse

Dans ce contexte scientifique, mon projet de thèse s'est articulé autour de la relation « structure fonction » de l'oncoprotéine E6 de HPV 16 (16E6). Plus précisément, j'ai effectué une étude structurale de la spécificité d'interaction entre 16E6 et le domaine PDZ1 de MAGI-1 (MAGI-1 PDZ1), ainsi qu'une analyse cinétique de la liaison entre MAGI-1 PDZ1 et des peptides de E6, issus de différents types d'HPV. A ces fins, nous devions produire des échantillons de 16E6 et de MAGI-1 PDZ1, éventuellement marqués isotopiquement pour des études RMN. Longtemps, la production de 16E6 et de ses domaines a été un facteur limitant pour toute étude *in vitro*. L'obtention d'un mutant de 16E6, nommé E6 (6C/6S), avait permis d'obtenir

des échantillons non marqués de 16E6 et de ses domaines. J'ai eu l'occasion de participer à la phase finale de la mise au point des conditions de production de E6 et de ses domaines marqués isotopiquement. J'ai également participé à la construction d'un nouveau système de vecteurs d'expression, nommés vecteurs pETM, à l'European Molecular Biology Laboratories (EMBL, Heidelberg, Allemagne), système que nous avons implémenté au laboratoire. Il nous a permis d'augmenter les rendements de production et, en parallèle, de faire un criblage des conditions d'expression. J'ai appliqué un certain nombre de méthodes d'évaluation de qualité de protéines, basées sur la mesure des quantités de protéine exprimées, de leur solubilité ainsi que de leur monodispersité (Nominé et al., 2001a, Nominé et al., 2001b). Ainsi, j'ai réussi à mettre au point un protocole standard de marquage et de purification du domaine C-terminal de E6 (6C/6S) (E6-C (4C/4S)) à des concentrations suffisantes permettant le calcul de sa structure atomique par RMN (Nominé et al 2005; Nominé et al., 2006). Ceci a ensuite permis l'analyse biophysique du domaine ainsi que le calcul d'un modèle putatif en hététodimère de la protéine entière. Concernant MAGI-1 PDZ1, j'ai procédé de manière similaire. J'ai conçu différentes constructions de manière semi-rationnelle, défini les conditions d'expression favorables et enfin, j'ai réalisé un criblage de ces constructions en fonction de leur solubilité, leur monodispersité et de leur activité biologique. A partir des constructions de MAGI-1 PDZ qui répondaient aux critères requis, j'ai produit un échantillon marqué au ^{15}N. Une analyse rapide de chacune de ces constructions par ^{1}H-^{15}N HSQC m'a permis d'évaluer rapidement leurs repliements et leur compatibilitépour procéder à une analyse structurale par RMN. J'ai retenu une construction MAGI-1 PDZ1, nommée PDZ1-5, qui satisfaisait tous ces critères.

La production et le marquage du PDZ1-5, de E6-C (4C/4S) et de deux peptides C-terminaux de E6 ont permis d'acquérir tous les spectres RMN nécessaires à l'attribution du domaine PDZ1-5 seul (Charbonnier et al., 2006) ainsi que du domaine PDZ1-5 complexé à E6-C ou un peptide C-terminal de E6-C.

L'attribution de spectres RMN est une étape indispensable pour toute étude structurale. La fréquence de chaque atome de la molécule visible par RMN est identifiée fournissant ainsi une carte de la molécule. Cette carte peut être utilisée dans un premier temps pour identifier des résidus impliqués dans le complexe E6-C/PDZ1. Ainsi nous avons pu délimiter une région de 11 résidus C-terminaux de E6-C

16

qui sont nécessaires et suffisants pour la spécificité de liaison au PDZ1-5. Nous avons acquis toutes les données nécessaires pour calculer la structure de PDZ1-5, lié au peptide identifié auparavant. Ces travaux sont actuellement en cours et aboutiront prochainement. D'ici la fin de l'année nous serons dans la mesure de posséder la structure tridimensionnelle de MAGI-1 PDZ1 lié à un peptide C-terminal issu de HPV 16 E6. Cette structure permettra de visualiser et de décrire avec précision l'interaction E6-C/PDZ1 d'un point de vue atomique.

L'étape suivante consistait à faire une étude cinétique de l'interaction E6-C/PDZ1. Pour cela, deux méthodes ont été mises au point au laboratoire.

(i) J'ai participé à la mise au point d'une méthode pour étudier les interactions protéine/protéine, utilisant la Résonance Plasmonique de Surface (SPR), plus communément connue sous le nom de Biacore. Le principe de cette méthode consiste à fixer un anticorps dirigé contre la Glutathione-S-Transférase (GST) sur une puce « sensor », qui est mise en contact avec des microcanaux dans lesquels circule un flux de tampon. Un faisceau lumineux monochromatique illumine la surface de la puce qui est recouverte d'une fine couche d'or. L'angle entrant de ce faisceau est déterminé afin que la réflexion de lumière soit totale. Dans ces conditions (réflexion totale sur une couche de métal riche en électrons), le phénomène de SPR est induit. Les photons incidents entrent en résonance avec les électrons du métal, produisant un cône d'ombre dans le faisceau de lumière réfléchi. L'angle de réflexion de ce cône d'ombre varie en fonction de la masse fixée à la surface de la puce. La variation de cet angle permet donc de quantifier en temps réel des changements de masse dus à des interactions moléculaires à la surface de la puce. Une expérience revient donc à injecter une première protéine (le ligand) qui est fusionnée à la GST, permettant un accrochage sur la puce. Ensuite, on injecte des volumes définis d'une deuxième protéine (l'analyte) à différentes concentrations. La SPR permet de suivre les changements de densité à la surface de la puce en temps réel, donnant accès au constantes cinétiques d'association (k_{on}) et de dissociation (k_{off}), ainsi qu'aux constantes d'association et de dissociation à l'équilibre (K_A et K_D, respectivement) (Charbonnier et al. 2006 (en cours de rédaction) ; Zanier et al., 2005). Cette méthode permet ainsi de cribler des ligands, tels que des mutants, des inhibiteurs ou d'autres molécules potentiellement thérapeutiques et d'évaluer leur affinité. Une limitation majeure de cette méthode est liée à la sensibilité de la technologie Biacore,

17

qui ne permet pas de mesurer avec précision des affinités à l'ordre d'un K_D d'une dizaine de µM et au delà.

(ii) Pour ces conditions spécifiques de faible affinité ou pour des k_{off} rapides, j'ai mis au point une deuxième méthode, nommée « holdup assay » qui permet d'étudier à l'équilibre des interactions d'une large gamme d'affinités, notamment de faible affinité (Charbonnier et al., 2006). Cette méthode est basée sur la rétention sélective chromatographique. En résumé, le ligand fusionné à une protéine porteuse est fixé sur des billes de résine d'affinité dans deux lots identiques. Un analyte purifié ou présent dans un extrait cellulaire brut y est ajouté et incubé jusqu'à ce que l'équilibre de l'interaction moléculaire soit atteint. Ensuite, dans un des lots, on injecte un compétiteur qui permet d'éluer le ligand de la résine (+), alors que l'autre lot reste inchangé (-). Finalement on extrait la totalité de la phase liquide des deux lots (+) et (-). Les extraits sont analysés soit par SDS PAGE classique dans le cas de protéines surexprimées en système bactérien, soit par western blot dans le cas de protéines issues d'extraits cellulaires. Si une interaction a lieu, l'analyte reste accroché en partie ou de manière quasi totale au ligand, fixé sur les billes de résine dans le lot (-). Le niveau d'analyte sera donc inférieur ou il aura complètement disparu dans le lot (-) par rapport au lot (+). Si une interaction n'a pas eu lieu, les niveaux d'analyte seront pareils dans les lots (-) et (+). Non seulement nous avons démontré que cette méthode fonctionne avec des protéines recombinantes produites en bactéries mais qu'elle est également adaptable pour l'analyse des protéines exprimées de manière transitoire dans des cellules eucaryotes. Elle permet de détecter des interactions de faible affinité, imperceptibles par une approche classique dite de « pulldown », qui élimine les protéines de faible affinité ou participant à des complexes transitoires dans les étapes de lavage et favorise la détection de complexes de forte affinité. Le « holdup assay » est adapté pour réaliser de façon rapide des criblages qualitatifs. Un avantage majeur est le fait de pouvoir travailler à l'équilibre, donnant accès à toutes les informations nécessaires pour l'estimation de K_D avec une précision acceptable. Dans l'exemple que nous présentons dans la partie résultat de ce mémoire, les mesures de K_D pour un couple ligand/analyte donné sont confirmées par des mesures Biacore (Charbonnier et al., 2006).

Nous avons appliqué ces deux méthodes et étudié la cinétique d'interaction de différents peptides C-terminaux de E6 issus de différents HPV, ainsi que de protéines

ayant été décrites dans la littérature comme liant MAGI-1 PDZ1. Nous avons mesuré leurs K_D par Biacore. Ainsi, l'analyse d'alignements de séquence des différents peptides C-terminaux en fonction de leur KD respectifs a permis d'identifier des résidus clés situés dans des positions conservées en amont de la séquence consensus, qui semblent être impliqués dans la spécificité d'interaction.

<u>Conclusions et perspectives</u>

L'interaction de la protéine E6 des HPV à haut risque est une des activités nécessaires pour l'induction d'hyperplasies *in vivo* ainsi que pour la progression maligne vers un cancer du col de l'utérus. Il a été démontré que la protéine E6 d'HPV à haut risque est plus impliquée dans l'étape de progression du cancer que dans l'étape de promotion (Mantovani & Banks, 2001 ; Simonson et al., 2005). Cette observation semble cohérente avec l'hypothèse que la dégradation de protéines à domaines PDZ est importante lors des phases tardives du développement d'un cancer. La dégradation des protéines PDZ contribuerait à la perte de la polarité cellulaire et de l'inhibition de contact entre cellules voisines et induirait ainsi la progression vers des phénotypes invasifs (Mantovani & Banks, 2001). Une étude approfondie de l'interaction de protéines E6 d'HPV à haut risque avec des domaines PDZ présente donc un intérêt majeur et permettrait de fournir des données utiles dans la quête de nouvelles cibles thérapeutiques.

Jusqu'à présent, aucune étude structurale et cinétique n'avait pu être faite, en raison du manque de protéine E6 recombinante purifiée. Nous avons réussi à produire et à marquer isotopiquement le domaine 16E6-C ainsi que le domaine PDZ1-5 à des concentrations suffisantes pour que des études structurales par RMN soient réalisées. Ainsi, l'étude RMN a permis de déterminer la structure de E6-C et d'identifier les 11 derniers résidus comme étant une région d'interaction avec le PDZ1-5. Nous avons attribué le domaine PDZ1-5 à l'état libre et lié au peptide C-terminal de E6-C et les calculs structuraux aboutiront à la fin de l'année 2006. La cinétique d'interaction pour différents peptides C-terminaux a été analysée par Biacore et a permis d'identifier des résidus clés pour la spécificité d'interaction.

Ces travaux présentent pour la première fois des données structurales concernant une activité oncogénique de HPV16 E6 et donnent accès au K_D de l'interaction 16E6-C/MAGI-1 PDZ1. La structure RMN du complexe permettra de

comprendre plus précisément le mode d'interaction et les déterminants de spécificité. Les méthodes Biacore et holdup, mises au point au sein de notre équipe, présentent des outils performants pour l'analyse cinétique ainsi que le criblage de mutants, de peptides ou autre molécule à potentiel thérapeutique.

Les travaux décrits ici ouvrent par ailleurs le champ sur une étude plus systématique. La maîtrise des connaissances concernant la manipulation de 16E6-C ainsi que du MAGI-1 PDZ1 permettront de produire aisément et d'étudier tout autre domaine PDZ liant E6 avec les méthodes utilisées durant ces travaux. Une telle étude systématique permettrait de décrire les domaines PDZ ciblés par les protéines E6 de virus haut risque et de décrire le mécanisme de d'interaction de manière précise afin d'identifier les déterminants de spécificité et de trouver des ligands à forte affinité. Ces données seront complémentaires à des études *in vivo* et permettront éventuellement, à long terme, d'identifier de nouvelles cibles thérapeutiques dans la lutte contre le cancer du col de l'utérus.

INTRODUCTION
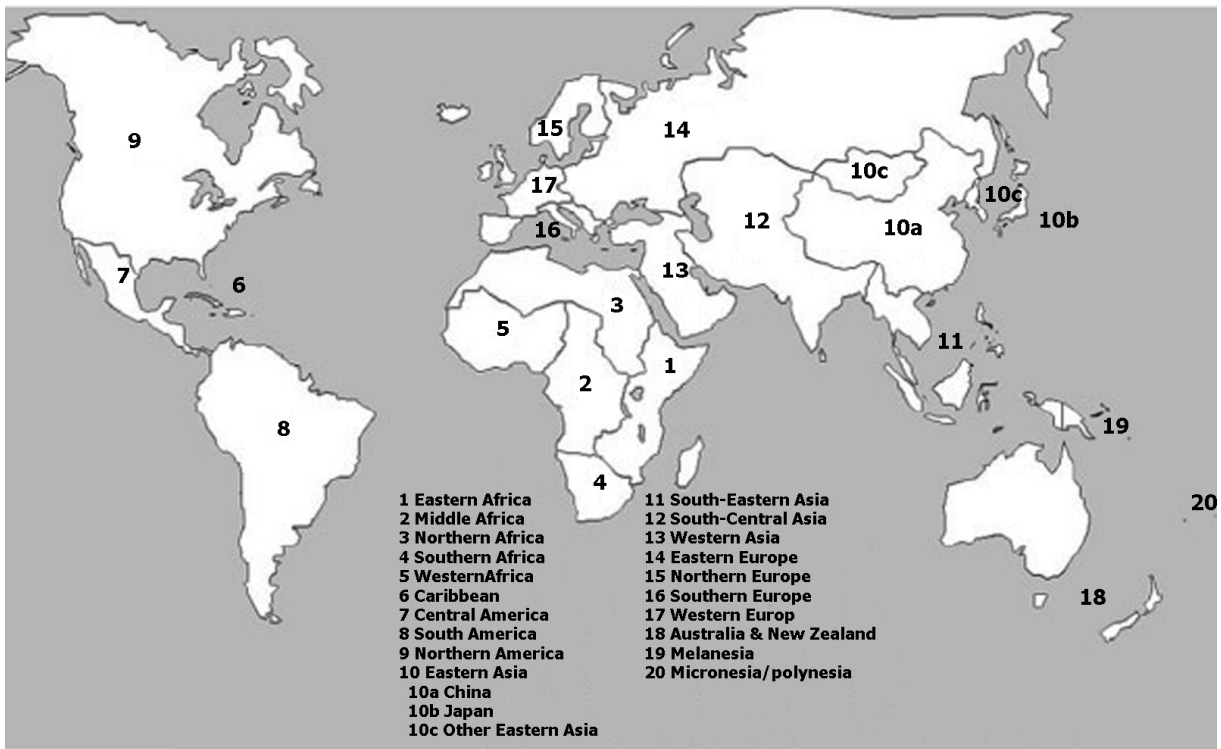
Figure 1.: The world areas as defined by the United Nations Population Division. This map will be used to illustrate the incidence and mortality rates of HPV induced cervical cancer worldwide (Parkin et al., 2006). A more detailed map according to individual countries is available in the following reference (Schiffman et al., 2005).
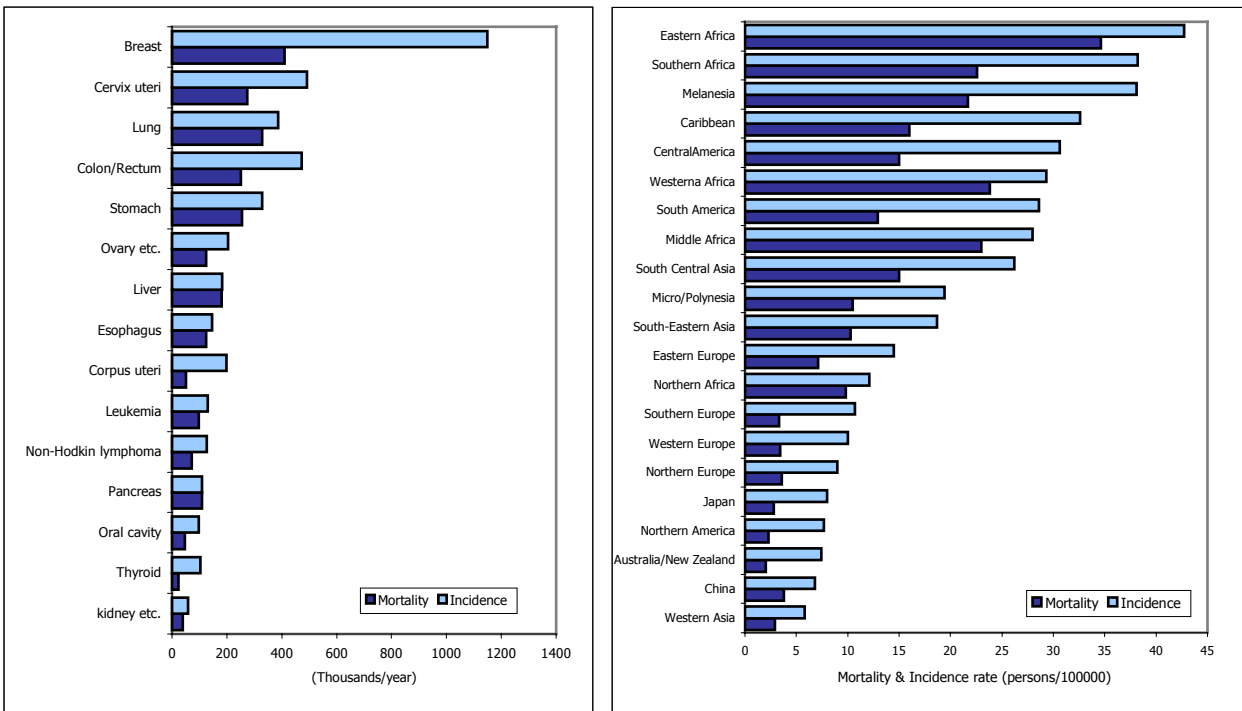
Figure 2.: A/ Estimated numbers of new cancer cases (incidence) and deaths (mortality) in 2002. Data shown in thousands for developing and developed countries by cancer site and female population.
B/ Age-standardised incidence and mortality rates for cervical cancer worldwide. Data shown as individuum per 100,000. Figures derived from Parkin et al., 2005 and recent IARC data.

Cervical cancer

Cervical cancer - worldwide

For the last 30 years the IARC[1] has been estimating the global cancer burden, providing global but still detailed estimations of incidence, mortality and prevalence for about 26 types of cancer in each national population worldwide for the year 2002[2].

Cervical cancer is the seventh cancer in frequency overall, but the second most common cancer in women worldwide (Parkin et al., 2005; Figure 2A). 493000 new cases and 274000 deaths were estimated for the year 2002 (Figure 2B).

Cervical cancer is more common in developing countries, where about 83% of the cases occur. In these countries cervical cancer counts for 15% of female cancers with a risk potential of 1.5% to be concerned before the age of 65. In developed countries this type of cancer counts for only 3.6% with a risk potential of 0.8% before the age of 65.

The highest incidence rates observed in developing countries are due to the lack of systematic screening approaches. The low incidence rates in developed countries are the consequence of the introduction of large screening programs in the 1960s, with age standardised rates less than 14.5 per 100000 (Figure 2B). Worldwide the ratio of mortality vs. incidence is about 55%, with good survival rates in developed countries (63% in Europe). However this is just a mean value and it is striking to see how close the mortality rates are in comparison to the incidence rates in developing countries (Figure 2B). This is true especially on the African continent, compared to the developed countries. The latter present globally lower incidence and

---

[1]IARC: International Agency for Research on Cancer

[2]   *Incidence* - is the number of new cases which occur in a year. It is expressed as the absolute number of cases per year or as a rate per 100000 persons per year. Primary prevention strategies generally aim at reducing the incidence.

      Prevalence - describes the number of persons alive at a particular time point with a particular disease.

      Mortality - is the number of deaths which occur in one year. The mortality rate is the number of deaths per 100000 persons per year. Mortality is calculated as the product of the incidence and the fatality of a given cancer. The mortality rate calculates therefore the average risk of dying from a specific cancer during a specific period (generally one year).

      Estimation methods have been described in the 1990s (Parkin et al., 1999; Pisani et al., 1999, Pisani et al., 2002). The results are presented in general as global totals by world area. A map of the world areas is displayed in figure 1. Countries generally labelled as "developed" comprise Northern America, Japan, Europe, Australia and New Zealand. The remaining countries are generally labelled as "developing".

mortality rates but also significantly lower incidence vs. mortality ratios. One explanation for this high incidence vs. mortality ratio in developing countries could be insufficient screening. The precancerous lesions in early stages are not detected and the chances to be treated with low risks of relapse are thus diminished. Another factor implicated is certainly the lack of suitable medical treatment once a cancer has been diagnosed. Therefore efforts must be done to provide the population of developing countries with information campaigns, systematic screening programs to prevent and reduce risk of infection and to make affordable treatment available once a precancerous lesion or cancer has been detected. Nevertheless, substantial declines in cervical cancer incidence and mortality have been observed mainly in western countries providing elaborate screening programs but also in some developing countries, especially in China, where the estimated age-standardised incidence rate dropped from 17.8% in 1985 to 6.8% in 2002. As a result of this trend, cervical cancer has ceded its place as the leading cancer in developing countries to breast cancer. However, when taking aside global statistics and looking more closely at individual countries cervical cancer remains still the main cancer affecting women in sub-Saharanian Africa, Central America, southcentral Asia and Melanesia (Parkin et al., 2005).

Cervical cancer – the etiological factors

The major etiological agents for the development of preinvasive and invasive cervical cancer are the oncogenic subtypes of HPV[3], the so-called "high risk" HPV. Their DNA can be detected in 99.7% of all cervical cancers (Walboomers et al., 1999). Already in the 1970s cervical cancer had been correlated to HPV but only in the 1980s were the first HPV types (HPV 16 and 18) isolated from cervical cancer tissue (Boshart et al., 1984; Durst et al., 1983). Direct correlation between HPV infection and cervical cancer was finally recognised by the WHO[4] and the NIH[5]. HPV 16 is the most common viral type found in cervical tumours (50-60%), followed by

---

[3] HPV: Human Papilloma Virus
[4] WHO: World Health Organization
[5] NIH: National Institutes of Health

HPV 18 (15%), HPV 45 (8%) and HPV 31 (5%). The remaining cervical cancers contain any of the other high risk HPV types[6].

HPV infection seems therefore to be the causative event for cervical cancer. However, it is important to note that HPV infection alone is likely to be insufficient to induce cervical cancer, since high-risk HPV infection is a common event in women with a normal cervical cytology and the majority of them will never develop a cervical cancer. Besides this major etiological factor, many cofactors play an important role in the development and the progression of cervical cancers.

A number of possible co-factors for progression of HPV infection into cervical cancer have been identified. There are genetic cofactors such as a sequence polymorphism in the sequence of the p53 tumour suppressor protein, which encodes either a P[7] or an R residue at position 72 (Matlashewski et al., 1987). P53 is a key protein in the cell cycle control and is inactivated in HPV induced cancers. It has been shown that p53 (72R) was twice as effective in preventing immortalisation of rodent cells and that this could be correlated to the induction of apoptosis (Crook et al., 1994; Thomas et al., 1996). It has also been shown that p53 (72R) was considerably more susceptible to HPV-dependent inactivation than p53 (72P). Genotyping seemed to show a prevalence for the homozygous genotype 72R/72R in cervical tumours (Storey et al., 1998). However it is not yet clear how important this polymorphism is, since there has been at least one conflicting report (Rosenthal et al., 1998).

Another co-factor might be immunosuppression. Immunosuppressed persons, such as organ transplant patients or HIV[8] infected persons, reveal increased levels of HPV infection and cervical cancer (Halpert et al., 1986; Matorras et al., 1991; Porreco et al., 1975; Schrager et al., 1989).

Smoking also seems to increase the risk of cervical cancer. Smoking has been described as decreasing the number of antigen-presenting Langerhans cells in the cervix, which might lessen the local immunity to HPV infection (Sasson et al., 1985). Furthermore by-products of cigarette smoke, which are present in the cervical

---

[6] High-risk anogenital HPV types confirmed officially by the International Agency for Research on Cancer: 16, 18, 31, 33, 35, 39, 45, 51, 52, 56, 58, 59, 66

[7] All amino acids will be written in the one letter code in this manuscript. A table displaying the one letter code can be found in the abbreviations at the beginning of the manuscript.

[8] HIV: Human Immunodeficiency Virus

mucus, could increase the cancer risk because of their mutagenic effects (Barton et al., 1988; Roche & Crum, 1991).

The steroid hormones progesterone and oestrogen have been described to be also important cofactors. Both hormones have been shown to influence HPV 16 and 18 gene expression (Chen et al., 1996; Kim et al., 2000; Mitrani-Rosenbaum et al., 1989; Mittal et al., 1993; Yuan et al., 1999;). Most cases of cervical cancer occur in a well defined region of the cervix, called the transformation zone, which is believed to be oestrogen sensitive (Elson et al., 2000). However the progesterone or oestrogen delivered by oral contraceptives seem not to increase the risk of developing HPV mediated cervical cancer (Brisson et al., 1994).

Finally, as cervical cancer is a sexually transmitted disease, sexual behaviour influences directly the risk of infection. It remains clear that people with greater numbers of sexual partners are at increased risk of developing genital HPV-related diseases. Co-infection with other sexually-transmitted pathogens, such as HIV, has already been described to increase the risk of developing HPV-related diseases (Palefsky, 2006).

Prevention and screening

Genital HPV infections may be distributed widely over genital skin and mucosal surfaces. Transmission can occur even when there are no overt symptoms. Several strategies could be employed to detect precancerous cervical lesions as soon as possible in order to minimise the risk of developing cancer caused by genital HPVs.

Papanicolaou screening, also known as "Pap" smear, was introduced in 1949 and is an effective strategy for reducing the risk of invasive cervical cancer (Papanicolaou, 1949). This technique is still used today (Burd et al., 2003; Jordan & Monaghan, 2004). However, the risk of obtaining false negative results remains the main drawback of this screening method. Only 51% of cervical intraepithelial lesions and invasive cancers can be detected accurately (Nanda et al., 2000).

Colposcopy may be indicated if a detailed inspection of the cervix is recommended due to detection of abnormal cells by routine Pap smear. It was developed by the German physician Hinselmann. This diagnostic procedure uses a colposcope, which is a kind of binocular, to examine a magnified and illuminated area of the cervix. Precancerous or malignant lesions can be detected, due to discernible

characteristics and site directed biopsies can be taken for further pathological examination.

The traditional cervical cytology based screening methods have revealed to be an effective way of preventing cancer. Figure 3 shows a classical approach of regular Pap-smears (blue arrows) combined with colposcopy and biopsies at statistically risky ages or in case of positive Pap-tests. However, these methods require good coverage of the target population (> 80%) and frequent intervals of 3-5 years to be efficient (Hakama et al., 1986; Hakama & Louhivuori, 1988). As mentioned before, cytological analysis shows a marked variability in its detection efficiency, with an estimated accuracy of 50% to identify high grade squamous intraepithelial lesions and cervical cancer (Nanda et al., 2000; Fahey et al., 1995). These limitations and the strong causal association between high risk HPV infection and the development of cervical cancer has led to propose DNA based screening approaches for the secondary prevention of cervical cancer, either alone or as an additional test to cytology.

Two tests were validated in large trials and epidemiological studies: first HC 2[9], which provides a viral profile of a number of different HPV types and second PCR[10]-based methods, using consensus primers to identify specific HPV types.

HC 2: In 1999 HC 1 was initially approved by the FDA as a complementary test to cytology for women with atypical squamous cells of undetermined significance (ASCUS). In 2003 the FDA finally approved the use of 13 HPV types detected by HC 2 as a complementary test to cytology for women aged 30 or older. RNA probes which are complementary to 13 high-risk (16, 18, 31, 33, 35, 39, 45, 51, 52, 56, 58, 68) and five low-risk HPV types (6, 11, 42, 43, 44) are used to detect presence or not of one of these HPV types in a standardised 96 well microplate hybridisation assay. DNA is denaturated and hybridised to the RNA probes. DNA-RNA hybrids are captured by specific antibodies bound to the wells of the microplates. Positive hybridisation events are detected by a series of reactions giving rise to luminescence which can be quantified with a luminometer. The measured RLUs[11] provide a semi-quantitative measure of the viral load. The cut-off for a positive test is 1.0 RLUs

---

[9] HC 2: Hybrid Capture 2
[10] PCR: Polymerase Chain Reaction
[11] RLU: Relative Light Unit

(approx 5900 genomes per test well). However, the result does not provide information about the specific types of HPV detected.

PCR-based assays: Today, no PCR-based tests are commercialised but they are in development and should soon be available. The advantages of these tests are that they require only small amounts of biological sample, that there is only a simple one-step treatment before the amplification reaction and their analytical sensitivity. As few as 10-100 copies of HPV genomes can be detected in the tissue being tested. Either a set of degenerated or specific primers against distinct viral DNA regions (usually conserved regions in the L1 gene) are used for a classical PCR reaction, leading to an exponential and reproducible increase in the nucleic acid sequences present in the biological sample. The primers used today are capable of detecting up to 40 different HPV types. Analysis of the amplification products can be performed by gel electrophoresis, dot blot, restriction fragment length polymorphism analysis or sequencing. PCR based DNA tests can therefore be used for primary screenings to identify women infected with high-risk types. These women could be assigned to a high-risk group of individuals which should perform regular cytological check-up for limiting the risk of precancerous cervical lesions. DNA tests could therefore be of use for the triage of ambiguous cytological results and finally for the follow-up of women post-treatment for cervical cancer precursors (Denny & Wright, 2005).

The Norwegian company NorChip commercialises the PreTect HPV-Proofer assay, which is based on real-time multiplex nucleic acid sequence-based amplification (NASBA). Briefly, total mRNA[12] is extracted from a biological sample and specific HPV mRNA are amplified in a DNA background with a real-time detection by measuring increasing fluorescence of HPV specific molecular beacon probes.

Another protein-based new approach for detecting presence of HPV in biological samples is being developed by the American company Arbor Vita Corporation. Briefly their test is based on specific protein domains called PDZ[13] domains which bind to a specific early protein of all high risk HPV. These PDZ domains can be used in an ELISA[14]-based approach to capture the early viral protein. Presence of viral protein is revealed by hybridisation with a first specific antibody and

---

[12] mRNA: messenger Ribonucleic Acid
[13] PDZ: PSD95/Dlg/ZO1
[14] ELISA: Enzyme-Linked Immunosorbent Assay

a second antibody directed against the first which is coupled to a marker. This diagnostic tool is in the late trials and should be commercially available in the next two years.

Prophylactic vaccination

Cervical cancer is exclusively linked to infection with HPV. Therefore a prophylactic vaccination approach is possible for this type of cancer (Walboomers et al. 1999).

Merck & Co. has developed a vaccine against four HPV strains, which is commercialised under the name of Gardasil™. It received the Food and Drug Administration approval on June 8, 2006. The HPV major capsid protein, L1, can spontaneously self-assemble into virus-like particles (VLPs). Gardasil contains recombinant VLPs assembled from the L1 proteins. Since VLPs lack the viral DNA, there is zero risk of inducing cancer. They exclusively trigger an efficient antibody response thus protecting recipients from HPV infection of the types represented in the vaccine (Munoz, 2006). Gardasil™ is targeting girls and women aged 9 to 26 since the vaccine only works if given before infection occurs. The use of the vaccine in men to prevent genital warts and interrupt transmission to women is initially considered only a secondary market. The high cost of this vaccine has been a cause of concern.

Glaxosmithkline has developed a vaccine called Cervarix™ which has been shown to be effective in preventing infection by HPV strains 16 and 18 for more than four years. Cervarix™ should be approved by year's end.

A few conservative religious groups publicly oppose the concept of making HPV vaccination mandatory for pre-adolescent girls, citing fears that vaccination against a sexually-transmitted disease might send a subtle message assuming pre-marital sex.

In addition to preventive vaccines, such as Gardasil and Cervarix, laboratory research and several human clinical trials are focused on the development of therapeutic HPV vaccines. In general these vaccines focus on the main HPV oncogenes, E6 and E7 (see later chapters for more details). Since expression of E6 and E7 is required for promoting the growth of cervical cancer cells (and cells within

warts), it is hoped that immune responses against the two oncogenes might eradicate established tumours.

Therapy

The choice of the appropriate treatment depends on the FIGO[15] stage of the squamous intraepithelial lesion. Different types of treatment can be distinguished according to if they are applied locally or in a systemic way.

Microinvasive cancer of stage IA can be treated by removal of the whole uterus including part of the vagina (hysterectomy). In case of stage IA2 lesions the hysterectomy includes in addition the lymph nodes. An alternative for patients who desire to maintain fertility is a local surgical procedure such as a cone biopsy. A cone biopsy is an extensive form of cervical biopsy where a cone-shaped wedge of tissue is removed from the cervix including a small amount of normal tissue so that a margin free of abnormal cells is left in the cervix.

If a cone biopsy was not able to produce clear margins, trachelectomy is a possible option left for those with early stage cervical cancer who would like to preserve their fertility. In this surgery, the cervix and the upper part of the vagina are removed but the rest of the uterus is left in place. The lymph nodes in the pelvis are also removed, usually by keyhole laparoscopic surgery, to see if the cancer has spread.

For those in stage I cervical cancer, which has not spread, this is a viable treatment option. It allows the preservation of the ovaries and uterus while surgically removing the cervical cancer. This operation can also be performed vaginally instead of abdominally. However, there are conflicting opinions as to which approach is better. A radical abdominal trachelectomy with lymphadenectomy usually only requires a 2 to 3 day hospital stay with most women recovering very quickly (approximately 6 weeks).

Recurrence in the residual cervix is a very rare event as long as the cancer has been cleared with the trachelectomy. Even though recurrence is rare, it is generally recommended for patients to practice vigilant prevention and follow-up care including pap screenings and colposcopy, with biopsies of the remaining lower uterine segment as needed to monitor for any recurrence.

---

[15] FIGO: Fédération Internationale de Gynécologie et d'Obstetrique (see chapter « HPV and cancer » for more details)

Type	Associated to:	Type	Associated to:
1	deep plantar warts	49	flat warts under immunosuppression or in case of EV
2	cutaneous warts (also tongue cancer)	50	benign EV lesions
3	"flat warts" (benign cutaneous lesions)	**51**	anogenital intraepithelial lesions & cancer
4	common skin warts (mainly on hand)	**52**	anogenital intraepithelial lesions & cancer
5	mainly benign/malign EV lesions	53	present in normal genital mucosa
6	warts in the genital and respiratory tract	54	condyloma acuminata
7	"butchers warts", oral & facial warts in HIV patients	55	condyloma acuminata
8	EV lesions	**56**	anogenital intraepithelial lesions & cancer
9	EV lesions	57	oral & inverted maxillary sinus papillomas
10	venign EV lesions / flat warts	**58**	anogenital intraepithelial lesions & cancer
11	warts in the genital and respiratory tract	**59**	anogenital intraepithelial neoplasia & cancer
12	benign EV lesions	60	keratinous plantar cysts
13	benign orogenital lesions (low-grade CIN)	61	vulvar intraepithelial neoplasia
14	flat wart like EV lesions	62	vulvar intraepithelial neoplasia
15	benign flat wart like EV lesions	63	punctate keratotic lesions/ICB
16	anogenital intraepithelial neoplasia & cancer	64	vulvar intraepithelial neoplasia
17	benign macular EV lesions	65	black verroucus lesions/ICB
18	anogenital intraepithelial neoplasia & cancer	**66**	cervical intraepithelial lesions & cancer
19	benign macular EV lesions	67	anogenital intraepithelial lesions
20	EV lesions / squamous cell carcinoma	68	cervical carcinoma
21	EV lesions	69	dysplastic lesions of the tongue
22	EV lesions	70	cervical condyloma
23	EV lesions	72	oral papilloma (HIV patient)
24	EV lesions	73	oral papilloma (HIV patient)
25	benign macular EV lesions	74	anogenital intraepithelial neoplasia
26	skin warts under immunosuppression	75	cutaneous warts
27	common warts	76	benign cutaneous warts
28	"butchers warts", "flat warts"	77	skin warts/squamous carcinomas of the skin
29	common warts	80	cutaneous warts/EV lesions
30	laryngeal carcinoma/condyloma accuminata		
31	anogenital lesions & cancer		
32	oral focal epithelial hyperplasia & papilloma		
33	anogenital intraepithelial lesions & cancer		
34	orogenital lesions (low oncogenic potential)		
35	anogenital lesions & cancers (medium risk)		
36	actinic keratosis / EV lesions		
37	keratocanthoma (basaloma)		
38	melanoma under immunosuppression		
39	anogenital intraepithelial neoplasia & cancer		
40	anogenital intraepithelial neoplasia & cancer		
41	cutaneous squamous cell carcinoma		
42	benign genital lesions (flat genital warts)		
43	anogenital intraepithelial neoplasia		
44	anogenital intraepithelial neoplasia		
45	anogenital intraepithelial neoplasia & cancer		
47	benign EV lesions		
48	cutaneous squamous cell carcinoma		

Table 1.: Characterisation of the different HPV types and their prefential localisation. A HPV genotype is defined as a viral DNA having less than 90% nucleotide similarity to any other HPV genotype in the L1 open reading frame (ORF). A HPV subtype is defined as a viral DNA having between 90 and 98% similarity and a HPV variant is defined as a viral DNA having more than 98% nucleotide similarity (The Human Papillomavirus Compendium 1995).

Early stages IB1 and IIA can be treated with radical hysterectomy with removal of the lymph nodes or radiation therapy. Radiation therapy is given as external beam radiotherapy to the pelvis and brachytherapy (internal radiation). For patients treated with surgery who have high risk features found on pathologic examination, radiation therapy with or without chemotherapy is given in order to reduce the risk of relapse.

Larger early stage tumours IB2 and IIA may be treated with radiation therapy and cisplatin-based chemotherapy, hysterectomy (which then usually requires adjuvant radiation therapy), or cisplatin chemotherapy followed by hysterectomy. Advanced stage tumours IIB-IVA are treated with radiation therapy and cisplatin-based chemotherapy.

Human papilloma viruses

HPV and cancer

Papillomaviruses were first identified in the early 20th century, when it was shown that skin warts, or papillomas, could be transmitted between individuals by a filterable infectious agent. In 1935 Francis Peyton Rous, who had previously demonstrated the existence of a cancer-causing sarcoma virus in chickens, went on to show that a papillomavirus could cause skin cancer in infected rabbits. This was the first demonstration that a virus could cause cancer in mammals. They represent a large and diverse group of non-enveloped DNA viruses which infect birds and mammalians (Moreno-Lopez et al., 1984). Papillomaviruses replicate exclusively in body surface tissues such as the skin, or the mucosal surfaces of the genitals, anus, mouth or airways (Doorbar, 2005).

More than 100 distinct human papillomavirus (HPV) types have been identified based on differences in their DNA sequence (Chan et al., 1995; De Villiers et al., 2004). A list of different HPV types and their preferential cellular localisation are displayed in table 1 (zur Hausen et al., 1996). Some HPV types cause benign skin warts, or papillomas, for which the virus family is named. HPVs are associated with the development of common warts and are transmitted environmentally or by casual
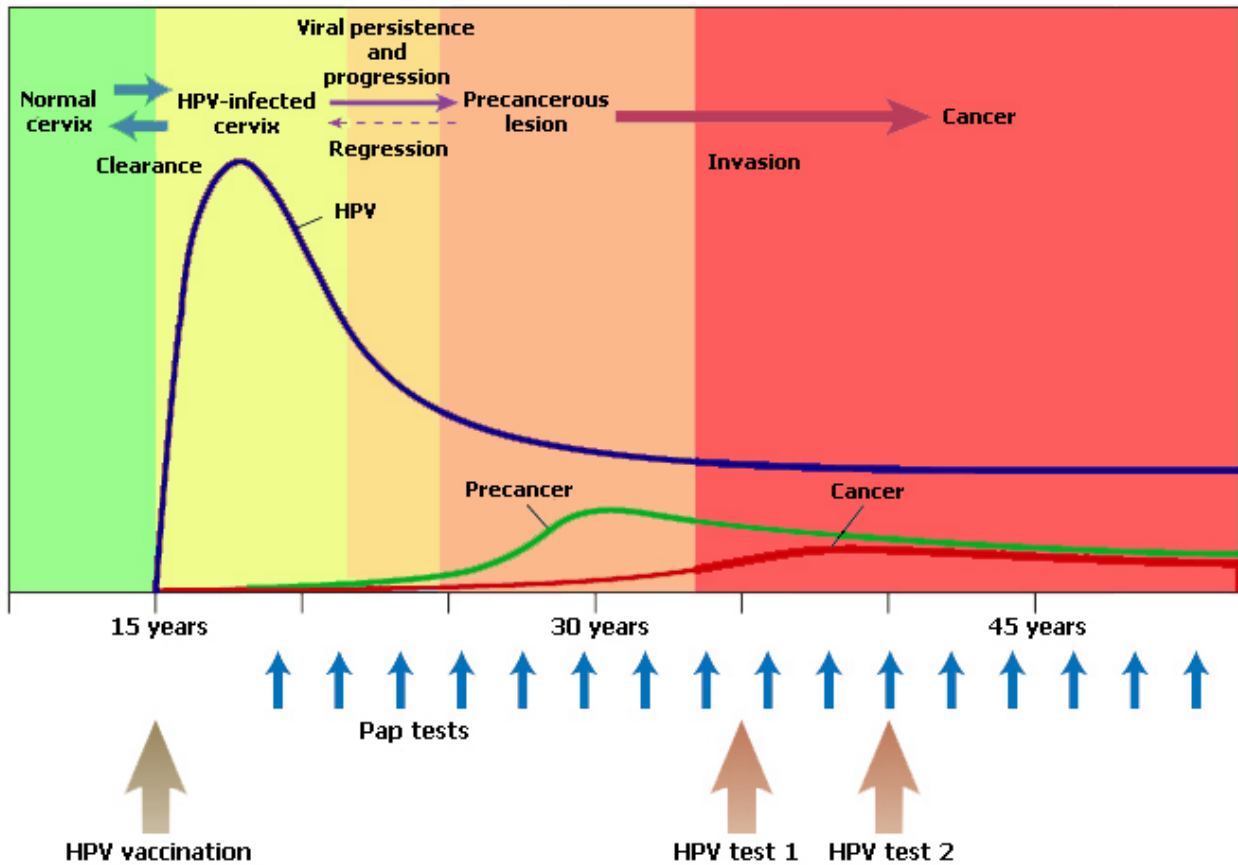
Figure 3.: The Natural History of HPV Infection and Cervical Cancer. The graph depicts a prevalence of transient infections with carcinogenic types of genital HPV (blue line) after the initiation of sexual activity. Most HPV infections are spontaneously cleared by the immune system. The development of pre-cancerous lesions (green line) occurs approximately 10 years later. Progression to invasive cervical cancer (red line) generally occurs after many years of persistent infection with a cancer-causing HPV type (peaks are not drawn to scale). The conventional model of cervical cancer prevention is based on repeated rounds of cytologic examination, including Papanicolaou smears and colposcopy (small blue arrows). Alternative strategies include HPV vaccination of adolescents (large grey arrow), one or two rounds of HPV screening at the peak ages of treatable precancerous conditions and early cancer (large redish-brown arrows) or both. The figure is based on information from Schiffman & Castle, (2005).

skin-to-skin contact. Most HPV types are adapted to infection of particular body surfaces[16].

A separate group of about 30 HPV types are typically transmitted through sexual contact. Genital HPV infection is very common, with estimates suggesting that up to 75% of women will become infected with one or more of the sexually-transmitted HPV types at some point during adulthood (Baseman & Koutsky, 2005). Some sexually-transmitted HPVs, such as types 6 and 11, can cause irregular genital warts also known as condyloma accuminata. In general, these wart-causing HPV types are not associated with an increased risk of cancer. Persistent infection with a subset of 15 to 20 so-called "high-risk" sexually-transmitted HPVs can lead to the development of SIL[17]. Persistant SILs can progress to high–grade SIL and finally to cervical carcinoma *in situ* (zur Hausen, 2002). These low-grade and high-grade SILs are precursor lesions of cervical carcinoma which are characterised by an abnormal growth and differentiation pattern of the cervical epithelium. They are distinguished into different levels of CIN[18] which occur exclusively in the epithelium and do not affect the basal membrane cell layer. CIN is the abnormal growth of precancerous cells in the cervix. Most cases of CIN stay the same or are eliminated by the host's immune system but a small percentage of cases progress to become cervical cancer, usually cervical SCC[19] (Agorastos et al., 2005). Three different levels of CIN are distinguished. Grade I, or condyloma or CIN1, is the least risky type and represents only mild dysplasia or abnormal cell growth (Agorastos et al., 2005) and is considered a LSIL[20] (Park et al., 1998). Grades 2 and 3, or CIN2 and CIN3, are considered HSIL[21] (Park et al., 1998). They show moderate dysplasia in CIN2, severe dysplasia in CIN3, and can finally progress into cancer (Agorastos et al., 2005). CINs have been proposed by some authors to progress toward cancer through these three stages in a linear fashion. However, evidence suggests that cancer can occur without

---

[16] For example, HPV types 1 and 2 tend to infect the sole of the feet or the palms of the hands, respectively, where they may cause warts (de Villiers et al., 2004). These could be common warts, which are often found on the hands and feet, but can also occur on other areas, such as the elbows or knees. Cutaneous HPV types do usually not cause genital warts and are not associated with the development of cancer. Another type are the plantar warts which are found on the soles of the feet as well as subungal or periungal warts from under the fingernail (subungal), around the fingernail or on the cuticle (periungal). They may be more difficult to treat than warts in other locations. Finally there are the flat warts, which are most commonly found on the arms, face or forehead. In people with normal immune function, flat warts are not associated with the development of cancer.

[17] SIL: Squamous Intraepithelial Lesion

[18] CIN: Cervical Intraepithelial Neoplasia

[19] SCC: Squamous Cell Carcinoma

[20] LSIL: Low grade Squamous Intraepithelial Lesion

[21] HSIL: High grade Squamous Intraepithelial Lesion

Normal	LSIL	HSIL		SCC
	CIN1	CIN2	CIN3	
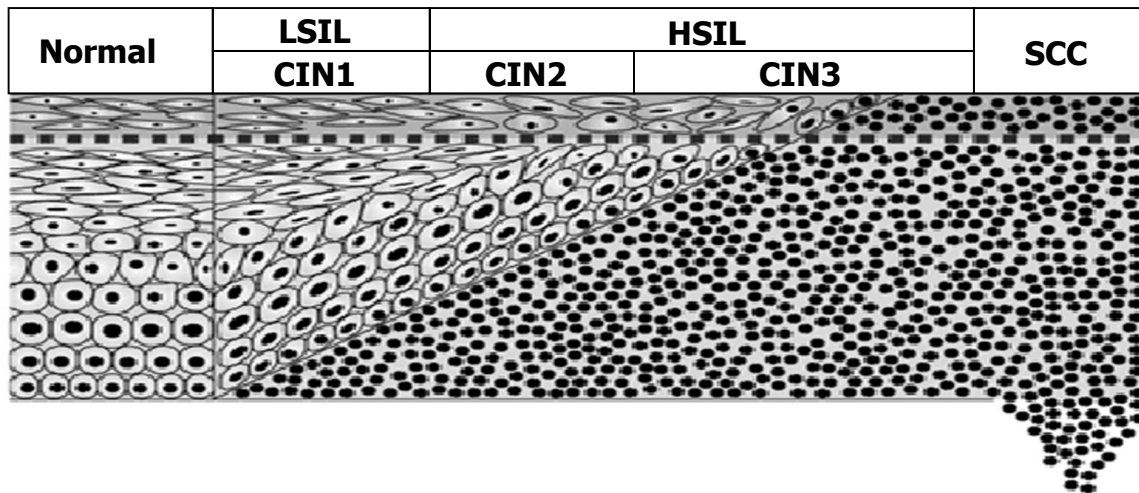
Figure 4.: Classification of the precancerous squamous intraepithelial lesions and a schematic representation of the changes in epithelial morphology. Starting from a normal epithelium, CIN1 is entered. Dysplastic cells are exclusively located in the inferior layers of the epithelium (basal and parabasal cells). In CIN2 dysplasic cells are located in the lower two third of the epithelium (reaching the intermediate cells). In CIN3, the dysplasia is localised over the entire epithelium (including superficial and squamous cells) with a loss of cell polarity. CIN3 can finally progress into invasive neoplasia and SCC (Figure adapted from a histopathology lecture of G. Fechter).

FIGO stage			Description
Stage 0			Preinvasive carcinoma *in situ* (epithelium but no invasion into the stroma)
Stage I			Confined cervical carcinoma (limited to the uterus)
	IA		Invasive carcinoma (diagnosed only by microscope, no visible lesions)
		IA1	Stromal invasion (less than 3mm in depth and 7mm in horizontal spread)
		IA2	Stromal invasion (between 3-5 mm in depth and less than 7 mm in spread)
	IB		Visible lesions (5 mm in depth and 7 mm in horizontal spread)
		IB1	Visible lesion (less than 4 cm)
		IB2	Visible lesion (more than 4 cm)
Stage II			Tumour invasion beyond the uterus
	IIA		Without parametrial invasion
	IIB		With parametrial invasion
Stage III			Tumour extends to pelvic wall or lower third of the vagina
	IIIA		Involves lower third of the vagina
	IIIB		Extends to pelvic wall and/or causes nephrosis or non-functioning of the kidney
Stage IVA			Tumour invades the mucosa of bladder or rectum and/or extends beyond pelvis
Stage IVB			Distant metastasis

Table2.: Classification of cervical cancer according to the FIGO staging system. Clinical parameters are used to classify cervical neoplasia following the Fédération Internationale de Gynécologie et d'Obstétrique. The alternative Tumour-Node-Metastasis (TMN) staging system is analogous to the FIGO stage.

first detectably progressing through these stages and that a high grade intraepithelial neoplasia can occur without first existing as a lower grade (Agorastos et al., 2005). Figure 4 illustrates the linear progression of a normal epithelium through the different CIN stadia ending in a SCC. Subsequent carcinoma progression concerning cervical carcinoma only are summarised in table 2 according to the FIGO[22] staging system, based on clinical examination. The IARC confirms as an evidence that infection with HPV of type 16, 18, 31, 33, 35, 39, 45, 51, 52, 56, 58, 59 or 66 can lead to cervical cancer, with HPV 16 and 18 displaying with 50-60 % the highest incidence in the case of cervical neoplasia. HPV infection is a necessary factor. However, the process of transforming normal cervical cells into cancerous ones is rare and slow. Cancer occurs in women who have been infected with HPV for a long time, usually over a decade or more (Sinal and Woods, 2005). Figure 3 illustrates the natural and age dependent infection cycle of HPV and the cancer risk. Cancers caused by a dozen or so "high-risk" HPV types kill several hundred thousand people per year worldwide and are a major focus of public health research (Schiffman & Castle, 2003).

Sexually-transmitted HPVs also cause a major fraction of anal cancers and a small fraction of cancers of the mouth and upper throat (known as the oropharynx). HPV types 6 and 11 can cause a rare pathology known as recurrent respiratory papillomatosis, in which warts form on the larynx or other areas of the respiratory tract. These warts can recur frequently, may require repetitive surgery, may interfere with breathing, and in extremely rare cases can progress to cancer (Sinal and Woods, 2005). Other HPV, such as HPV 5 or 8, could be identified in skin cancers occurring in patients suffering from EV[23].

HPV – the virus particle

Papillomaviruses are small double stranded DNA viruses belonging to the papovavirus family. They are non-enveloped viruses. The capsid of the virus is not

---

[22] FIGO: Fédération Internationale de Gynécologie et d'Obstétrique

[23] EV, epidermodysplasia verruciformis, is a rare, lifelong, autosomal recessive hereditary disorder affecting the skin, correlated with chronic infection with HPV. Flat-to-papillomatous wartlike lesions and reddish-brown pigmented plaques on the trunk, the hands, upper and lower extremities as well as in the face are characteristic and may transform into malignant skin tumours, mainly on sun-exposed areas. EV patients are usually infected with multiple types of more than 30 cutaneous HPV. HPV types 5 and 8 are found in more than 90 % of EV-associated SCC. Individuals with EV have a specific impaired cellular immunity which make them susceptible to widespread viral infection. Many of the HPV types found in EV lesions are non-pathogenic to the general population. Renal transplant patients and immunosuppressed patients have an increased risk of developing EV lesions. In addition failure of programmed cell death may play an important role in malignant transformation of the epithelium.
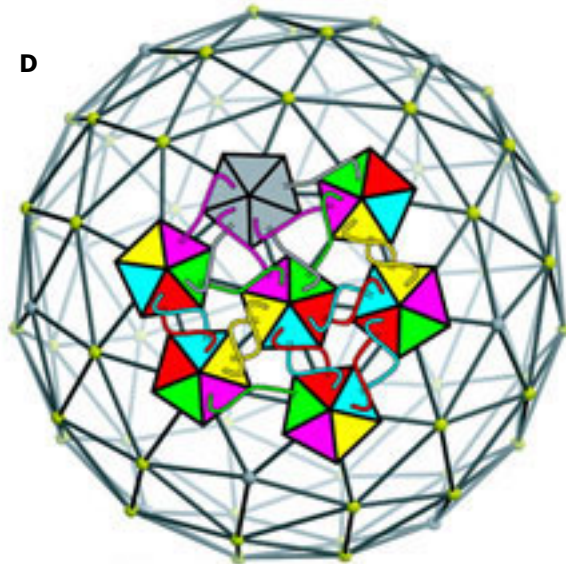
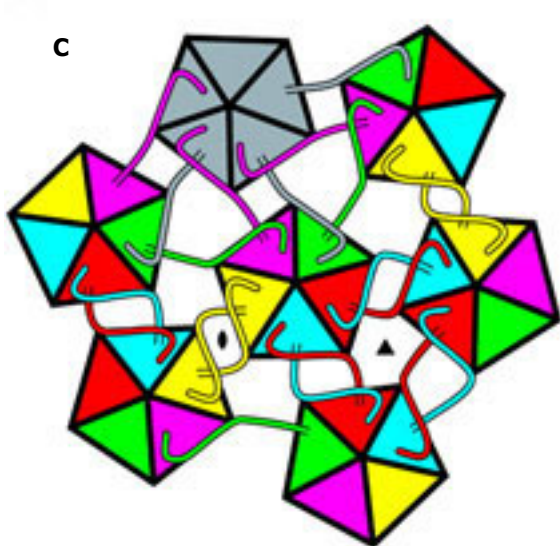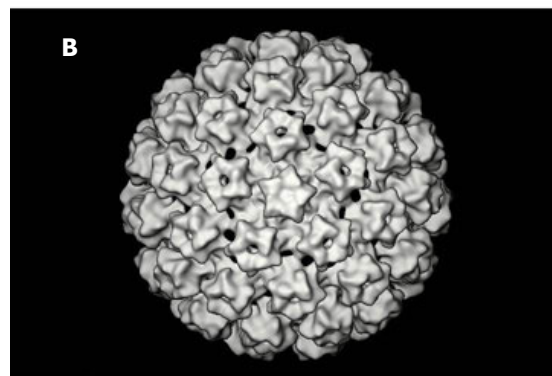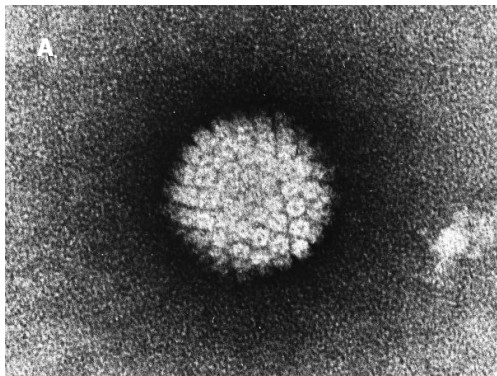**Figure 5.:** **Papilloma virus particles A/** Electron micrograph of a negatively stained HPV which occurs in human warts. **B/** Structural features of bovine papillomavirus capsid revealed by a three-dimensional reconstruction to 9 Å resolution (Trus et al., 1997). **C/** Pentamer contacts in papillomavirus particles. **D/** Same diagram as in C/ superimposed on the full papillomavirus lattice (Modis et al., 2002).

covered by a lipid membrane. A single viral protein, known as L1, is necessary and sufficient for formation of a 60 nanometer capsid composed of 72 capsomers (Figure 5). Like most non-enveloped viruses, the capsid is geometrically regular and exhibits icosahedral symmetry[24]. The papillomavirus genome is a double-stranded circular DNA molecule which is packaged within the L1 shell.

The papillomavirus capsid also contains a less-abundant viral protein known as L2. Although it is not clear how L2 is arranged within the virion, it is known to perform several important functions, including the packaging of the viral genome into nascent virions and facilitating the infectious entry of the virus into new host cells[25].

HPV – the life cycle

High risk papillomaviruses replicate exclusively in keratinocytes, which form the surface layers of the walls of the vaginal mucosa. These surface tissues, also known as stratified squamous epithelia, are composed of stacked layers of differentiating keratinocytes, which become gradually specialised. These keratinocytes are continuously renewed by less-specialised stem cells resident in the epithelial basal layer (Figure 6, left panel). Less-differentiated keratinocyte stem cells are thought to be the initial target of productive papillomavirus infections. Subsequent steps in the viral life cycle are strictly dependent on the process of keratinocyte differentiation. As a result, papillomaviruses can only replicate in differentiating body surface tissues. HPVs are thought to gain access to keratinocyte stem cells through microwounds, symbolised by a deep crevasse through epithelial cell layers in Figure 6. After successful infection of the basal cell layer, the virus expresses very low levels of early viral proteins, which are responsible for enhancing proliferation of the basal layer (zur Hausen, 2002) and replicating and maintaining the viral DNA as a low copy episome (10-100 copies per cell) (Oh et al., 2004). Undifferentiated keratinocyte cells of the epithelial basement layer can maintain papillomavirus genomes in a latent state for decades (Doorbar 2005). Induction of the viral life cycle as well as early and late gene expression occurs only in differentiating cells. The expression of the viral late genes, coding for the capsid

---

[24] Self-assembled virus-like particles composed of L1 are the basis of a successful group of prophylactic HPV vaccines designed to induce a immune response in the host and production of virus-neutralising antibodies which protect against initial HPV infection (see later paragraph for more detail).
[25] L2 is of interest as a possible target for more broadly-protective HPV vaccines.
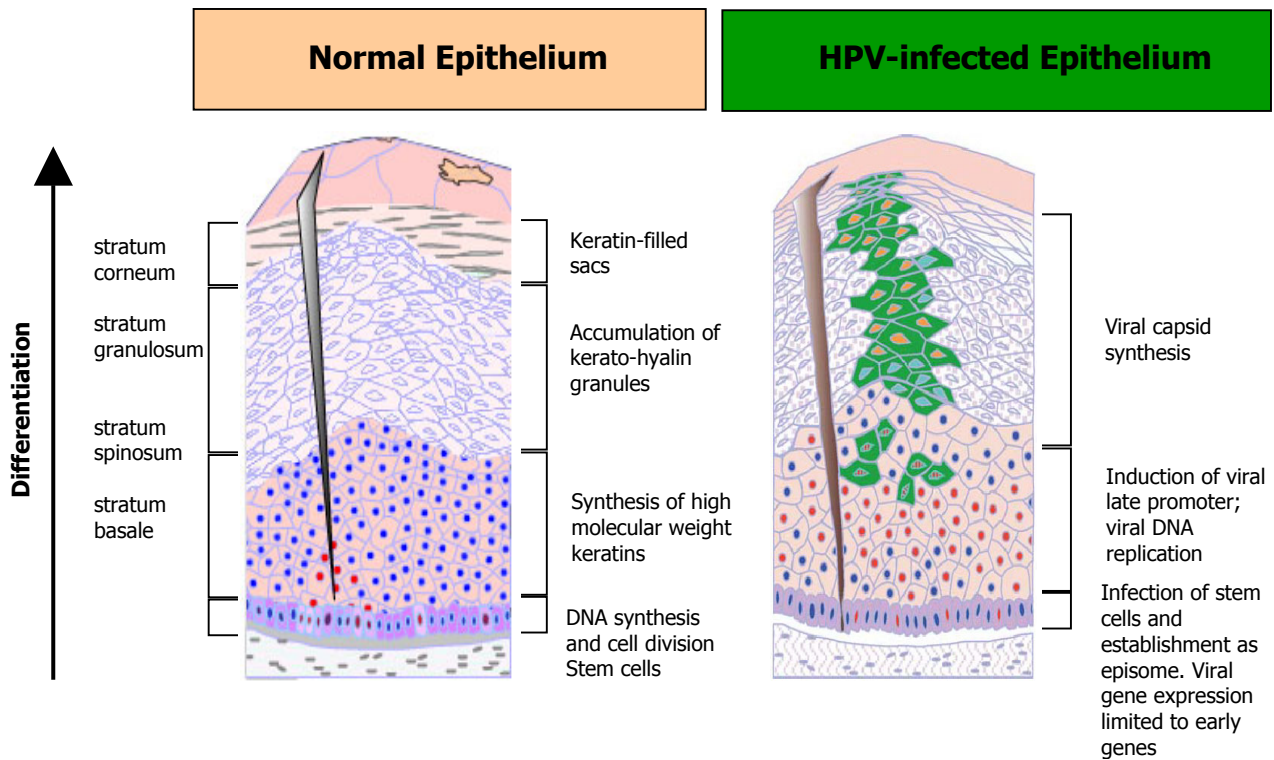
Figure 6.: Comparison of differentiation dependent functions in a normal and HPV-infected stratified squamous epithelum. Normal keratinocytes lose progressively their nuclei as they differentiate into the surface cells of the stratum corneum. HPV are suspected to infect the epithelium through microwounds. HPV-infected cells, however, maintain their nuclei and induce expression of the late viral genes as they move through the various differentiating stages. Finally, new virus particles are released by natural desquamation (not shown). Adapted from Stubenrausch & Laimins, 1999 and Doorbar, 2005.
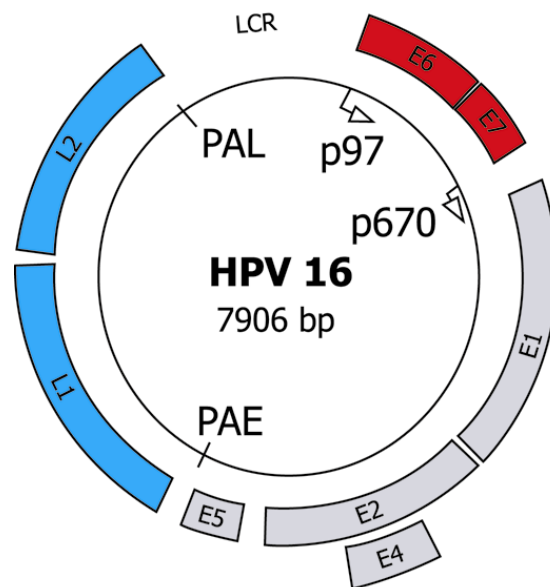
Figure 7.: Genome organisation of the high-risk HPV 16. The localisations of the open reading frames for the major proteins are indicated with boxes. Proteins labelled E1 – E7 are early viral proteins. Proteins labelled L1 and L2 are the late proteins. They constitute the „early region" and the „late region" on the viral genome, respectively. The long control region (LCR) is situated between the L2 and the E6 ORF. PAE and PAL depicts the polyadenylation signals of the early and the late region, respectively. Proteins translated from transcripts at the constitutive early promoter p97 are expressed in the basal layers of the epithelium. Transcripts initiated by the inducible promoter p670 are expressed in differentiated cells and during the replication of viral DNA and the assembly of viral particles.

proteins is exclusively restricted to differentiating keratinocytes and their expression levels correlate with an increase of the viral episome copy number. New infectious progeny virus particles are assembled in the surface layers of the stratified squamous epithelium and released by desquamation (Figure 6, right panel). It has been proposed that the differentiation-dependence of the HPV life cycle has been associated to yet unidentified cellular factors (Longworth & Laimins, 2004b). In addition high-risk and low-risk HPV can be distinguished according to their spatial and temporal difference in DNA replication within the epithelium. Low-risk HPVs have the tendency to replicate DNA in the less differentiated basal layer of the epithelium, where the cellular DNA replication machinery is present and active. High-risk HPVs replicate in higher levels of the epithelium and thus need more aggressive strategies to reactivate the normally quiescent cellular replication machinery, which could possibly be correlated with the malignancy of the high-risk HPVs (Sterlinko & Banks, 2004).

HPV – genome organisation and viral proteins

The HPV genome is a double stranded circular DNA of about 7900 base pairs (Figure 7) and varies in exact size according to the HPV type. The viral ORFs[26] occupy only one strand of the DNA and encode for early (E) and late (L) proteins, according to their time dependent expression during the viral life cycle. Most of the ORFs are polycystronic. Viral transcripts are differentially processed through alternative splicing and translated by ribosome scanning. Several splice variants rendering truncated or fused proteins have also been identified. The papillomavirus genome can be divided into three regions: 1) a LCR[27] or URR[28], 2) an early region (E) and 3) a late region (L). The LCR contains the origin of replication as well as regulatory signals for the viral transcription. It bears also one of the major promoters, which initiate expression of the early or late gene products, designed as p97. Another major promoter is p670 with respect to its nucleotide number which lies outside the LCR in the early E7 reading frame (Figure 7). p97 is a constitutively active promoter whereas p670 is induced differentiation-dependent.

---

[26] ORF: Open Reading Frame
[27] LCR: Long Control Region
[28] URR: Upstream Regulatory Region
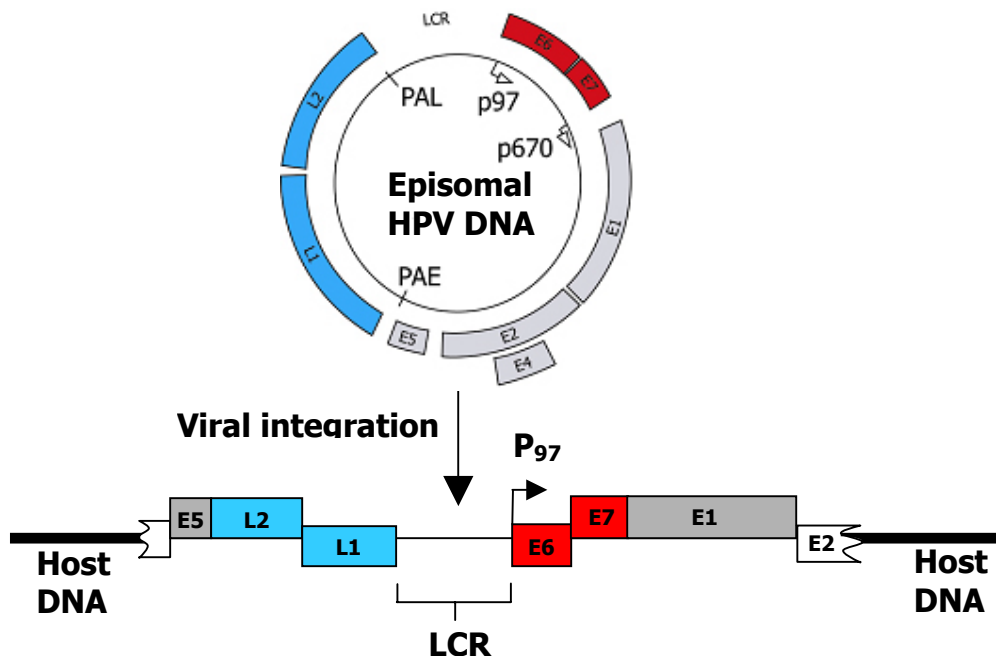
Figure 8.: Schematic representation of the integration of the viral genome into the host DNA. Integration occurs in most of the cancerous cells and seems to be a decisive event for progression and cancer. Integration often disrupts the E2 ORF and deletes the E4 ORF (Adapted from Dell & Gaston, 2001). In addition, parts of the E5 ORF or even the entire E5 ORF and parts of the L2 ORF frequently get disrupted during integration (Zur Hausen, 2002).
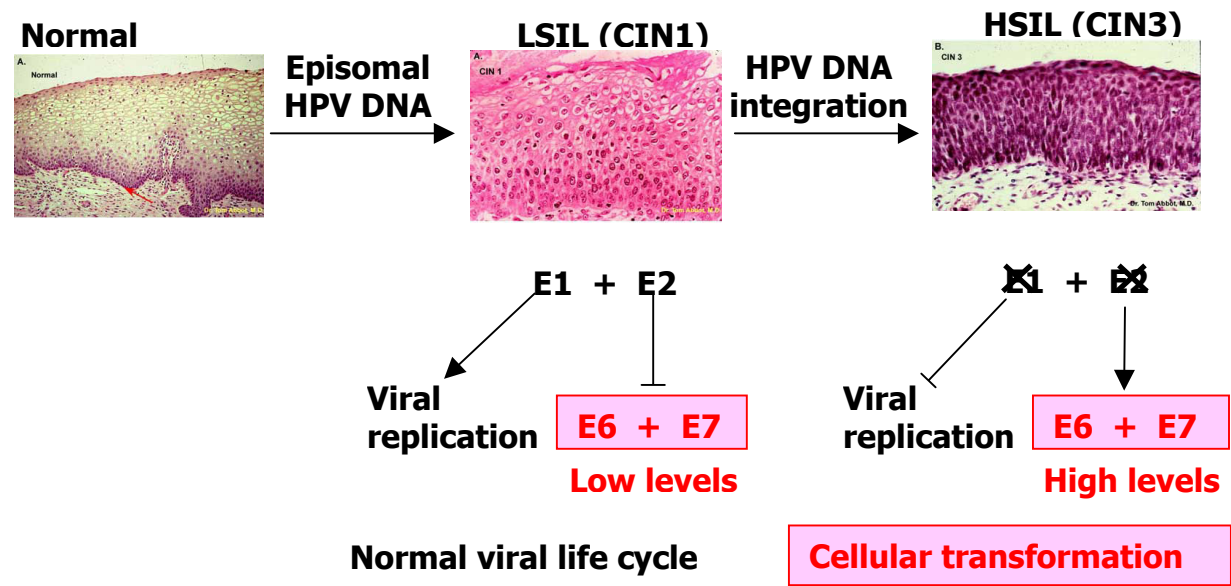
Figure 9.: Schematised representation of the integration dependent transformation process. During normal HPV life cycle and in LSILs, E6 and E7 levels are triggered to a basal level by E2 transcriptional repression. HPV genome integration on the host chromosomes induces disruption of the E2 ORF and loss of negative regulation of the E6 and E7 expression levels. Higher expression levels of the latter seems to correlate with malignant progression (based on Blachon & Demeret, 2003).

HPV induced cellular transformation

In benign HPV induced epithelial lesions or simply during the normal HPV live cycle, viral DNA is present as episomal DNA. In contrast, in HSIL and cervical cancer cells, viral DNA is found episomal as well as integrated on the host cell chromosomes. Therefore, integration of the viral DNA seems to play a decisive role for generation of HSIL, progression and cervical cancer (Dell & Gaston, 2001; Durst et al., 1985; Kalantari et al., 2001; Schwarz et al., 1985). Integration of the viral genome seems to occur at specific fragile sites. These sites are genomic regions prone to chromosome breaks, facilitating viral DNA integration. About 192 individual HPV integration sites have been described to date. Integration occurs late in the progression of HSIL and seems to be correlated to destabilisation of the genome integrity in replicating epithelial stem cells expressing the E6 and E7 oncoproteins. However, even if in some cases, integration of the viral genome disrupts cellular genes known to be involved in tumour development or in other cancer entities (*myc, tp63, nr4A2, amp-1, fancc, tnfaip2* or *htert*), no preferential integration sites could be identified in the host genome. Progression of HSIL and cervical cancer seem to be essentially linked to modifications of the viral genome and deregulation of viral protein expression upon integration and in a lesser extent in the disruption of cellular genes at specific loci (Wentzensen et al., 2004).

Viral genome integration on the host chromosomes results in a linearisation of the viral DNA and commonly in the disruption of the E2 ORF and consequently the loss of the E2 and E4 proteins (Figure 8). Disruption of the E5 and L2 gene and subsequent loss of these proteins has also been observed (Baker et al., 1987; Corden et al., 1999; Scheffner et al., 1994; zur Hausen, 2000). Loss of the E2 protein might be important in cervical tumorigenesis since it down-regulates the transcription of the E6 and E7 oncoproteins. Increased levels of E6 and E7 seem to correlate with progression of HSIL and cancer (Figure 9) (Dell & Gaston, 2001; Munger et al., 2004). However, other studies show that this hypothesis might not be the whole story and that changes in the expression levels of E2 might also play a direct role in tumorigenesis (Bellanger et al. 2005; Blanchon et al., 2005; Grm et al., 2005).

Different studies show clearly that there is a correlation between high-risk HPV infection and deregulated E6 and E7 expression in transformed cells (Goodwin & DiMaio, 2000; Wells et al., 2000). Mainly E6 and E7 genes are consistently expressed

in cervical cancers, suggesting a model in which only a minimal fragment of the viral genome is required for maintenance of proliferative tumour cells (Hawley-Nelson et al., 1989; Hudson et al., 1990; Munger et al., 1989a). This theory has been affirmed by different experimental approaches aiming at inactivating the biological functionof E6 and E7 in vivo and in vitro. Experiments using shRNA[29] against HPV 16 E6 showed reduced expression of E6 and E7 in SiHa cells and suppression of cell proliferation, tumorigenecity and in vitro cell invasiveness (Bai et al., 2006). Similar experiments were performed using either rybozymes or siRNA. It has been shown that transformed phenotypes of HeLa cells could be inhibited by an anti HPV 18 rybozyme (Chen et al., 1996) and that a hairpin rybozyme against HPV 16 E6 prevents immortalisation of the cells (Alvarez-Salas et al., 1998). Finally si RNA directed against HPV 16 E6 in SiHa or Hela cells showed suppression of E6 activity and restoration of apoptosis (Butz et al., 2003).

However, focusing on E6 and E7 remains quite a reductionist view of the topic and it can be assumed that HPV infected cells need still additional mutations in the host cell genome or other factors besides viral genome integration.

The HPV proteins

As already mentioned in the paragraph above, there are two groups of viral proteins according to their expression during the viral life cycle. There are the early proteins, namely E6, E7, E1, E2 transcribed from the early promoters, as well as later in the viral life cycle E4 and E5 and the late proteins L1 and L2. The late proteins are structural proteins and important for the capsid assembly and the viral DNA packaging. The early proteins are mainly regulatory proteins. Among these, some have oncogenic properties in high-risk HPV. This paragraph aims to describe briefly the function of each of these proteins in the natural viral life cycle.

HPV late proteins

L1 and L2 proteins (55 kDa and 75 kDa, respectively) are implicated into viral encapsidation. They are transcribed and translated in cells of the stratum granulosum and stratum corneum during advanced stage of the viral life cycle

---

[29] shRNA: small hairpin RNA
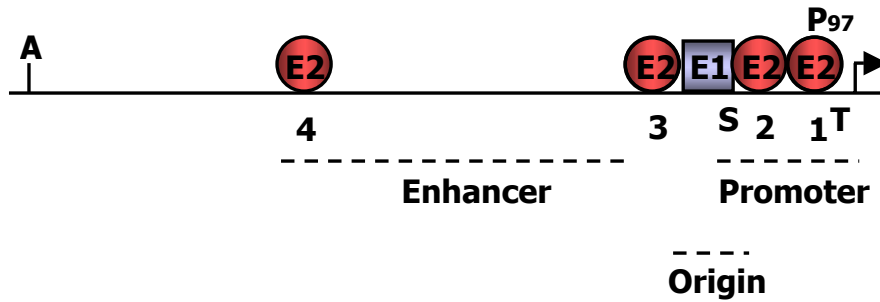
Figure 10.: Magnification of the LCR region of the HPV 16 genome. The bent arrow represents the transcription start of the p97 promoter. The four E2 binding sites are numbered 1-4. The localisation of the TATA box (T), the SP1-binding site (S) as well as of the enhancer, the origin of replication and the promoter are indicated (Dell & Gaston, 2001).
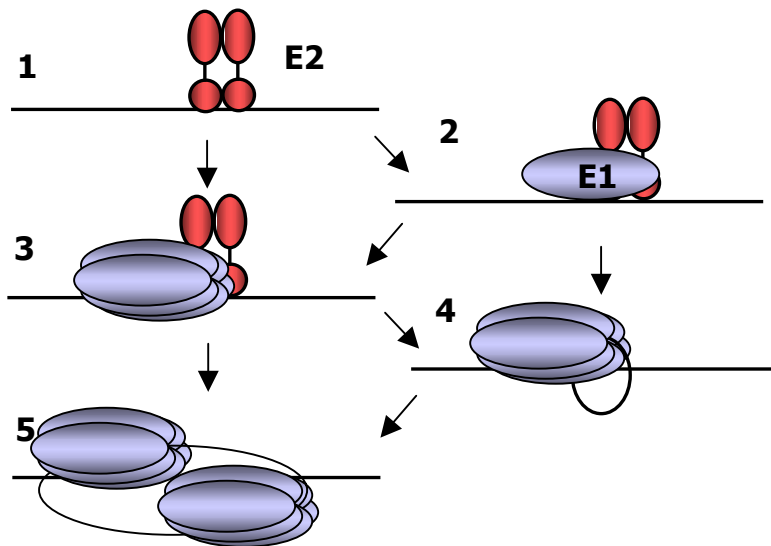
Figure 11.: Schematic representation of a possible HPV replication complex. Site specific binding of E2 in the LCR (1) leads to either E1 monomer (2) or hexamer (3) recruitment. The E2-E1-hexamer complex (3) may lose E2 (4) to be converted to a double E1-hexamer complex (5). The second E1-hexamer might be recruited by the E2-E1-hexamer complex. The double E1-hexamer complex may also form spontaneously after E2 displacement (5) (Figure adapted from Dell & Gaston, 2001).

(Figure 6). They localise to the nucleus via a NLS[30] where DNA packaging and self assembly to viral particles occurs spontaneously. L1 is the major capsid protein which is sufficient for capsid formation. Purified L1 protein alone can form capsids, which are stabilised by disulfide bonds between neighbouring L1 molecules. L1 capsids assembled in vitro are the basis of prophylactic vaccines against several HPV types. Compared to other papillomavirus genes, the amino acid sequences of most portions of L1 are well-conserved between types. L2 represents a minor capsid protein. L2 has been identified as being able to bind DNA (Zhou et al., 1994) and it is thought to cooperate with L1 for packaging viral DNA into the viral particles. In addition it has been shown that the L2 interacts with cellular proteins during the infectious entry process. After binding to the cell, L2 is cleaved by a cellular protease (Richards et al., 2006), the virus internalised into the endosome pathway and released by acidification of the endosomes into the cytoplasm. After endosome escape, L2 and viral DNA are thought to be imported into the nucleus.

The viral particle measures 50-55 nm in size and is built by 72 capsomers assembling to a icosahedral capsid (Modis et al., 2002). Each capsomer is built of 5 copies of the L1 protein.

HPV early proteins

E1 and E2 proteins

The early proteins E1 and E2 are the first viral proteins expressed in the basal cell layers during the natural viral life cycle. Transcription is induced by cellular transcription factors via the early constitutive promoter p97.

E1 is an ATP dependent helicase of approximately 68 kDa which recognises and binds regions in the ORI[31] of the viral genome located at the LCR (Figure 10). This complex is able to recruit cellular proteins necessary for viral genome replication (Bonne-Andrea et al., 1995; Park et al., 1994). Interestingly, E1 has been reported to bind only poorly to the ORI in the absence of the E2 protein (Frattini & Laimins, 1994). Once E2 has enabled E1 binding to the ORI two distinct complexes have been described to be formed at the ORI (Figure 10, 11). One complex is formed of an E2 dimer and one E1 molecule which is thought to form an initiation complex (Lusky et

---

[30] NLS: Nuclear Localisation Signal
[31] ORI: Origin of replication

al., 1994). For the second complex, only E1 molecules assemble into hexamers which displace E2, thus forming the final initiation complex for further unwinding of the DNA and strand separation in the 3' to 5' direction which is necessary for viral genome replication (Fouts et al., 1999). It has been described that in absence of DNA, E1 can exist in a monomeric form. The monomer to hexamer transition could play a regulatory role for ORI recognition (Chen & Stenlund, 1998). Helicase activity seems to be divided into a N-terminal DBD[32] and a C-terminal ATP-binding domain (Titolo et al., 1999). The structure of an E1 DBD from BPV[33] and HPV 18 have been determined by X-ray crystallography (Enemark et al., 2000; Auster & Joshua-Tor, 2004).

E2 is known to be the major transcription factor which regulates the viral genome in the natural viral life cycle. E2 is a multi-functional protein of about 50 kDa which regulates transcriptionally the expression of virtually all HPV genes, including the oncogenes E6 and E7. It has been proposed that E2 can mediate cell proliferation by down-regulating the expression of the latter (Nishimura et al., 2000; Francis et al., 2000) or by direct effects which are not totally understood yet (Webster et al., 2000). When E2 expression levels are low, it acts as a transcription activator, whereas at high expression levels it acts as a transcriptional repressor by blocking binding of cellular transcription factors (Bouvard et al., 1994; Steger & Corbach 1997). As previously mentioned, interaction with E2 enhances the binding specificity of E1 to the ORI and activates helicase activity for the rolling circle DNA replication process. E2 has four binding sites in the LCR (Figure 10), which are conserved among the different high-risk HPV types suggesting that they are important for normal viral function. Two sites lay one next to the other between a Sp1[34]-binding site and the TATA-box[35] (Figure 10, sites 1 and 2, Figure 11). It has been described that binding of E2 to these sites results in transcriptional repression by hindering binding of Sp1 and TBP[36] (Dostatni et al., 1991; Tan et al., 1992; Tan et

---

[32] DBD: DNA Binding Domain

[33] BPV: Bovine Papilloma Virus

[34] SP1 is a human transcritpion factor of 80 kDa. It contains a zinc binding motif and binds directly to GC-rich elements usually located at −40 to −90 of the promoter and enhances gene transcription.

[35] TATA-box: DNA sequence in the promoter region of genes, usually at −20 to −25 bp of the promoter, where transcritpion factors bind. It contains a highly through evolution conserved 5'-TATAA-3' sequence. The TATA-box is essential for binding RNA polymerase in prokaryotes as well as RNA polymerase II in eucaryotes.

[36] TBP: TATA-box Binding Protein. In eucaryotes this 30 kDa protein normally binds to the TATA-box in the process of transcription initiation. It is known as a commitment factor, which determines promoters for transcription. It binds different

al., 1994). The two other sites are further upstream (Figure 10, sites 3 and 4) and binding of E2 to these sites has been described to activate transcription (Ham et al., 1994; Ushikai et al., 1994). The repressing or activating activity is probably monitored by E2 expression levels and affinity to the different binding sites (Bouvard et al., 1994; Steger & Corbach, 1997). HPV transformed cells lack the E2 protein and show an expected increased level of E6 and E7 which has been related to abnormal cell proliferation. Expression of BPV E2 protein induces growth arrest in cervical carcinoma cells (Dowhanick et al., 1995; Goodwin et al., 1998; Hwang et al., 1993; Hwang et al., 1996). In addition, the E2 proteins of HPV 16, 18 and 33 have been described to induce apoptosis in cervical carcinoma cell lines (Desaintes et al., 1997; Frattini et al., 1997; Webster et al., 2000), which is supposed to be partially related to the changes in expression levels of E6 and E7 (Nishimura et al., 2000; Sanchez-Perez et al., 1997). E2 seems to act like a tumour suppressor protein, by controlling viral protein expression and thus balancing pro- and anti-apoptotic functions of other early viral proteins. Disruption of the E2 ORF in infected cells increases the proliferation and the risk of developing CIN and cervical cancer.

E4 protein

The E4 mRNA is one of the major transcripts in HPV-induced lesions and is late promoter dependent (Chow et al., 1987; Longworth & Laimins, 2004b). E4 is a 17 kDa protein which is localised within the differentiating layers of the epithelium (Doorbar & Gallimore, 1987) and is not required neither for transformation nor for episomal persistence of viral DNA (Neary et al., 1987). The E4 ORF lacks a transcription initiation codon and is transcribed as a splice product starting from the E1 initiation codon resulting in an E1^E4 protein product containing the five first amino acids of E1 and the E4 ORF expression product (Knight et al., 2004). It is one of the most highly expressed HPV proteins, especially in the differentiated basal layers (Doorbar et al., 1997; Middleton et al., 2003; Peh et al., 2002). E1^E4 is mainly localised in the cytoplasm and has been described to form multimers (Pray & Laimins, 1995; Roberts et al., 1997; Sterling et al., 1993; Wang et al., 2004). It has been demonstrated to interact with cytokeratins which leads to the collapse of the

---

transcription associated factors, important for transcription initiation and allows the proper binding of the RNA polymerase to the transcription start.

cytokeratin network (Doorbar et al., 1991; Roberts et al 1993; Roberts et al., 1994a; Roberts et al., 1997; Wang et al., 2004). E1^E4 has been identified as a zinc finger protein (Roberts et al., 1994b). E1^E4 has been shown to play a role in both early stages and late stages of the viral life cycle. E1^E4 spliced mRNA could be identified in early viral transcripts and the protein itself could be localised in cultured basal-like human keratinocytes (Nakahara et al. 2005). Expression of E1^E4 correlates with viral genome replication, suggesting that this protein might play a role in viral genome amplification (Doorbar et al., 1997; Nakahara et la. 2005; Palefsky et al., 1991). It might also play a role in expression regulation by interacting with a cellular helicase, which is implicated in mRNA splicing and translation initiation (Longworth & Laimins, 2004b). In late stages E1^E4 has been shown to be implicated in suprabasal DNA synthesis (Peh et al., 2004). In differentiated cells expression of mutant E1^E4 exhibited reduced levels of viral DNA amplification and late gene expression (Wilson et al., 2007). In addition E1^E4 has been shown to be implicated in the perturbation of differentiation in the superficial layers of stratified squamous epithelia (Nakahara et al., 2005). In addition, E1^E4 of low and high risk HPVs have been shown to induce cell cycle arrest.

E5 protein

E5 transcripts have been found to be both transcribed by the constitutive promoter p97 as well as by the differentiation-dependant promoter p670. However, p670 seems to be the more important promoter *in vivo* (Glahder et al., 2003). In contrast to BPV E5 protein, the HPV E5 protein has been less studied than its BPV orthologue. HPV E5 has a weak transforming activity in contrast to the BPV E5 protein which is the major transforming protein (Rabson et al., 1986; Schiller et al., 1986; Valle & Banks, 1995). HPV E5 is a 9.4 kDa hydrophobic protein located mainly within the Golgi apparatus, and in less extent in the plasma membrane (Burkhardt et al., 1989; Tsai & Chen, 2003). It has been shown that BPV 1 and HPV 16 E5 associate with the membrane-bound proton ATP-ase, a component of the gap-junction complex (Conrad et al., 1993; Finbow & Pitts, 1993). HPV 16 E5 has also been found to interact with the EGF[37] receptor and the PDGF[38] receptor (Conrad et

---

[37] EGF: Epidermal Growth Factor
[38] PDGF: Platelet-Derived Growth Factor

al., 1994) giving rise to some transforming abilities. However, the open reading frame coding for E5 is frequently deleted in human cervical cancers (Schwarz et al., 1985). On the other hand there are relatively large amounts of E5 mRNA and E5 protein in anogenital LSIL (Kell et al., 1994; Stoler et al., 1992), giving rise to the speculation that E5 is not required for latent HPV infection but plays a role in early steps of HPV infection. It seems to play a role in regulating DNA synthesis for viral DNA replication (Genther et al., 2003) and to prevent DNA-damage induced apoptosis by stimulating cell growth via interaction with the afore mentioned receptors (zur Hausen, 2002).

E7 and E6 proteins

E7 and E6 are under control of the same promoter and are translated from a bicistronic RNA (Smotkin et al., 1989). The E6/E7 mRNA is present at low levels in the basal layer of epithelia and its expression seems to correlate with cellular immortalisation (Steenbergen et al., 1998). Higher levels of E6/E7 mRNA can be found in tumours and seem to be correlated with transformation (Stoler et al., 1992). Up to now, little is known about whether this correlates with higher levels of E6 and E7 proteins. Expression of E7 from the polycystronic mRNA seems to depend on ribosome scanning through the E6 ORF with a translation initiation at a specific start codon (Stacey et al., 2000) as well as from spliced mRNAs which do not encode for the full-length E6 protein (Smotkin et al., 1989).

E7 protein

The E7 protein is one of the two major oncoproteins of high–risk HPV, responsible for immortalisation and progression into cancer. Together with E6, E7 serves to suppress apoptosis and to promote cell cycle progression, thus rendering replication of the viral DNA possible. Expression of E7 is required for survival of cancer cell lines which are derived from HPV-induced tumours. E7 is the subject of intense research interest and is believed to exert a wide variety of other effects on infected cells.

A

```
HPV 45    MHGPRETLQE IVLHLEPQNE LDPVDLLCYE QLSES-EEND EADGVSHAQL PARR
HPV 16    MHGDTPTLHE YMLDLQP--- -ETTDLYCYE QLNDSSEEE-D EIDGPAGQAE PDR-
HPV 18    MHGPKATLQD IVLHLEPQNE I-PVDLLCHE QLSDS-EEND EIDGVNHQHL PARR
HPV 11    MHGRLVTLKD IVLDLQPP-- -DPVGLHCYE QLEDSSEDEVD KVDKQD--AQ PLT-
HPV 06    MHGRHVTLKD IVLDLQPP-- -DPVGLHCYE QLVDSSEDEVD EVDGQD--SQ PLK-


HPV 45    AEPQRH KILCVCCKCD GRIELTVESS AEDLRTLQQL FLSTLSFVCP WCATNQ
HPV 16    ---AHY NIVTFCCKCD STLRLCVQST HVDIRTLEDL LMGTLGIVCP ICSQKP
HPV 18    AEPQRH TMLCMCCKCE ARIELVVESS ADDLRAFQQL FLNTLSFVCP WCASQQ
HPV 11    ---QHY QILTCCCGCD SNVRLVVECT DGDIRQLQDL LLGTLNIVCP ICAPKP
HPV 06    ---QHF QIVTCCCGCD SNVRLVVQCT ETDIREVQQL LLGTLNIVCP ICAPKT
```

B

C

Figure 12.: Solution structure of the structured part of the E7 oncoprotein. A/ Sequence alignment of chosen E7 sequences. The CR1, CR2 and CR3 regions are outlined orange, red and magenta, respectively. Secondary structure elements of the folded C-terminus are indicated. Arrows stand for â-strand and bars for á-helices. Hydrophobic residues stabilising the hydrophobic core and the dimer interface are coloured blue. The zinc chelating cysteine residues are colored red. **B/** Schematic representation of the HPV 16 E7 protein. The CR1, CR2 and CR3 domain are depicted by arrows. The LXCXE motif and the conserved cysteines are typed in red. **C/** Solution structure of the folded part of the HPV 45 E7. Dimer subunits are coloured dark blue and light red. Residues of the hydrophobic core and the interface are shown in blue. The zinc chelating cysteines are coloured red (figure partly based on Ohlenschläger et al., 2006).

Structural description

E7 proteins contain approximately 100 amino acids (HPV 16 E7 contains 98 amino acids) and can be divided into three conserved regions labeled CR1, CR2 and CR3 (Figure 12 A,B) (Barbosa et al., 1990; Dyson et al. 1992). The solution structure of a HPV 45 E7 dimer has been measured and calculated by NMR[39] (Figure 12 C) Figure 12 A shows a sequence alignment of several E7 proteins against HPV 45, including HPV 16 (Ohlenschläger et al., 2006). The structure of this dimer is highly homologous to the X-ray structure of the low-risk HPV1a CR3 domain (Liu et al., 2006). The CR1 and CR2, comprising residues 1 to 15 and 16 to 38 in HPV 16 respectively, show partial sequence and functional homology to the adenovirus E1A protein as well as to the SV 40 large T-antigen (Phelps et al., 1992). The CR2 contains a conserved LXCXE[40] motif which is correlated with binding of E7 to the pRb[41] tumour suppressor protein (Figure 12 A,B) (Munger et al., 1989; Barbosa et al., 1990; Chellappan et al., 1992; Lee et al., 1998; Singh et al., 2005). The CR3 of HPV 16 spans the residues 39 to 98 and contains two zinc-binding motifs, which are separated by 29 residues. The latter mediate zinc dependent dimerisation. (Barbosa et al., 1989; Clements et al., 2000; Alonso et al., 2002) and participate in the interaction of E7 with a large number of diverse cellular proteins involved in cell cycle regulation and apoptosis, among them pRb and the CDKI[42] p21[CIP1] (Zwerschke & Jansen-Durr, 2000; Munger et al., 2001; Ohlenschläger et al., 2006 (for review)). Analytical gel filtration, NMR HSQC[43] and relaxation experiments showed clearly that the HPV 45 E7 oncoprotein forms a symmetric dimer of 12 kDa in solution at NMR buffer conditions (Ohlenschläger et al., 2006). NMR measurements on the CR3 domain alone and on the full length E7 protein showed evidence for a disordered N-terminus of each dimer, comprising the CR1 and CR2 as well as a highly structured CR3 domain with a novel zinc binding fold (Figure 12 C). NMR measurements and EXAFS[44] measurements in conjunction with structure calculations allowed to identify C66, C69, C99 and C102 as the zinc co-ordinating atoms in HPV 45, which correlates with the residues C58, C61, C91 and C94 in HPV 16. These findings are strongly

---

[39] NMR: Nuclear Magnetic Resonance

[40] residues 22-26; L, C and E represent the amino acids in the one letter code and X any amino acid

[41] pRb: Retinoblastoma protein. The "p" indicates that this protein is mainly present as a phosphoprotein in cells.

[42] CDKI: Cyclin-Dependent Kinase Inhibitor

[43] HSQC: Heteronuclear Single Quantum Coherence

[44] EXAFS: Extended X-ray Absorption Fine Structure

supported by a mutational study (McIntyre et al., 1993). Each HPV 45 E7 monomer exhibits a β1-β2-α1-β3-α2 topology. Briefly, β1 and β2 form an antiparallel β-sheet which is packed on top of the α1 helix. The C-terminus of the α1-helix elongates in the β-strand-like β3 structural element which ends in the α2-helix. The zinc co-ordinating CXXC motifs are located in the turn connecting β1 and β2 as well as in the C-terminal α2-helix. Dimerisation of E7 involves the α1-helices as well as formation of an intermolecular two-stranded antiparallel β-sheet between the β2-strand of one monomer and the β3-strand of the other (Figure 12 C). E7 alone or in cooperation with E6 has many additional features and functions, besides pRb and p21^{CIP1} interaction, which will be discussed later in the context of HPV mediated cell immortalisation and transformation.

Disruption of the cell cycle control and episomal maintenance

The major activity of the E7 oncoprotein consists in the proteasome mediated degradation of the retinoblastoma tumour suppressor family members or pocket proteins, namely pRB, p107 and p130 (Boyer et al., 1996). Pocket proteins are a family of closely related proteins which regulate the cell cycle and cellular proliferation (Grana et al., 1998). They contain a conserved "pocket" domain which enables them to bind members of the E2F family of transcription factors and other proteins. Each of these pocket proteins has a particular affinity for a set of proteins. pRb interacts with E2F-1, E3F-2, E2F-3 and E2F-4, whilst p107 and p130 interact with E2F-4 and E2F-5 (Muller & Helin, 2000). Rb proteins are major regulators of the natural cell cycle (Figure 13). The interaction of each pocket protein with E2F proteins is controlled via phosphorylation events mediated by CDKs[45] (Dell & Gaston, 2001). Hypophosphorylated Rb binds E2F and thus represses E2F dependent gene expression. This controls transition from G1 to S-phase by transcriptional repression of many genes involved in S-phase entry. During late G1-phase, pRb becomes hyperphosphorylated by cyclin E-CDK2 and cyclin D-CDK4 or cyclin D-CDK6, resulting in the release of E2F and activation of E2F dependent gene expression. Cyclin-CDK complexes are themselves regulated by kinase inhibitors such as p16^{INK4} or p21^{CIP1} for cyclin D-CDK or cyclin A-CDK and cyclin E-CDK, respectively (Sherr, 2000) (Figure 13).

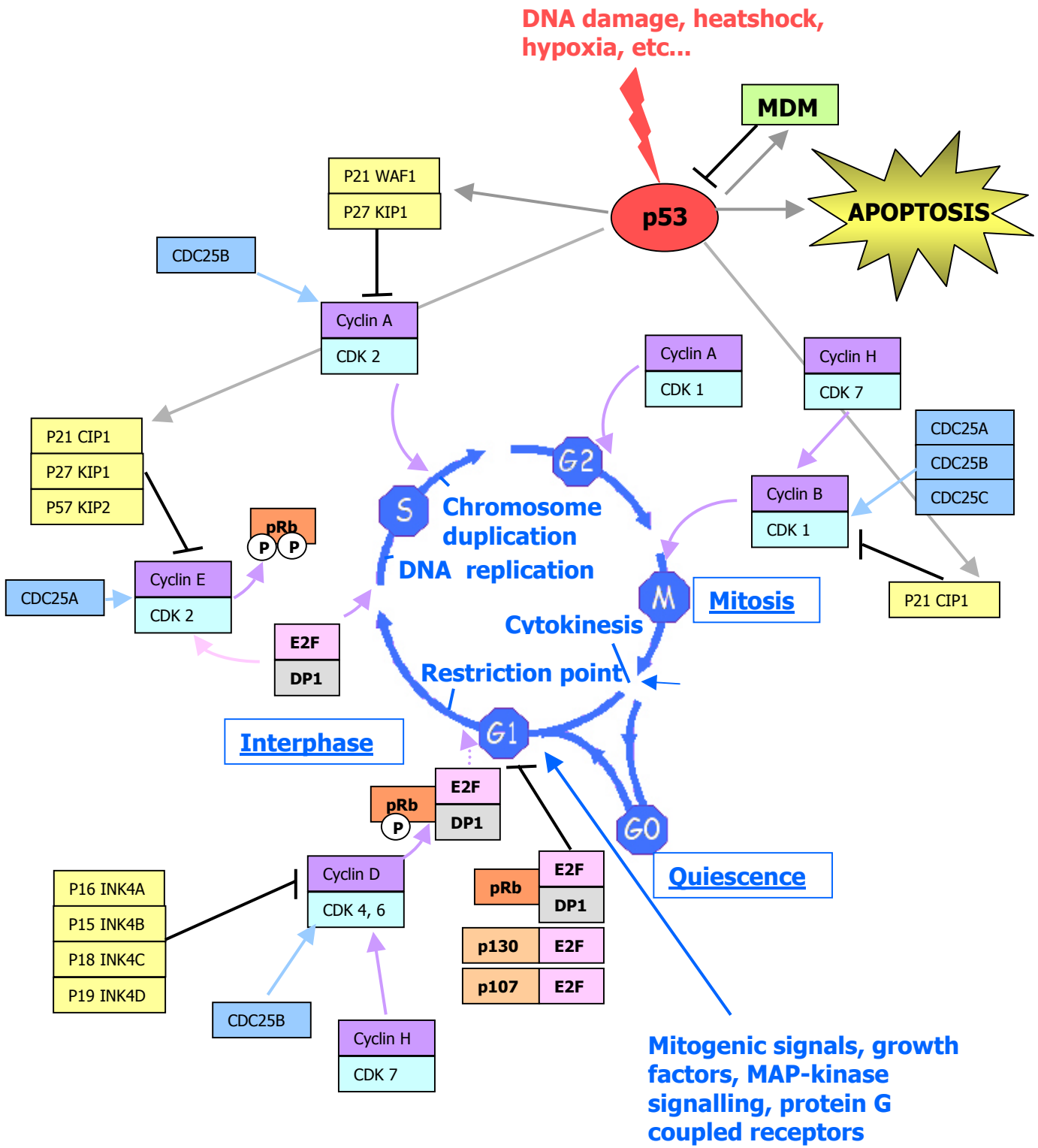
[45] CDK: Cyclin Dependent Kinase

Figure 13.: Schematic representation of the cell cycle.

Binding of E7 to these pocket proteins is mediated through the afore mentioned LXCXE motif located in the CR2 domain of E7, whereas both the CR1 and CR2 domains are required for E7 to induce the degradation of the pRb protein. In general, high-risk HPV E7 bind pRb with higher affinity than low-risk HPV E7. Mutations in the LXCXE consensus affect this interaction and influence the immortalisation ability of E7 (Flores et al., 2000). Although E7 and E2F do not bind the same site on the pocket protein, binding of E7 provokes the release of E2F (Lee et al., 1998) and E7-expressing cells keep on cycling even in the presence of high levels of $p16^{INK4}$ and $p21^{CIP1}$ (Jian et al., 1998; Martin et al., 1998). Therefore, binding of E7 to Rb and subsequent proteasome-mediated Rb degradation overrides these kinase inhibitors. In consequence the continuous E2F release brings permanent activation of the E2F-dependent genes inducing uncontrolled G1 to S-phase entry and thus cell proliferation (Figure 13). In this way, E7 also short-circuit p53 induced $p21^{CIP1}$-mediated cell cycle arrest in response to DNA damage (Figure 13). Interestingly the pRb degradation seems to be the decisive event as low-risk HPV 1 E7 binds to pRb but does not degrade it. (Jones et al., 1997). It has also been suggested that additional cellular proteins are required for efficient E7-mediated degradation of pRb (Berezutskaya et al., 1997; Longworth & Laimins, 2004b). E2F also positively regulates mitotic control genes. As a consequence, deregulation of E7 has a positive effect on the G2 to M-phase checkpoint (Thierry et al., 2004). However, the ability of pRb degradation and E2F release seems to be necessary but not sufficient for cell transformation of certain cell types (Banks et al., 1990). This means that E7 must deregulate other cellular pathways than the Rb-mediated ones.

Interaction with other proteins and other effects of E7

E7 mediated up-regulation of E2F has been correlated with abnormal centrosome duplication and aberrant spindle pole formation which seems to lead to genomic instability. HPV E6 oncoprotein seems not to have a direct influence on the centrosome number but it seems to potentiate the effects of E7 (Duensing et al., 2000). Thus correct chromosome segregation during mitosis in suprabasal cells is disturbed, which has become a marker for high-risk HPV lesions (Duensing & Munger, 2004).

E7 also binds to HDACs[46] via its C-terminal zinc-binding domain for which the integrity of the fold and the original amino acid at position 67 has been demonstrated to be critical (Brehm et al., 1999). HDACs repress transcription of genomic promoters through deacetylation of core histone proteins and in this way it remodels chromatin. E7 repression of HDACs has been associated with the maintenance of viral episomal DNA in HPV-infected cells (Longworth and Laimins, 2004a). Therefore, E7 reduces negative regulation and increases positive regulation of the cell cycle.

Furthermore, E7 has been described to interact directly with cyclin-kinase complexes as well as with CDKIs. E7 seems to be able to bind cyclin A-CDK2 and to interact indirectly with cyclin E-CDK2 by means of the p107 protein (McIntyre et al., 1996; Tommasino et al., 1993) and additionally, to increase the cellular levels of cylcins A and E (Ruesch & Laimins, 1998) (Figure 13). As both kinases hyperphosphorylate pRb *in vivo*, thereby inducing the release of E2F, it is believed that the interaction of E7 with these cyclins serves to enhance the activity of these kinases and keep the cells cycling.

E7 also binds the CDKIs $p27^{KIP1}$ and $p21^{CIP1}$, thus additionally blocking negative cell cycle control points (Funk et al., 1997; Zerfass-Thome et al., 1996).

E6 protein

The E6 oncoprotein is the second major oncoprotein of high risk HPV, responsible for transformation. It is a short lived protein which can be phosphorylated by PKA[47] (Kuhne et al., 2000) and PKN[48] (Gao et al., 2000). E6 is a multifunctional protein and has been a major focus of research. Since the expression of E6 is strictly required for maintenance of a malignant phenotype in HPV-induced cancers, it is an appealing target for diagnostic or therapeutic applications.

Structural description

E6 proteins contain approximately 150-160 amino acids and include 2 zinc-finger like motifs (Barbosa et al., 1989; Cole & Danos, 1987). E6 from HPV 16 is

---

[46] HDAC: Histone Deacetylase
[47] PKA: Protein Kinase A
[48] PKN: Protein Kinase N

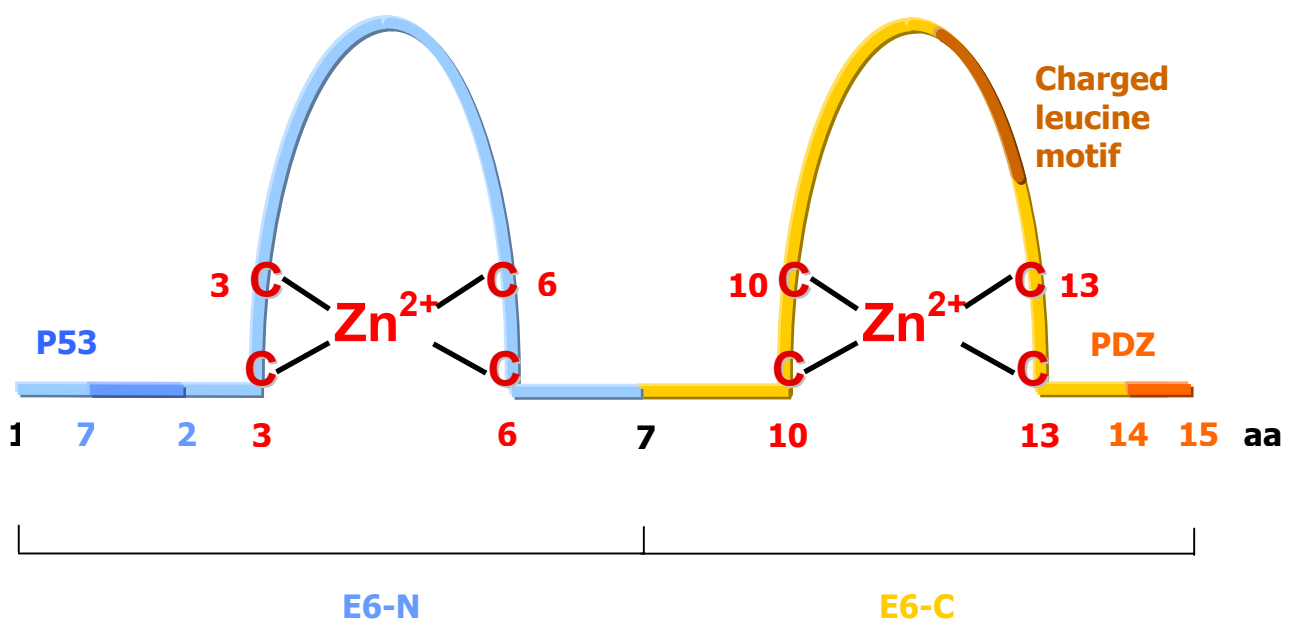**Figure 14.: Schematic representation of the HPV 16 E6 oncoprotein.** The E6-N domain is coloured in light blue and the E6-C domain in yellow. The eight conserved cysteine residues implicated in zinc coordination are represented in red. Annotations indicate interacting domains which were identified by mutagenesis experiments. E6-C has been described to bind to four-way DNA junctions (Ristriani et al., 2001).

made up of 158 or 151 amino acids, according to which methionine is used as a start codon (Figure 14). The E6 ORF shows clear homology to the E7 ORF and it has been supposed that they have been formed by duplication of a 33-37 residues domain containing a cysteine doublet (CXXC). E6 proteins contain four conserved CXXC motifs and sequence alignments suggested the presence of two zinc binding motifs (Cole & Danos, 1987). The spacing between the CXXC motifs in HPV is usually 29-30 amino acids, which is quite large. Zinc blotting assays (Barbosa et al., 1989; Cole & Danos, 1987; Grossman & Laimins, 1989; Kanda et al., 1991), site directed mutagenesis experiments (Lipari et al., 2001) as well as investigation by atomic and electronic absorption (Demers et al., 1994) showed clearly that E6 was a zinc-binding protein. High risk HPV produce also truncated forms of E6 by alternative splicing of the mRNA. Four spliced ORFs could be identified for HPV 16 and labelled E6I to E6IV (Doorbar et al., 1990; Sherman & Alloul, 1992; Smotkin et al., 1989). However, only the protein encoded by the E6I ORF, called E6*, could be identified (Schneider-Gadicke et al., 1988) and showed to have some transactivation activities (Shirasawa et al., 1994). In addition an E6* splice variant from HPV 18 has been detected, which encompasses residues 1-43 (equivalent to residues 1-42 of HPV 16) followed by 14 or 11 additional residues, according to the acceptor site for the alternative splicing[49]. It has been described to bind to full length E6 from HPV 16 and 18 as well as to E6AP but not to p53. It has been demonstrated that the binding of E6* to these proteins correlated with an antiproliferative activity of E6* in CaSKI cells by inhibiting E6 mediated degradation of p53 (Pim et al., 1997; Pim & Banks, 1999).

However, during a long period, remarkably little was known about the structure of E6. First E6 sequences were already described in the 1980s, but since then this small protein and its domains proved very difficult to produce in large quantities and as soluble and monomeric, correctly folded entities. In 1997 a research group claimed to have produced stable recombinant E6 protein both in bacteria and insect cells. This E6 protein seemed to be biologically active and was able to form stable homodimers (Daniels et al., 1997). However the precise expression and purification conditions were never published and no data about protein stability was available. A first improvement consisted in expressing E6 as a

---

[49] The RNA for the 14 additional residues codes for VPAVPETVESSRKT, which is the E6* splice variant which is usually referred to. The RNA for the 11 additional residues codes for ERLQRRRETQV. (personal communication from Dr. Murielle Masson)

fusion to a carrier protein. Frequently used carrier proteins are GST[50], MBP[51] or TRX[52]. Expression as a fusion protein usually increases expression levels and allows affinity purification. In addition, the afore mentioned carrier proteins have been described to facilitate folding of the passenger protein (E6), which otherwise would accumulate in solid inclusion bodies (Kapust & Waugh, 1999; Lorenzo et al., 1997; Shih et al., 2002). In this way, by using a GST fusion vector system, another group managed to purify a mutant of the 151-residue form of the E6 protein bearing a short C-terminal deletion (E6 Δ143-151) at very low concentrations[53] as well as its independently folding N-terminal domain (E6-N, residues 1-77). The E6 protein seemed to be biologically active. Again E6 Δ143-151 seemed to be able to oligomerise in a salt dependent manner with a monomer-dimer equilibrium being the best fit of the experimental data collected. ESI-MS[54] and Zn^{2+} substitution with Co^{2+} followed by atomic absorption analysis demonstrated that E6 binds 2 equivalents of Zn^{2+} and E6-N 1 equivalent. Circular dicroism measurements revealed a secondary structure content of 26% α-helix and 27% β-sheet for E6 Δ143-151 and 25% α-helix and 29% β-sheet for E6-N (Lipari et al., 2001). However, E6 yields after proteolytic cleavage of the carrier protein were always substantially lower than expected due to aggregation and nonspecific binding to surfaces during purification (Lechner & Laimins, 1994; Lipari et al., 2001). MBP had been described as a folding helper carrier protein (Kapust & Waugh, 1999) but strategies based on increasing the fraction of folded E6 protein failed. Furthermore, light scattering and ELISA[55] experiments demonstrated that MBP is able to keep misfolded passenger proteins as soluble inclusion bodies (SIB) in solution, rendering classical pellet/supernatant assays for protein quality monitoring unreliable (Nominé et al., 2001a). These SIBs could be formed by up to 400 fusion protein moieties and renaturation attempts suggested that they were stabilised mainly by hydrophobic interactions between misfolded E6 (Nominé et al., 2001a). In addition misfolding of E6 could partly be related to an unusually high proportion of cysteine residues in this protein (16 in HPV 16) which could promote intermolecular disulfide bonding and aggregation (Ristriani

---

[50] GST: Glutathion S-Transferase
[51] MBP: Maltose Binding Protein
[52] TRX Thioredoxin
[53] 30 nM were obtained in this study; concentration required for NMR analysis at that time was 600-1000 nM
[54] ESI-MS: Electrospray Ionisation Mass Spectrometry
[55] ELISA: Enzyme-Linked ImmunoSorbent Assay

HPV16Rct	CYSLYGTTLEQQYNKPLCDLLIRCINCQKPLCPEEKQRHLDKKQRFHNI.RGRWTGRCMSCCRSSR...T..RRETQL
HPV18Rct	SDSVYGDTLEKLTNTGLYNLLIRCLRCQKPLNPAEKLRHLNEKRRFHNI.AGHYRGQCHSCCNRARQERLQRRRETQV
HPV34ct	NQSVYGRTLENLTNKQLCNILIRCGKCQKPLCPLEKQRHVDENKRFHQI.ADQWTGRCTQCWRP.........SATVV
HPV26ct	TCSVYGATLEALTKKSLCNLLIRCHRCQMPLGPEEKQRIVDEKRRFHEI.AGQWKGLCTNCWRPRRQ......TETQV
HPV53ct	NCSVYGASLEALTKKKLSDLSIRCYRCQHPLTPEEKQLHCDYKKRFHKI.SHMWTGSCLTCWRHTTA.....TESAV
HPV40ct	RYAAYAPTVEEETGLTILQVRIRCCKCHKPLSPVEKTNHIVKKTQFFKL.KDSWTGYCLHCWKKCMEKGQR..SETLC
HPV27ct	HYSCYGDTVETETGIPIPQLFMRCYICHKPLSWBEKEALLVGNKRFHNI.SGRWTGHCMQCGSTCTAPDPA..SRTLH
HPV11Rct	NYAAYAPTVEEETNEDILKVLIRCYLCHKPLCEIEKLKHILGKARFIKL.NNQWKGRCLHCWTTCMEDLLP.......
HPV72ct	TYSGYGPTVEQETGKSLAELYIRCHACCKPLSCQEKEYQVQTGIHFHKI.SGLWTGRCCQCRGACTARWQP.......
HPV10ct	DYSYYVEGVEEETKQSIYTQLIRCYMCHKPLVREEKDRHRNERRRLHKI.SGYWRGSCEYCWSRCTVRIPQ.......
HPV32ct	DRSAFWHTVEQETGLLLEEQIIRCAICQKPLSPSEKDHHIYNGRHFRFI.LNRWTGRCTQCRE.........
HPV12ct	QQTVLGRDIELATGKSIFDLKIRCQTCLSFLDTIEKLDSCGRGLPFHKV.RDRWKGICRQCKHLYLNNDR........
HPV20ct	ESTVLGRDIBQVTGKSVFDIDVRCYTCMKFLDSIEKLDICGRKRPFYLV.RGSWKGICRLCKHFQ.... ....
HPV17ct	EQSVSGRELEEIBHKPIGEIPIRCKFCLKKLDLLEKLDTCYRHQQFHKV.RRNWKGLCRHCGSIG.... ....
HPV23ct	QLTVYGREIEQEEQRPIGQICIRCQYCLKSLDLIEKLDICSFNQPFHKV.RNHWKGRCRHCKEIE.... ....
HPV1Act	QESYEVPEIBEILDRPLLQIELRCVTCIKKLSVAEKLEVVSNGERVHRV.RNRLKAKCSLCRLYAI...........
HPV48ct	VCTAKSHLLTGLVKKELSDINIRCQHCYSFLDYLEKLYHLYNDVDFLLI.RGTWRGVCRNCISHEGR.......
HPV16Rnt	dPQERPRKLP..QLCTELQTTIHDIILECVYCKQQLLRREVYDFAFRDLCIVYR.DGNPYAVCDKCLKFYSKISEYRHY....
HPV18Rnt	dPTRRPYKLP..DLCTELNTSLQDIEITCVYCKTVLELTEVFEFAPKDLFVVYR.DSIPHAACHKCIDFYSRIRELRHY....
HPV34nt	nPEERPYKLP..ALCEEVNISIHEIELDCVYCERQLYRCEVYDFIPRDLCVVYR.KGKPLGVCQPCLLFYSKVRQYRRY....
HPV26nt	dPRERPRTLH..ELCESLNTTLQNLQVQCVYCKETLQWADVYNFAICDLRVVYR.DRSPYAACKRCVIFYSKITEYRRY....
HPV53nt	nTEERPRTLH..QLCEVVNKPLLELQLGCVFCKKALTASEVYNFAYTDLRVVYR.DGYPYGVCKFCLLFYSKVRKLRYY....
HPV40nt	rCGSQARTLY..ELCDQCNITLPTLQIDCVFCKTVLKTAEVLAFAFRELYVVWR.DDFPHAACPRCLDLHGKVNQYRNF....
HPV27nt	eENPCPRNIF..LLCKQYGLELEDLRLLCVYCRRALSDADVLAFAIKELSVVWR.KGFPFGACGKCLIAAGKLRQYRHW....
HPV11Rnt	dASTSATSID..QLCKTFNLSLHTLQIQCVFCRNALTTAEIYAYAYKNLKVVWR.DNFPFAACACCLELQGKINQYRHF....
HPV72nt	MGLHNPTNIW..LLCKEIEVDLEDLRITCIFCKNELTTEELLAIAIKELQIVWR.DNWPPGVCAPCLARATKVRELRYW....
HPV10nt	MGAQEPRNIL..LLCRNCGIPLEDLRLCCIFCTKQLTAAELAAFALRELYLVWR.AGVPYGACARCLLLQGIVRRLKYW....
HPV32nt	sASSQPSTLY..QLCKDFGLTLRNLQICCIWCKNHLTSAEAYAYHFKDLHVVWK.KGFPYAACAFCLEFYSKVCALRHY....
HPV12nt	sTPELPTTIK..ELADLLDIPLVDCLVPCNFCGKFLDPLEVCDFDKKQLTLIWK.GHFVTACCRSCCAATAIYEFNEFY....
HPV20nt	1EPPLPATIC..GLAKLLEIPLDDCLIPCNFCGNFLTHLEVCEFDEKKLTLIWK.DHLVFACCRVCCSATATYEFNQFY....
HPV17nt	MDRPKPQTVR..ELADTLCIPLVDILLPCRFCNRFLAYIELVAFDLKGLQLIWTEEDFVFACCSSCAYATAQYEFSKFY....
HPV23nt	MDSTRPLTVQ..QLSDKLTVPVVDLLLPCRFCSRFLTYLELREFDYKHLQLIWTEEDFVFACCSGCAYASAQFEIQQFY....
HPV1Ant	.MATPIRTVR..QLSESLCIPYIDVLLPCNFCNYFLSNAEKLLFDHFDLHLVWR.DNLVFGCCQGCARTVSLLEFVLYY....
HPV48nt	MEPQFPTDLD..SYCKYFNISFFDLVLKCIFCKFSVSIVDLASFHNKRLSVIWR.DNTPFACCTKCLRLTALYEKDNFF....

Figure 15.: Sequence alignment of E6 sequences from different HPV types. Alignment of E6C (upper half) and E6N (lower half) domains from 17 E6 sequences representative of the main HPV species. The conserved cysteines are coloured brownish (Nomine et al., 2006, see supplementary data at the end of the manuscript).

et al., 2000). Mutation into serine of the 6 non conserved cysteine residues of wild type E6 which are not involved in Zn^{2+} coordination, referred to as E6 (6C/6S), combined with systematic optimisation of the expression and purification protocol finally allowed to purify monodisperse and biologically active E6 (6C/6S) (Nomine et al., 2001b) (Figure 15). However, until now, it has not been possible to perform NMR experiments on full length E6 (6C/6S), since it is not stable at the required concentrations. Recent results seem to show that purified and monomeric samples of E6 protein tend to aggregate into fibres even at low concentrations after 12-24 h (Zanier et al., unpublished data). However, it has been shown that E6 contains 2 zinc binding domains (Nominé et al., 2003) and the C-terminal domain, named E6-C (4C/4S), revealed to be monomeric and stable upon purification, whereas the E6-N (2C/2S) tended to precipitate. Further optimisation of the E6-C (4C/4S) expression and purification protocol as well as setting up a standard operation protocol for producing isotopically labelled E6-C (4C/4S) for NMR analysis were part of my Ph.D. work and are presented in the results section (Nominé et al., 2003; Nominé et al., 2005). In the following paragraph I will briefly introduce the structure of E6-C (4C/4S) and the derived model of full length E6. Detailed information can be found in the publication which is added in the supplementary data (Nominé et al., 2006).

Most of the hydrogen and nitrogen frequencies of E6-C (4C/4S) were assigned and the solution structure measured and calculated by NMR spectroscopy (Nominé et al., 2006) (Figure 16). The use of the E6-C (4C/4S) mutant for structure determination can be justified by the facts that the mutated cysteins were not conserved in the sequence alignment and that they were replaced by serine residues which represents a very isosteric substitution. Furthermore the mutations did not prevent the domain to adopt a stable fold, all mutations concern exposed residues and the full length E6 (6C/6S) retains full biological activity with respect to p53 degradation *in vitro* and *in vivo* (Nominé et al., 2001b). Translational diffusion experiments at different concentrations as well as measurement of global rotational correlation time indicated clearly a monomeric molecule in solution (Nominé et al., 2005). E6-C (4C/4S) represents a novel zinc-binding fold starting at Tyr 81 and ending at Arg 141, which is consistent with domain boundary definition by mild proteolysis experiments (Lipari et al., 2001; Nominé et al., 2003) as well as with medium range NOE analysis (Nominé et al., 2005). It consists of a three-stranded β-
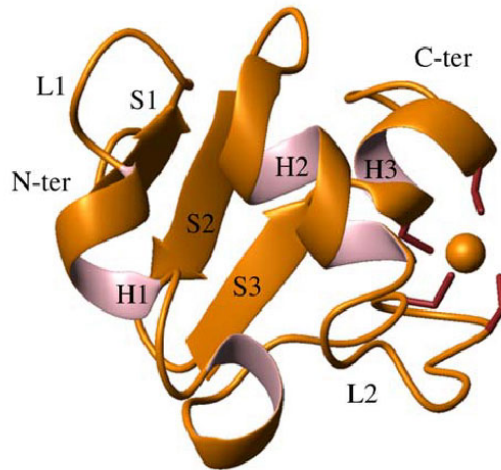
Figure 16.: Ribbon representation of the E6-C NMR structure. E6-C represents a novel zinc binding fold consisting of a three stranded â-sheet (S1, S2, S3) with two short helices (H1, H2) packed on one of its sides. The zinc coordinating cysteine residues are coloured red (figure based on Nominé et al., 2006).
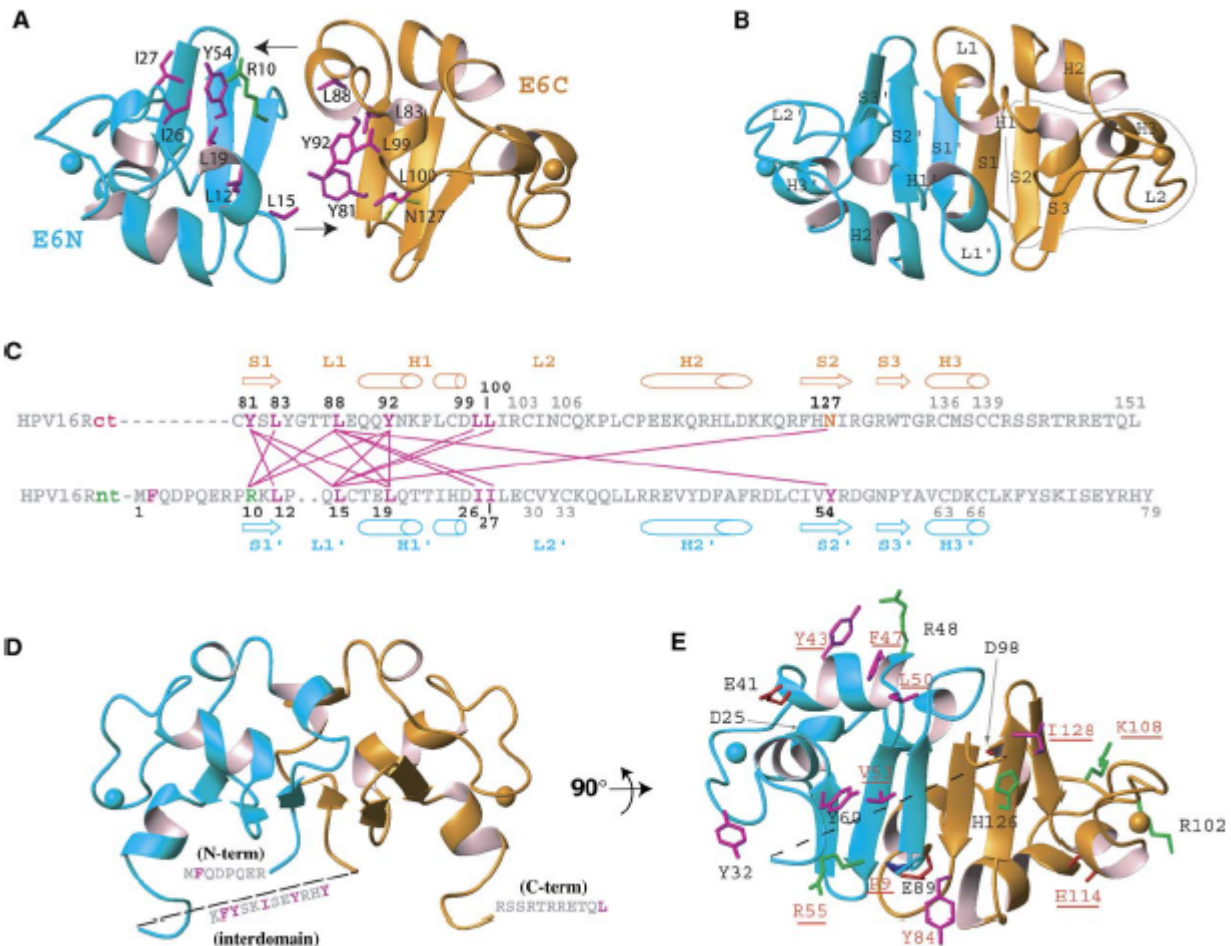
Figure 17.: The pseudodimeric model of E6. A/ E6-N model and E6-C structure positioned symmetrically with exposed hdrophobic patches facing each other. **B/** Proposed pseudodimeric arrangement of E6-N and E6-C. The grey closed line indicates the DNA binding region of E6-C which was mapped in the present work. **C/** Contact map of hydrophobic residues involved in the interface of the pseudodimeric model. **D/** Position of the three flanking regions which are not predicted by the model. **E/** A view of the core model highlighting surface residues conserved in all E6 proteins. Previously mutated residues are indicated with underlined red labels (Nominé et al., 2006).

sheet (S1, S2, S3) with two short helices (H1, H2) packed on one of its sides (Figure 16). The last ten C-terminal residues (S142-L151) are not structured. They encompass a binding site for PDZ[56] domain containing proteins and are supposed to fold upon binding to such a domain (Cowburn, 1997; Hung & Sheng, 2002). PDZ domains are deliberately just mentioned here and will be discussed more in detail in later paragraphs of this manuscript. Cysteines 103, 106, 136 and 139 are implicated in Zn^{2+} coordination (Figure 16). E6-C is a very basic domain with a mostly positive surface charge potential, which correlates with its ability to bind cruciform DNA molecules. The binding site for this kind of interaction could be mapped to the zinc binding region (Nominé et al., 2006). The molecule shows unusual intrinsic molecular motions, which were a limiting factor for the precision of structure calculations. These regions could be mapped to the β1-L1-α1-region and the α2-L3-β2-region which overlaps with a hydrophobic pocket bearing partly exposed and the two totally exposed hydrophic residues Y80 and L88 (Figure 16, 17). NMR titration experiments with peptides conferring the minimal binding sequences of E6AP[57] and p53[58] showed no binding at all. Interaction with these proteins therefore probably requires the full length folded E6 protein. On the other hand, the isolated E6-C domain fully retains binding to PDZ domains via its last C-terminal residues. This will be discussed in greater details in the next chapters. Sequence alignments between E6-C and E6-N showed that all hydrophobic residues in the E6-C domain are well conserved in the E6-N domain, confirming the hypothesis that the two domains may originate from gene duplication (Cole & Danos, 1987). In addition circular dicroism experiments indicated the same α-helix and β-strand content in the two domains. Based on these data a plausible model of the E6-N domain could be modelled by homology to the backbone fold of the E6-C domain. Analysis of the E6-C structure and the E6-N model according to exposed hydrophobic residues and the backbone dynamics of E6-C around the exposed hydrophobic patch, allowed to propose a model of the full length E6 protein, which masks most of the hydrophobic residues allowing optimal fit of the secondary structure elements (Figure 17). In this model of full length E6, the two domains face each and assemble as a symmetrical pseudo-dimer. Note that the

---

[56] PDZ: PSD95/DLG/ZO1; small globular protein-protein interaction domain which was identified first in the three listed proteins

[57] The E6AP peptide used for NMR titrations was : ESSELTLQELLGEER (Elston et al., 1998; Huibregtse et al., 1993)

[58] The p53 peptide used for NMR titrations was: TLEDSSGNLLGR

UBIQUITINATION-DEGRADATION	E6AP
TRANSCRIPTIONAL REGULATION	BRCA1
	P53
	Myc
	CBP/p300
	ADA3
	Tuberin
	TRIP-Br1
	NFX1
	GADD34/PP1
	Zyxin
	Gps2
	PML-IV
DNA REPAIR	XRCC1
	MGMT
DNA REPLICATION	hMCM7
APOPTOSIS	p53
	p21
	Bak
	FADD
	Bax
	TNRF1
EPITHELIAL DIFFERENTIATION & ORGANISATION	Paxillin
	E6BP/ERC-55
CELL SIGNALLING	IRF3
	PKN
	E6TP1
	TIP-1
	TIP-2
NUCLEAR IMPORT	Importins
CELL ADHESION AND POLARITY	MAGI-1,2,3
	hDLG
	hScrib
	MuppI
	Fibulin-1
	PatJ

Table 3. Cellular proteins targeted by HPV 16 E6

conformations of the N-terminal and the C-terminal tails as well as the linker region could not be predicted. This model is supported by HSQC-NMR mapping experiments when measuring the effect of unlabelled E6-N on isotopically labelled E6-C (unpublished data). The data is poor due to the tendency of E6-N to aggregate but seems to indicate implication of the hydrophobic exposed residues Y80-L82. Analysis of conserved exposed residues of the full length E6 models of different HPV evolutionary groups with respect to their level of exposure to the solvent revealed clusters of specialised surface residues. The physico-chemical properties of residues in these clusters vary from one HPV species to another, suggesting that in this way the E6 proteins of different HPV species modulate the ability to bind distinct sets of cellular targets. However, until now generic E6 activities could not be clearly defined, since E6 proteins from most HPV species have not been characterised.

Biological activity of the HPV 16 E6 oncoprotein

In the 1980s, E6 has been identified as a viral oncoprotein, since studies on cervical tumours and derived cell lines showed evidence of steady presence of E6 (Androphy et al., 1987; Banks et al., 1987; Schwarz et al., 1985). E6 has been described to induce immortalisation of primary mammary epithelial cells (Band et al., 1991; Wazer et al., 1995). In addition, expression of high-risk HPV E6 alone has been shown to induce transformation of mammary fibroblasts as well as immortalisation of epithelial cells (Kiyono et al., 1998; Liu et al., 1999; Song et al., 1999).

E6 is a multifunctional protein which main activities are mediated by interacting with more than 33 cellular proteins (table 3). In addition E6 has been described to activate transcription of specific genes and to bind to four-way DNA junctions. The following chapters will describe the best known activities of E6.

Interaction of high-risk HPV E6 with p53 and E6AP and its biological effects

Degradation of P53 and subsequent interference with p53-mediated cell-cycle regulation is certainly the most cited and one of the best-characterised functions of high-risk HPV E6. p53 is a multifunctional tumour suppressor protein which is transcriptionally upregulated and post-translationally modified in response to
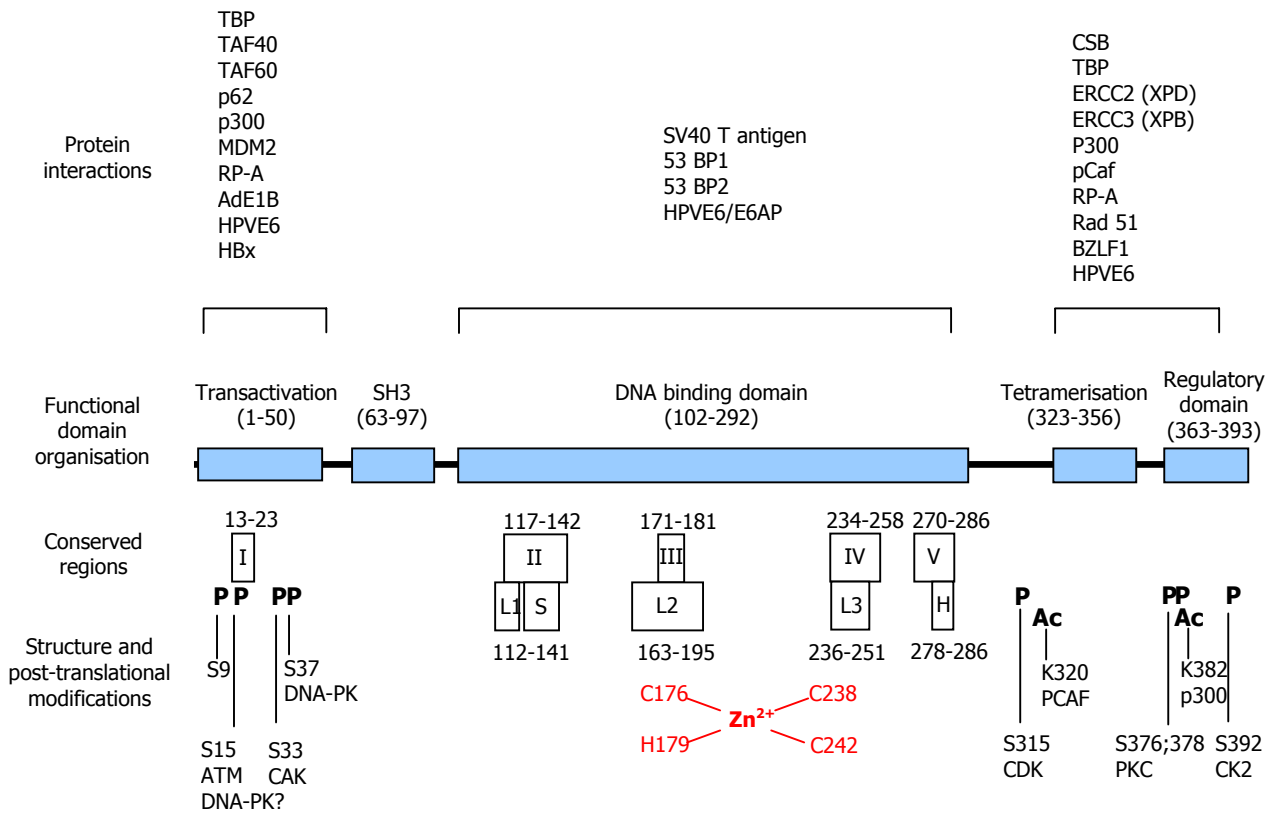
Figure 18.: Schematic representation of the p53 tumour suppressor protein. A non-exhaustive list of viral and cellular proteins are listed according to the functional domains displayed as light-blue squares. The 5 conserved regions are numbered I to V. Some phosphorylation and acetylation sites are shown together with their corresponding kinases. Some secondary structure elements of the core domain are depicted on the bottom of the scheme and labelled L1, L2, L3 for loops, S for â-sheet and H for á-helix. The zinc coordinating atoms are listed. The numbers refer to the amino acid residues (Cho et al., 1994; May & May, 1999).

genotoxic or cytotoxic stress. It is involved in regulation of both G1/S and G2/M cell cycle checkpoints (Oren, 2003). DNA damage or other cellular stress activates p53-mediated transcription and translation of cellular regulators which induce cell-cycle arrest and/or apoptosis (Figure 13). In addition it has been shown to be an important regulator in other cellular processes such as genome stability, differentiation, senescence or angiogenesis. P53 downregulates the expression from several promoters, among which are *c-jun*, *c-myc* and *β-actin*. The fact that p53 is inactivated in nearly 50% of all cancers (Cox & Lane, 1995; Hollstein et al., 1996) emphasises the importance of this protein as a tumour suppressor.

Human p53 is a protein of 393 residues and 53 kDa, located in the nucleus. Five highly conserved regions could be identified by cross-species comparison spanning the amino acids 13-23, 117-142, 171-181, 234-258 and 270-286, termed domains I-V (Soussi et al., 1990; Soussi & May, 1996). In human p53, four major functional domains were described: a N-terminal transcriptional activation domain (residues 1-42), a central sequence specific DNA binding domain, also known as core domain (residues 102-292), a C-terminal tetramerisation domain (residues 323-356) and a C-terminal regulatory domain (residues 360-393) (May & May, 1999) (Figure 18).

The acidic N-terminal transcriptional activation domain (residues 1-42) (Unger et al. 1992) allows p53 to recruit several cellular proteins, among which the basal transcriptional machinery, including the TBP and TAFs[59] (Lu & Levine, 1995; Thut et al., 1995). Adenovirus E1B-55 kDa protein, the hepatitis B virus X protein as well as the human MDM2[60] protein bind to the N-terminal region of p53 and inhibit its transactivation function (Levine, 1997; Momand et al., 1992; Oliner et al., 1993; Yew & Berk, 1992;). MDM2 is a cellular negative regulator of p53. By binding the N-terminal domain of p53, it inhibits the transactivation activity of p53 (Oliner et al., 1993) and provokes its polyubiquitination, leading to its proteasome-dependent degradation (Haupt et al., 1997; Kubbutat et al., 1998). The *mdm2* gene itself is activated by p53 in a negative feedback control loop (Wu et al., 1993). The N-terminal region of p53 also contains a proline rich region (residues 63-97) with

---

[59] TAF: TBP-Associated Factors
[60] MDM2: Murine Double Minute 2
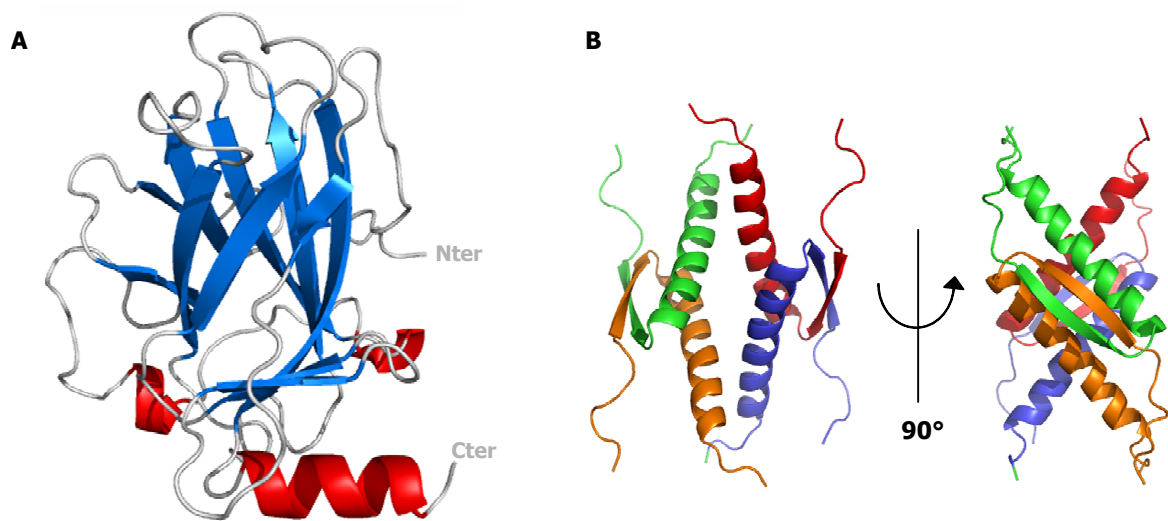
Figure 19.: Structures of the core domain and the tetramerisation domain of p53. A/ X-ray structure of the core domain of p53. It consists of a four- and a five-stranded antiparallel â-sheet and a loop-sheet-helix motif, which comprises a loop, a three stranded â-sheet and the á-helix. A second loop which is interrupted by a short helix and a third loop are stabilised by a zinc ion (Cho et al., 1994). **B/** NMR structure of the tetramerisation domain of p53. The structure shows the arrangement of the monomers in the dimer dimer conformation (Clore et al., 1994).

similarity to SH3 domains, which is required for p53 mediated apoptosis (Sakamuro et al., 1997), for tumour cell growth suppression (Walker & Levine, 1996) and for high risk HPV E6 mediated degradation (Li & Coffino, 1996). Interestingly the afore mentioned p53 polymorphism encoding either P or A at position 72 is situated in this proline rich region (Matlashewski et al., 1987; Rosenthal et al., 1998; Storey et al., 1998;).

The central region or core domain of p53 (residues 102-292) is an autonomous folding domain containing four conserved sub-regions (II – V). This region is resistant to proteolysis (Bargonetti et al., 1993; Pavletich et al., 1993) and contains the DNA binding domain, which recognises and binds a consensus target sequence (El-Deiry et al., 1992). 80-90% of the tumour mutations are found within this region. In addition this region functions as a protein binding domain interacting with the simian virus 40 (SV40) large T antigen as well as with the cellular proteins 53BP1 and 53BP2 (Gorina & Pavletich, 1996; Iwabuchi et al., 1994; Jenkins et al., 1988; Ruppert & Stillman, 1993). The X-ray structure of the core domain was solved in 1994 (Cho et al., 1994) (Figure 19A).

The C-terminal region of p53 comprises a flexible linker (residues 300-318) connecting the core domain to the tetramerisation domain (residues 323-356) and the basic regulatory domain (residues 363-393). The structure of the tetramerisation domain was solved both by NMR and crystallography and revealed formation of stable tetramers by an assembly of "dimer dimers" (Clore et al., 1994; Jeffrey et al., 1995) (Figure 19B). It is well established that p53 forms tetramers and that tetramerisation is essential for efficient transactivation *in vivo* and for p53-mediated suppression of growth of carcinoma cell lines (Kraiss et al., 1988; Pietenpol et al., 1994). Three NLS[61] have been identified in the C-terminal region. Mutation of NLS1 (residues 316-325) leads to complete cytoplasmic localisation of p53, whereas mutation of NLS2 and NLS3 (residues 369-375 and residues 379-384, respectively) leads to a both nuclear and cytoplasmic localisation (Dang & Lee, 1989; Shaulsky et al., 1990). The basic regulatory domain has been described to be possibly implicated directly in induction of apoptosis by interacting with helicases of the basal transcription repair complex TFIIH (Wang et al., 1996), a transcriptional regulatory domain (Wang & Prives, 1995) or a DNA damage recognition domain. This region

---

[61] NLS: Nuclear Localisation Signal

E6-dependent ubiquitination

P53-binding

1   289   527   790   841 875

IPESSELTLQELLGEERR
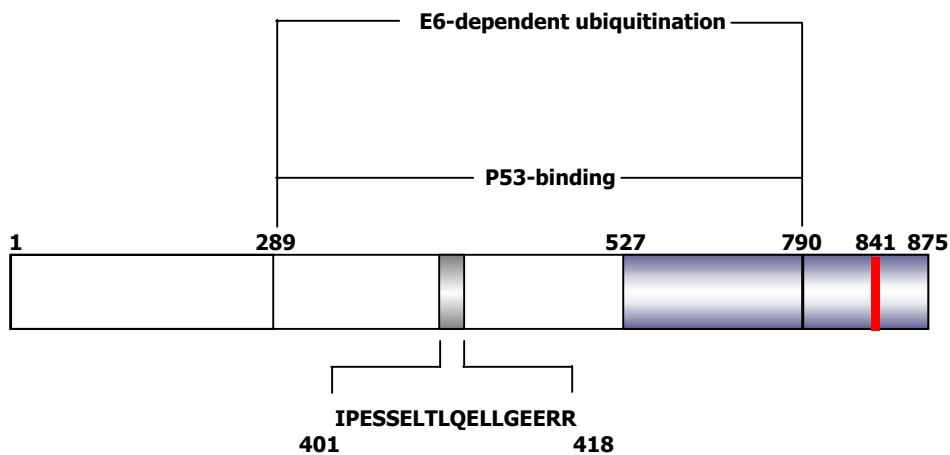401   418

Figure 20.: Schematic representation of E6AP. The E6 binding charged leucine motif is coloured in light grey. It was the first protein of a familiy of proteins containing an LXXLLX motif implicated in E6-binding. The regions having been described as being implicated in E6-dependent p53 binding and polyubiquitination are labelled. The catalytical HECT domain is coloured blueish. The structure of the HECT domain bound to the UbcH7 ubiquitin-conjugating enzyme has been published (Huang et al., 1999), whereas the structure of the residues 1-527 remains unknown. The catalytical cysteine residue at position 841 is labelled with a red line. A substitution of this cysteine residue by an alanine disables the ubiquitin-ligase activity. The amino acid numbering corresponds to the isoform II of E6AP (Yamamoto et al., 1997). Figure adapted from Huibregtse & Beaudenon, 1996.

E6
E6AP

P53
E6AP

E2

E2

E1

P53

Proteasome mediated degradation

Ubiquitin-ATP

E1

AMP + PPi
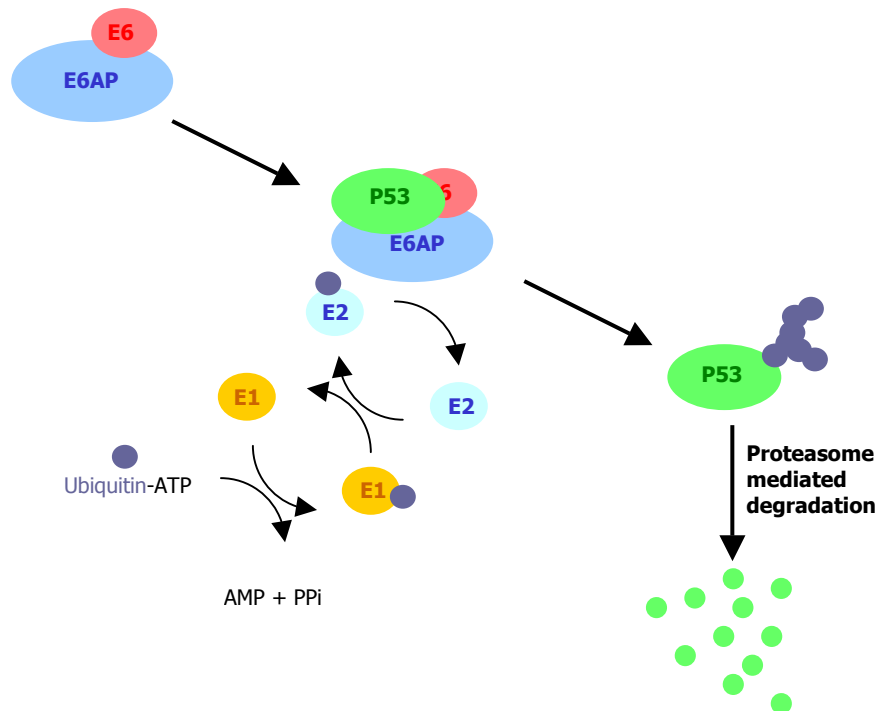
Figure 21.: Schematic representation of the E6AP dependent E6 mediated degradation of p53. P53 is selectively targeted for degradation by covalently linkage to ubiquitin in a three-step E6AP dependent pathway. A ubiquitin-activating enzyme (E1) binds ubiquitin in an ATP-requiring reaction via a thioester bond. Ubiquitin is transferred on a sulfhydryl group of a ubiquitin–conjugating enzyme (E2). The ubiquitin is then transferred via a ubiquitin protein ligase (E3 or in this special case E6AP) on the target protein. Repetition of this process leads to polyubiquitination of the target protein (p53) mediating its proteasomal degradation.

seems to act as a negative regulator of p53 sequence specific binding (Abarzua et al., 1996; Hupp & Lane, 1995; Selinova et al., 1997; Selinova et al., 1999; Shaw et al., 1996).

In addition activation of p53 *in vivo* has been correlated with phosphorylation (Ko & Prives, 1996), acetylation (Sakaguchi et al., 1998), glycosylation (Shaw et al., 1996) as well as with proteolytic removal of the C-terminal domain (Okorokov et al., 1997). Activation by phosphorylation or acetylation stabilises p53 and induces tetramerisation. In general, most post-translational modifications on p53 (apart from polyubiquitination) are thought to induce conformational changes in p53 thus activating specific functions of the protein.

High-risk HPV E6 has recently been described to repress p53-dependent gene activation. High and low risk E6 proteins show similar activity in repressing p53-dependent chromatin transcription via modulation of acetylation on p53 and nucleosomal core histones. Instead of disrupting pathways leading to the assembly of active transcription complexes on chromatin, HPV E6 seems to alter the function of pre-existing transcription complexes to repressors of transcription. This may point to a general strategy for oncoviruses: viral oncoproteins would incorporate into existing complexes with minimum disturbance to allow cellular differentiation, which is essential for viral replication. In case of HPV, viral particles are finally released by cellular desquamation (Thomas & Chiang, 2005).

High-risk HPV E6 has been described to achieve p53 degradation via formation of a stable complex with another cellular protein called E6AP[62] (Scheffner et al., 1993) (Figure 20, 21). E6AP is a protein of 875 amino acids and 100 kDa which was found initially associated to p53 (Huibregtse et al., 1991; Yamamoto et al., 1997). E6AP is a E3 ubiquitin-ligase (Huibregtse et al., 1995) which has been localised both in the cytoplasm and in the perinuclear region of the cells where p53 degradation occurs (Daniels et al., 1998). In addition, E6AP has been described as a transcriptional coactivator interacting with nuclear receptors and it has been implicated in the Angelman Syndrome (Nawaz et al., 1999). E3 ubiquitin-ligases, by definition, are enzymes which bind to specific protein substrates and promote the

---

[62] E6AP: E6 Associated Protein

transfer of ubiquitin from a thioester intermediate to amide linkage with the substrate proteins. This way they mono- or polyubiquitinate them, the latter being a signal for proteasomal protein degradation. E3 ubiquitin-ligases can be divided into 3 major classes, based upon domain structure and substrate specificity. The first class are the N-end rule ubiquitine-ligases which target protein substrates bearing specific destabilising residues in their N-terminus (Varshavsky, 2003). The second type are the HECT[63] domain proteases. E6AP was detected and described as the first member of this class of ubiquitine-ligases (Figure 20). The last and most represented class are the RING[64] ubiquitine-ligases which bear a metal chelating RING finger domain. E6AP alone is able to bind E6 via a charged leucine consensus motif (LXXLL(A/G)X) (Huibregtse et al., 1993) but not p53 (Figure 20). Only in the presence of E6 does it form a trimeric complex with p53 which results in specific poly-ubiquitination and subsequent proteasome-mediated degradation of p53, thus abrogating p53-mediated cell cycle control and allowing viral replication (Cooper et al., 2003). The full-length structure of E6AP has not been determined yet but it has been possible to map the E6 binding site which is located between the residues 391-408 (Huibregtse et al., 1993) and to acquire and to assign by NMR the ^{1}H, ^{15}N and ^{13}C frequencies of the HECT domain (Wu & Domaille, 2001; BMRB acc.:5013). In addition, this 18 amino acid peptide could be characterised structurally. The peptide exhibits α-helical tendencies in solution (Be et al. 2001). On a helical model of the peptide, the L residues of the LXXLL consensus motif are exposed at the same side of the helix with a sterical requirement of a G or an A residue at the position after the LXXLL motif (Be et al., 2001; Liu et al., 2004). Importance of these residues could be demonstrated by mutational analysis in the same study (Be et al., 2001), as well as independently confirmed (Bohl et al., 2000; Liu et al., 2004). It has been described that both low-risk and high-risk HPV E6 oncoproteins could bind to the C-terminal basic region of p53 (Li & Coffino, 1996). However, formation of the trimeric high-risk HPV E6/E6AP/p53 complex promotes binding of E6 to the core domain of p53 which has been demonstrated to be required to induce proteasome mediated degradation of p53 (Li & Coffino, 1996).

---

[63] HECT: Homology to E6AP C-Terminus
[64] RING: Really Interesting New Gene

Binding to E6BP

E6BP[65], also described as ERC-55, is another member of the charged leucine motif binding proteins. It is a calcium binding protein of the CREC[66] family of proteins which is localised in the endoplasmic reticulum (Honore et Vorum, 2000; Weis et al., 1994). In vitro experiments demonstrated that E6BP interacts specifically with high–risk HPV (Rapp & Chen 1998). Through interaction with E6BP, E6 may inhibit terminal differentiation of epithelial cells, establishing the necessary environment for the viral DNA to replicate (Sherman et al., 2002). The minimal binding domain of E6BP to E6 has been mapped via mutagenesis experiments to a LEEFLGD peptide (Chen et al., 1998) which was independently confirmed by yeast two-hybrid experiments (Elston et al., 1998).

Binding to MCM7

Multicopy maintenance protein 7 (MCM7) is also a member of the charged leucine motif binding proteins (Kuhne & Banks, 1998). It is also referred as minichromosome maintenance 7 protein (Kukimoto et al., 1998) and is a protein which is essential for DNA replication and regulates that DNA replication occurs only once in a cell cycle (Kukimoto et al., 1998). Both E6 and E6AP seem to interact with MCM7 via a LLG motif leading to its degradation (Kuhne & Banks, 1998).

Binding to Paxillin

Paxillin also belongs to the family of charged leucine motif binding proteins. It is implicated in the formation of focal adhesion protein complex, which is located at the extracellular matrix and plays an important role in cellular morphology, division and metastasis. E6 interacts with paxillin and prevents association with its regular cellular partners (Tong et al., 1997). E6 binds via the LDXLLXL consensus motif, which is present four times in the paxillin sequence (Vande Pol et al., 1998), but the most N-terminal one seems the most important concerning E6 binding (Tong et al., 1997).

---

[65] E6BP: E6 Binding Protein

[66] The CREC protein family is a recently discovered protein family with similar sequences but unknown functions. It consists of a number of low-affinity calcium binding proteins containing multiple EF-hand Ca2+ binding motifs. These proteins have been shown to interact with a number of other proteins, and have been associated with a number of disease states, including cancer, malaria, and toxin mediation.

Binding to TNF R1

TNF[67] is again a member of the charged leucine motif binding proteins. It is a pro-apoptotic signalling protein which induces intracellular apoptotic signals through its cell surface receptor, TNF R1. Apoptosis is a genetically programmed cell death and is integral component of the normal cell maintenance. It also serves to eliminate radiation damaged, aberrantly growing as well as virally infected cells which will threaten integrity of the entire organism. In case of virally infected cells, the early cell death would not allow the virus to replicate. E6 binds to the C-terminus of TNF R1 which presents two (E/D)L(L/V)G motifs (Elston et al., 1998; Filippova et al., 2002). E6 binding interferes with the TNF R1-associated death domain which normally binds to TNF R1 and is necessary for apoptotic signalling.

Binding to Tyk and IRF-3

IRF-3[68] is a transcription factor which regulates expression of the interferon α and β-controlled genes. It belongs to a group of structurally related transcription factors, which mediate cellular responses to interferons. High-risk HPV E6 binds to IRF-3 and thus disrupts cellular anti-viral responses (Ronco et al., 1998). ISGF3α, another protein which is homologous to IFR-3 is activated via phosphorylation by the tyrosine kinase TyK2. High-risk HPV E6 proteins interact with TyK2 and block their activity thus deregulating other interferon signalling cascades (Li et al., 1999).

Binding to Bak

Bak is a pro-apoptotic member of the Bcl-2 family. High-risk and low-risk HPV E6 bind directly to Bak. However poly-ubiquitination mediated proteasomal degradation occurs via a trimeric E6/E6AP/Bak complex (Thomas & Banks, 1999). In normal cells Bak turnover is regulated by E6AP-mediated poly-ubiquitination (Jackson et al., 2000).

Binding to c-Myc and expression of hTERT

In normal somatic cells the number of cell divisions is partly limited by the shortening of the chromosomal telomeres. Telomerases are enzymes which elongate

---

[67] TNF: Tumour Necrosis Factor
[68] IRF3: Interferon Regulatory Factor-3
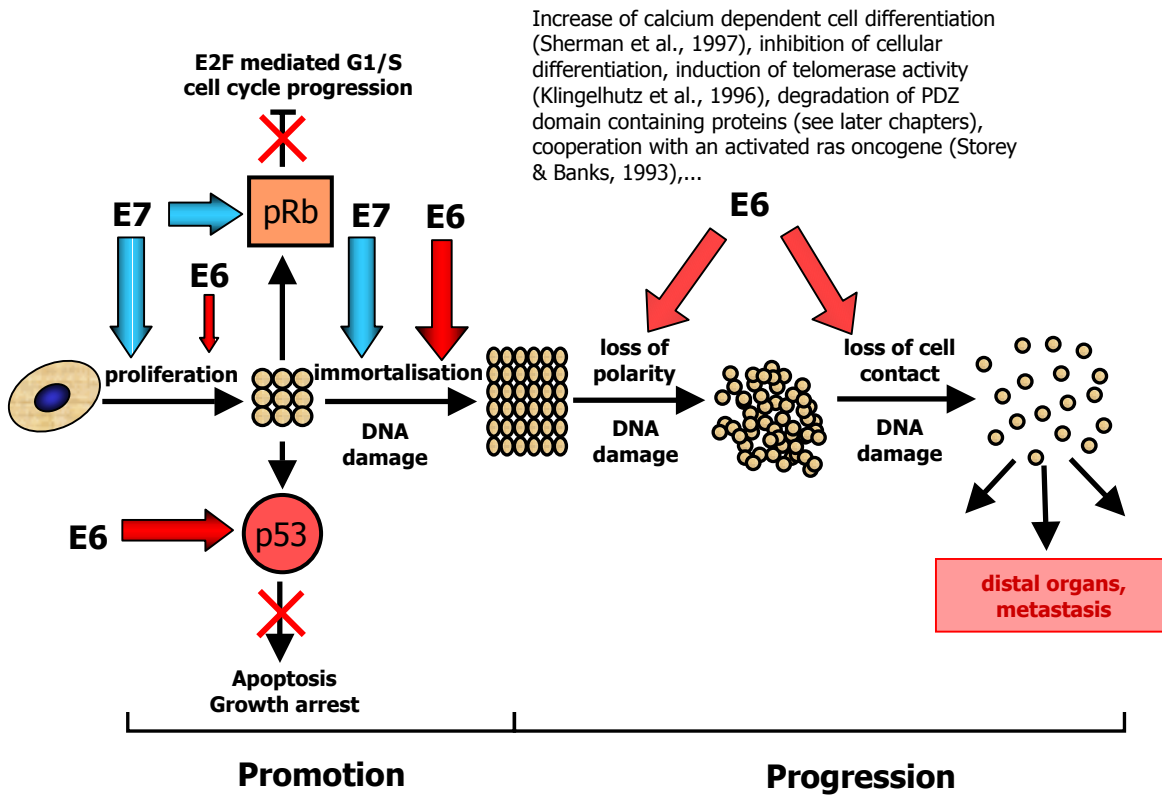
Increase of calcium dependent cell differentiation (Sherman et al., 1997), inhibition of cellular differentiation, induction of telomerase activity (Klingelhutz et al., 1996), degradation of PDZ domain containing proteins (see later chapters), cooperation with an activated ras oncogene (Storey & Banks, 1993),...

Figure 22.: Schematic representation of the different contributions of E6 and E7 to transformation and cancer (Mantovani & Banks, 2001).

telomeres by adding specific DNA repeats. Immortalised cells maintain their telomere length by having constantly activated telomerase. A major player is the catalytic protein subunit of human telomerase, the reverse transcriptase (hTERT). By interacting with c-Myc, E6 activates the hTERT gene expression (Veldman et al., 2003). However there are conflicting reports which cannot find correlation between c-Myc expression and hTERT induction by E6 (Gewin et al., 2004). Recent experiments show evidence that high-risk HPV E6 induction of hTERT requires the presence of E6AP. It has been shown that E6 recruits E6AP to the telomerase promoter (Kelley et al., 2005). In addition, two splice variants of NFX1, a known transcriptional repressor were identified as being targets of the E6/E6AP complex for proteasomal degradation. NFX1-123, the first splice variant, strongly coactivates with c-Myc at the hTERT promoter, whereas NFX1-91, the second splice variant, represses the activity of this promoter. It has been shown that coexpression of E6 with either isoform, increased the activity of the promoter, suggesting that E6 could co-operatively activate hTERT with NFX1-123 and could decrease the repressive effects of NFX1-91 (Gewin et al., 2004).

Binding to PDZ domain containing proteins

In the late 1990s another important feature of high-risk HPV E6 oncoproteins was discovered. High-risk mucosal HPV E6 have the ability to interact with PDZ[69] domains of distinct cellular proteins via a conserved X[T/S]X[L/V] consensus binding motif located at their C-terminus. This C-terminal domain has been described as being not involved in p53 binding and degradation. However, deletion mutants of high risk E6 lacking this binding motif were impaired in transforming cells (Kiyono et al., 1997). This correlates with experiments showing that a functional PDZ binding motif correlates partly with EMT[70] and disruption of intercellular junction formation in immortalised keratinocytes (Watson et al., 2003). In addition, experiments in transgenic mice showed that E6 can induce hyperplasia and tumorigenesis by pathways independent of p53 degradation (Song et al., 1999) and that a functional

---

[69] PDZ: PSD95/SAP90 - Dlg - ZO1

[70] EMT: Epithelial-to-Mesenchymal Transition. It is a program in cellular development of biological cells characterised by loss of cell adhesion, repression of E-cadherin expression, and increased cell mobility. EMT is essential for numerous developmental processes including mesoderm formation and neural tube formation. Initiation of metastasis involves invasion, which has many phenotypic similarities to EMT, including a loss of cell-cell adhesion mediated by E-cadherin repression and an increase in cell mobility.
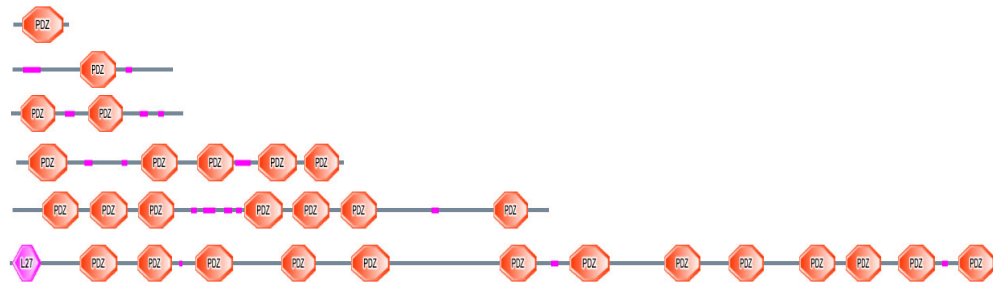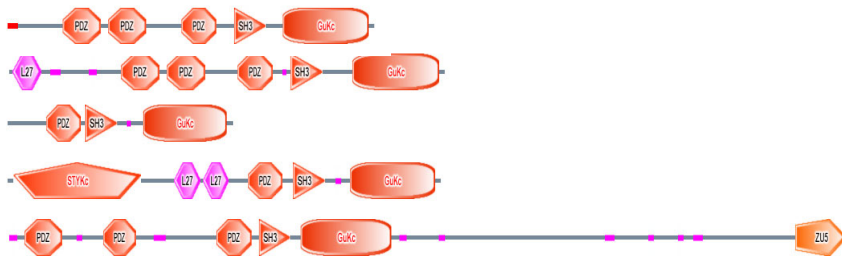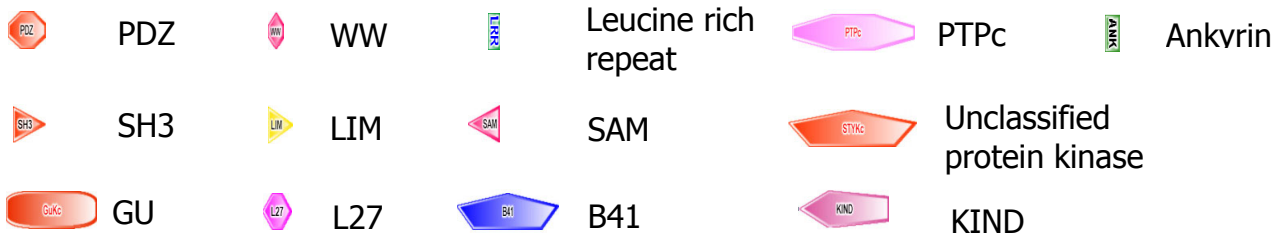
Figure 23.: Modular organisation of PDZ-containing proteins exemplified by representative members of described PDZ families. Figure adapted from Jelen et al., 2003; Ham & Hendriks, 2003. Protein domain organisation was generated with SMART (Schultz et al., 2000).

C-terminal PDZ binding motif is necessary for induction of hyperplasia and suprabasal DNA synthesis and might be implicated into malignant progression (Nguyen et al., 2003; Simonson et al., 2005). Figure 22 summarises the different contributions of E6 and E6 to immortalisation and malignant transformation before going into detail of E6s interaction with PDZ domains.
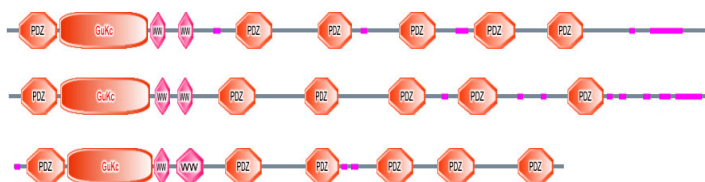
PDZ domains

PDZ domains - a general introduction

PDZ domains were first identified as regions of sequence homology found in diverse signalling proteins (Cho et al., 1992; Woods & Bryant, 1993; Kim et al., 1995). PDZ domains are the most common protein interaction modules representing 0.2 to 0.5% of open reading frames in three ore more currently sequenced metazoan genomes (Jelen et al., 2003; Schultz et al, 1998; Schultz et al., 2000). Originally PDZ domains were recognised as ~90 amino acid sequences in the synaptic protein PSD95/SAP90, the drosophila septate junction protein DLG[71] as well as the epithelial tight junction protein ZO-1[72], hence the acronym PDZ (Kennedy, 1995). PDZ domains have also been described in the past as DHRs[73] or as GLGF repeats[74]. PDZs are small globular domains which are specialised for binding to the C-terminus of proteins which are mainly localised at cell-cell junctions such as transmembrane receptors and channel proteins, or other PDZ domains. The key function of PDZ domains was found by discovering the interaction of the PDZ domains 1 and 2 of PSD 95[75] with a C-terminal consensus motif E(T/S)(D/E)V of Shaker-type K^+ channels (Kim et al., 1995) and NMDA[76] receptor NR2 subunit (Kornau et al., 1995; Niethammer et al., 1996). At the same time a PDZ domain of PTP1E/FAP-1[77] was described to interact with the C-terminus of the cell surface receptor Fas[78] (Sato et al., 1995). These were the first discoveries which allowed to define the function of PDZ domains as protein-protein

---

[71] DLG: Disc Large protein from Drosophila melanogaster

[72] Zo-1: Zonula occludens 1 protein; It is involved in maintenance of epithelial polarity

[73] DHR: Discs-large Homology Region

[74] GLGF repeat: denoted according to a GLGF signature sequence found in most PDZ domains

[75] PSD 95: Postsynaptic Density protein 95; It is a 95 kDa protein involved in signalling at the post-synaptic density

[76] NMDA: N-methyl-D-aspartate

[77] PTP1E: Protein-Tyrosine Phosphatase 1E; FAP-1: Fas-Associated Phosphatase 1

[78] Fas: Cell surface receptor which is expressed on a variety of normal and neoplastic cells and which is also known as APO1 or CD95

interaction domains. They interact through a sequence-specific manner to C-terminal peptides. This kind of interaction often localises proteins to specific subcellular regions, enabling the formation of local supramolecular complexes. These modular protein domains can easily be put together in order to form multi-PDZ proteins or combined with other modular protein interaction domains in order to generate more complex scaffolds.

Indeed, one striking aspect of PDZ domains is the multiplicity of PDZ domains within a single polypeptide chain. 18% of the human PDZ domain containing proteins have three or more PDZ domains. That is much higher than for other modular signalling domains (Harris & Lim, 2001). PDZ-domain containing proteins can be split into groups according to the arrangement of the PDZ domains as well as other conserved domains. A first group comprises proteins containing only PDZ domains with different specificities within a single polypeptide chain and is referred to as multi-PDZ proteins or PDZ-only proteins (Figure 23). Examples for PDZ-only proteins include single PDZ domain proteins such as PICK1[79] or Par6 as well as multi PDZ domain proteins such as NHERF[80] (2 PDZ), CIPP[81] (4 PDZ), INAD[82] (5 PDZ), GRIP[83] (7 PDZ), PATJ[84] (10 PDZ) or even MUPP1[85] with 13 PDZ domains. The other groups include proteins possessing single or multiple PDZ domains in combination with other functional domains. One group is represented by the MAGUK[86] proteins (Figure 23). They contain invariably one or three PDZ domains, a SH3[87] domain and a GuK[88] domain. A variation of the MAGUK group is represented by the MAGI[89] proteins, which contain mainly the same functional domains but in an inverted order. In addition the SH3 domain is usually replaced by WW[90] domains (Dobrosotskaya et al., 1997). Other PDZ-containing proteins present more diversified combinations with a

---

[79] PICK1: Protein Interacting with C-Kinase 1

[80] NHERF: Na^+/H^+ Exchanger Regulatory Factor

[81] CIPP: Channel Interacting PDZ domain Protein

[82] INAD: Inactivation No Afterpotential D

[83] GRIP: Glutamate Receptor-Interacting Protein

[84] PATJ: Pals-1 Associated Tight Junction protein

[85] MUPP1: Multi PDZ domain Protein 1

[86] MAGUK: Membrane-Associated Guanylate Kinase

[87] SH3: Src Homology 3 domain; They are small protein domains which bind to proline-rich peptides in their respective binding partner.

[88] GuK: Guanylate Kinase homology domain.

[89] MAGI: Membrane Associated Guanylate kinase with Inverted domains

[90] WW domain: Protein domain which is named after a pair of conserved tryptophanes and which binds to proline-rich peptides. They are highly compact (35-45 residues) modular domains which adopt an antiparallel three stranded fold (Macias et al., 1996; Sudol et al., 1996; Zarrinpar & Lim, 2000).

variety of interaction domains such as WW, LIM[91], CaMK[92], DH/PH[93], ankyrin or leucine-rich repeats. These proteins were either gathered into one group referred to as PDZ domain containing proteins with other signalling domains (Harris & Lim, 2001) or splitted into several groups according to the nature of the additional signalling domains, namely the PDZ-LIM group, a LAP[94] group and a Shank group (Figure 23) (Jelen et al., 2003).

Besides the interactions via a C-terminal consensus motif, binding to internal motifs which are presented as a ß-hairpin structure or to other PDZ domains have been described (Hillier et al., 1999; Tochio et al., 2000).

A high amount of PDZ-containing proteins were associated with the plasma membrane (Fanning & Anderson, 1999). Since they have been initially described, PDZ and PDZ-like domains have been identified in numerous proteins from diverse organisms such as bacteria, plants, yeast, metazoans and drosophila (Hung & Sheng, 2002; Ponting, 1997). Analysis of human genomes allowed to estimate the presence of more than 440 PDZ domains in more than 250 different proteins (Schultz et al., 2000; Sheng & Sala, 2001). Since PDZ domains mainly recognise short C-terminal peptides they are able to interact with the great majority of transmembrane proteins if they present a PDZ consensus binding motif on their free C-termini exposed to the cytosol. In addition the combination of PDZ domains with different binding specificities in one scaffold protein will determine the composition of the protein complex assembled around the scaffold. The ability of PDZ-containing proteins to multimerise might increase the size and potentially the heterogeneity of PDZ-based complexes.

One functional complex was described in 1995, showing that PSD95/SAP90 clustered Shaker subfamily K+ channels at presynaptic membranes (Kim et al., 1995).

Another PDZ domain containing protein is GRIP, which links AMPA[95] receptor subunits to down-stream effector proteins and plays a critical role in clustering these AMPA receptors at excitatory synapses (Dong et al., 1997).

---

[91] LIM: Protein domain named from the Lin-11, Isl-1 and Mec-3 genes. It is a protein-protein interaction domain presenting a zinc finger structure.
[92] CaMK: Calmodulin Kinase domain
[93] DH/PH: Dbl Homology/Pleckstrin Homology domain
[94] LAP: Leucine rich repeat And PDZ
[95] AMPA: α-Amino-5-hydroxy-3-Methyl-4-isoxazole Propionic Acid
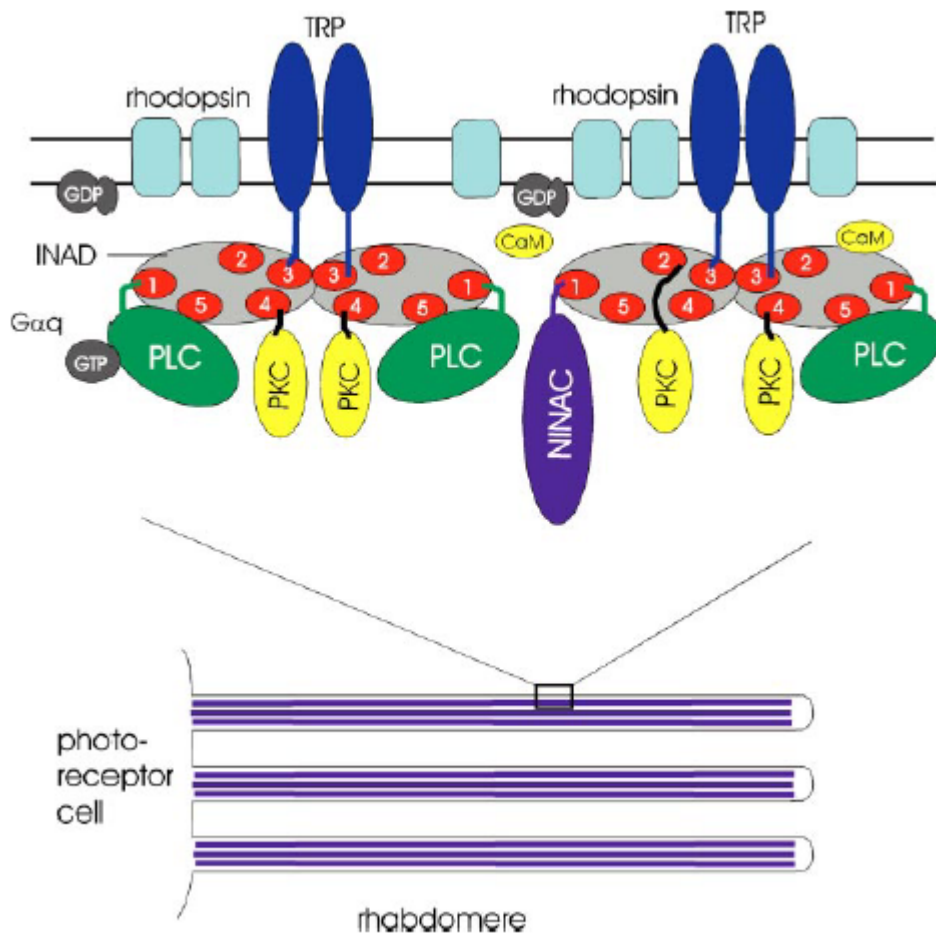
Figure 24.: Schematic representation of the INAD mediated phototransduction complex in *D. melanogaster* photoreceptors. A schematic actin-filled microvilli of rhabdomeres is depicted at the bottom. INAD is shown beneath the plasma membrane multimerised via its PDZ3 and PDZ4. The major interacting proteins are shown binding to the specific PDZ domains. PDZ domains are represented by red ovals, F-actin by purple lines (Sheng & Sala 2001).

In fact PDZ domain containing proteins are mainly scaffolding molecules which cluster proteins at specific cellular localisations in order to form functional complexes. The following paragraphs will give a short overview concerning the best characterised functions:

Organisation of signal transduction in the fly eye:

The protein INAD from *D. melanogaster* is composed of 5 PDZ domains and has been described as scaffold protein for the G-protein-coupled phototransduction cascade in the fly eye. *InaD* mutant flies show in fact mislocalisation of several proteins in the phototransduction signalling cascade in the rhabdomeres of their eye (Chevesich et al., 1997; Shieh & Zhu, 1996; Tsunoda et al., 1997; Tsunoda et al., 2001). INAD seems therefore to be required for the assembly and function of the phototransduction signalling pathway. The assembling function of INAD correlates with the fact that this protein essentially consists of PDZ domains and appears to have no catalytic function. Extensive studies revealed that the five PDZ domains in INAD interact with individual proteins of the signal transduction pathway such as PLC-β[96], PKC[97] or TRP[98] (Figure 24) (Bahner et al., 2000; Harteneck, 2003; Kiselyov et al., 2005; Li & Montell, 2000; Tsunoda et al., 2001).

Signal transduction at synapses:

Several PDZ-domain containing proteins are implicated in organising the signalling machinery at synapses and neuromuscular junctions. At the dendritic side of the neuronal synapse there is a dense complex of proteins, termed PSD[99], which contains different signalling components. Among the scaffold proteins implicated in the PSD is the MAGUK protein PSD95 (Figure 23) which is one of the best-studied in the PSD (Kim et al., 1996; Kistner et al., 1993; Kornau et al., 1995; Muller et al., 1996; Niethammer et al., 1996). The PDZ1 and PDZ2 have been described to bind the C-terminus of the NR2B subunit of the NMDA glutamate receptor (Figure 25) (Kornau et al., 1995; Niethammer et al., 1996). The PDZ2 seems also to participate

---

[96] PLC-β: Phospholipase C-β
[97] PKC: Protein Kinase C
[98] TRP: a cation channel found in the photoreceptors of *D. melanogaster*.
[99] PSD: Postsynaptic Density
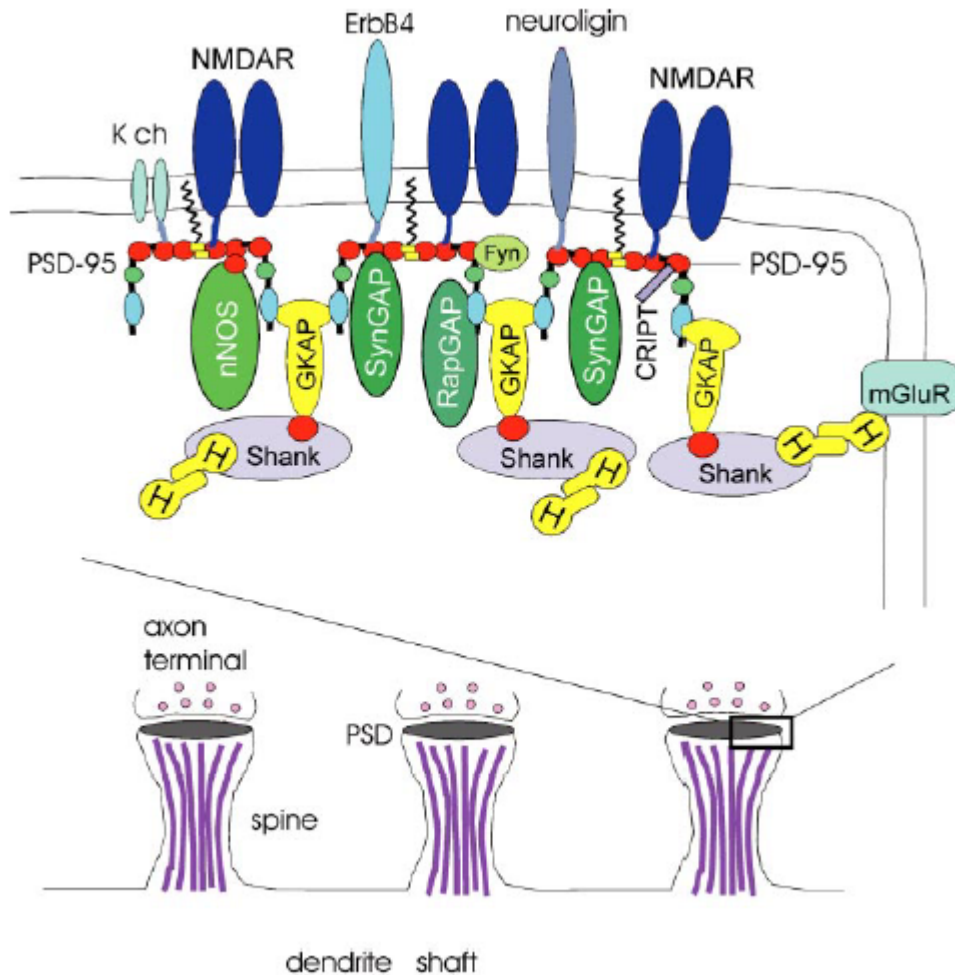
Figure 25.: Schematic representation of the organisation of the PSD at excitatory synapses. The PSD is normally located at the dendritic spines and is directly aligned with the active zone of the presynaptic terminal (bottom). PSD95 is shown beneath the postsynaptic membrane multimerised via ist N-terminal region. The figure shows an incomplete list of proteins binding to PSD95. PDZ domains are represented ba red ovals, F-actin by purple lines (Sheng & Sala 2001).

in a heterodimerisation with the PDZ from nNOS[100] (Brenman et al., 1996). The GuK domain participates in interactions with other PSD components such as GKAP[101] (Naisbitt et al., 1999). The PDZ-mediated co-localisation of nNOS with the NMDA receptor complex is thought to allow efficient synthesis of the second messenger nitric oxide in response to local Ca^{2+} influxes resulting from NMDA channel opening (Figure 25). In addition PSD95 has been implicated in synapse formation, receptor clustering, AMPA receptor clustering and trafficking by interacting with SAP97[102], regulating dendrite outgrowth and branching (Cai et al., 2006; Charych et al., 2006; El-Husseini et al., 2000). However, its precise biological role is still not clear, since PSD95 knockout mice do not show severe neurological or developmental defects. This might be because of possible overlapping functions of the different MAGUK proteins present in neurons (Migaud et al., 1998). Another protein implicated in the organisation of the PSD is GRIP, a scaffold protein which is able to dimerise and to interact with AMPA receptors (Dong et al., 1997; Srivastava et al., 1998; Wyszynski et al., 1999) as well as SHANK which has been described to crosslink PSD95 and GRIP and which has been proposed to be promoting scaffolds of scaffolds (Kennedy, 2000; Naisbitt et al., 1999; Sheng & Kim, 2000; Tu et al., 1999).

Organisation of membrane protein activity and trafficking:

Another activity of multiple PDZ containing proteins is the regulation of membrane protein activity and trafficking. NHERF is one example of a PDZ protein which uses different mechanisms to control the function of several cell surface proteins. NHERF has been proposed to maintain the Na+/H+ antiporter NHE3 in an inactive state by linking it to PKA, which is thought to downregulate it (Lamprecht et al., 1998; Weinman et al., 2000). In addition NHERF has been described to be required for activation of NHE3 upon stimulation of the β2 adrenergetic receptor, which is a G-protein coupled receptor (Hall et al., 1998a; Hall et al., 1998b). NHERF can also directly regulate the activity of the β2 adrenergetic receptor by controlling its recycling and subcellular localisation (Cao et al., 1999). The most striking functional aspect of NHERF is probably the direct modulation of CFTR[103] chloride

---

[100] nNOS: neuronal Nitric Oxide Synthase

[101] GKAP: Guanylate Kinase Associated Protein

[102] SAP97: Synapse Associated Protein 97

[103] CFTR: Cystic Fibrosis Transmembrane conductance Regulator

channel via binding of its two PDZ domains. NHERF as well as CAP70[104], another PDZ domain containing protein, have been described to bind to the C-terminus of the CFTR channel, thus increasing cellular conductance by binding two separate channel subunits via the tandem PDZ of NHERF, inducing an activating conformational change (Bezprozvanny & Maximov, 2001; Raghuram et al., 2001; Wang et al., 2000).

Organisation of cell-cell junctions and cell polarity:

In addition many PDZ domain containing proteins, such as the ZO[105] proteins have been described to cluster and organise protein complexes at cellular junctions, such as adherens junctions or tight junctions. They mediate cell signalling, organise cell-cell barriers or apico-basal compartimentation of cells and cell polarity. These examples are not mentioned here and will be discussed in later chapters.

Cellular localisation of PDZ domains targeted by high-risk E6

Most of the PDZ domain containing proteins are located at the membrane-cytoskeleton interfaces of cell-cell contacts or at cell junctions. They form multiprotein (signalling) complexes at the inner interface of the membrane in order to modulate cell growth, cell polarity and cell adhesion in response to cell contact (Craven & Bredt, 1998; Fanning & Anderson, 1999). Especially epithelial cell junctions are implicated in the regulations of cell proliferation and differentiation and their components include several oncogenes such as APC[106], β-catenin and E-cadherin (for review see: Barth et al., 1997). Two common sites of cell-cell contact in epithelial cells are the adherens junction and the tight junction. The next paragraphs will provide basic information about these two types of cellular junctions, prior to go into details of E6 interactions with PDZ domain containing proteins and their possible biological consequences.

Adherens junctions

Adherens junctions are built of protein complexes that occur at cell-cell junctions in epithelial tissues. They have been described either as bands which

---

[104] CAP70: CFTR Associated Protein of 70 kDa
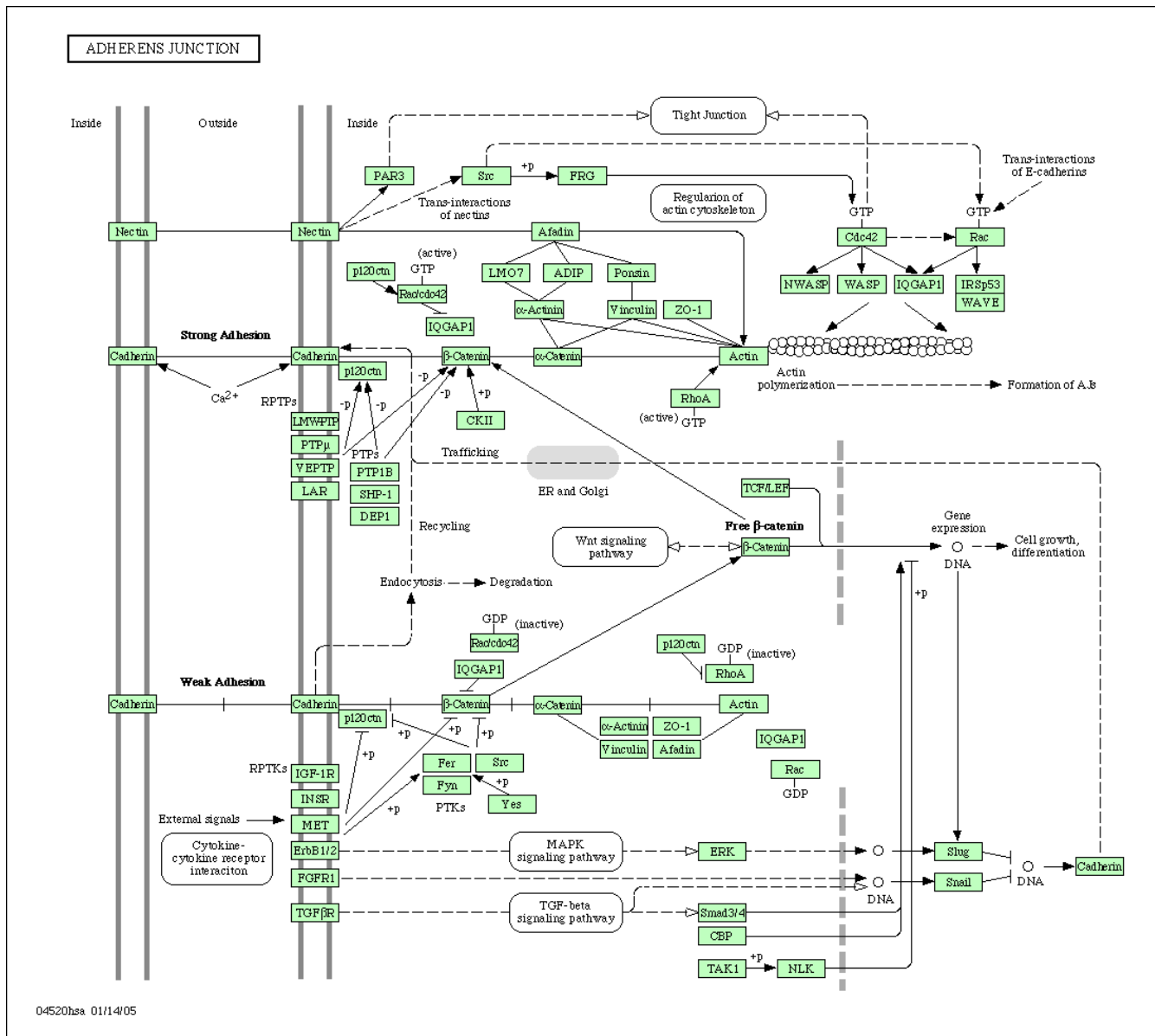[105] ZO: Zonula Occludens
[106] APC: Adenomatous Polyposis Coli

Figure 26.: Schematic representation of the protein complexes at adherens junctions. Figures are taken from the Kyoto Encyclopedia of Genes and Genomes (KEGG) (http://www.genome.jp/kegg/)

encircle the cell (zonula adherens) or as spots of attachment to the extracellular matrix (adhesion plaques). Three major proteins are implicated into the formation of these junctions, namely different types of cadherins, β-catenin and α-catenin. Cadherins can be divided into classic cadherins (E- for epithelial, P- for placental and N- for neuronal) which are found in the adherens junctions and desmosomal cadherins which are found in the desmosomes (Yap et al., 1997). Classical cadherins have an extracellular part comprising 5 domains which interacts in the extracellular space with cadherins of adjacent cells in a calcium dependent manner. In the cytosol cadherins present a conserved cytoplasmic domain which associates with intracellular proteins and actin filaments thereby forming the junctional structure (Gates & Pfeifer 2006; Takeichi, 1995). The cytoplasmic domain of E-cadherin binds to β-catenin or γ-catenin which interacts with α-catenin. Finally α-catenin forms either a direct linkage to the actin cytoskeleton or an indirect one through α-actinin, vinculin, ZO1 or spectrin and a number of other molecules associated to cadherin complex (Gates & Peifer, 2006; Yamada & Geiger, 1997). Therefore adherens junctions serve as a bridge connecting the actin cytoskeleton of neighboring cells through direct interaction and may serve as a regulatory module to maintain the actin contractile ring with which it is associated in microscopic studies (Figure 26).

Tight junctions

Tight junctions, also known as zonula occludens, are the connections between epithelial and endothelial cells that comprise various tissues of the body. They occur where adjacent cells are closely associated and form virtual impermeable barriers to fluid in the cellular interstitium by tightly joining the cellular membranes. When using conventional electron microscopy, tight junctions appear as regions of tight apposition of neighbouring cells, where the plasma membrane of adjacent cells appear to fuse together (so-called kissing points). Their main functions can be summarised in three points. Firstly they hold epithelial cells together. Second they block the movement of integral membrane proteins between the apical and the basolateral surfaces of the cell. Thus, they confer apico-basal polarity to the cells allowing specialised functions on each surface such as for example endocytosis at the apical side and exocytosis at the basolateral side. Third, they regulate the passage of molecules across these natural barriers by hindering or controlling the diffusion of
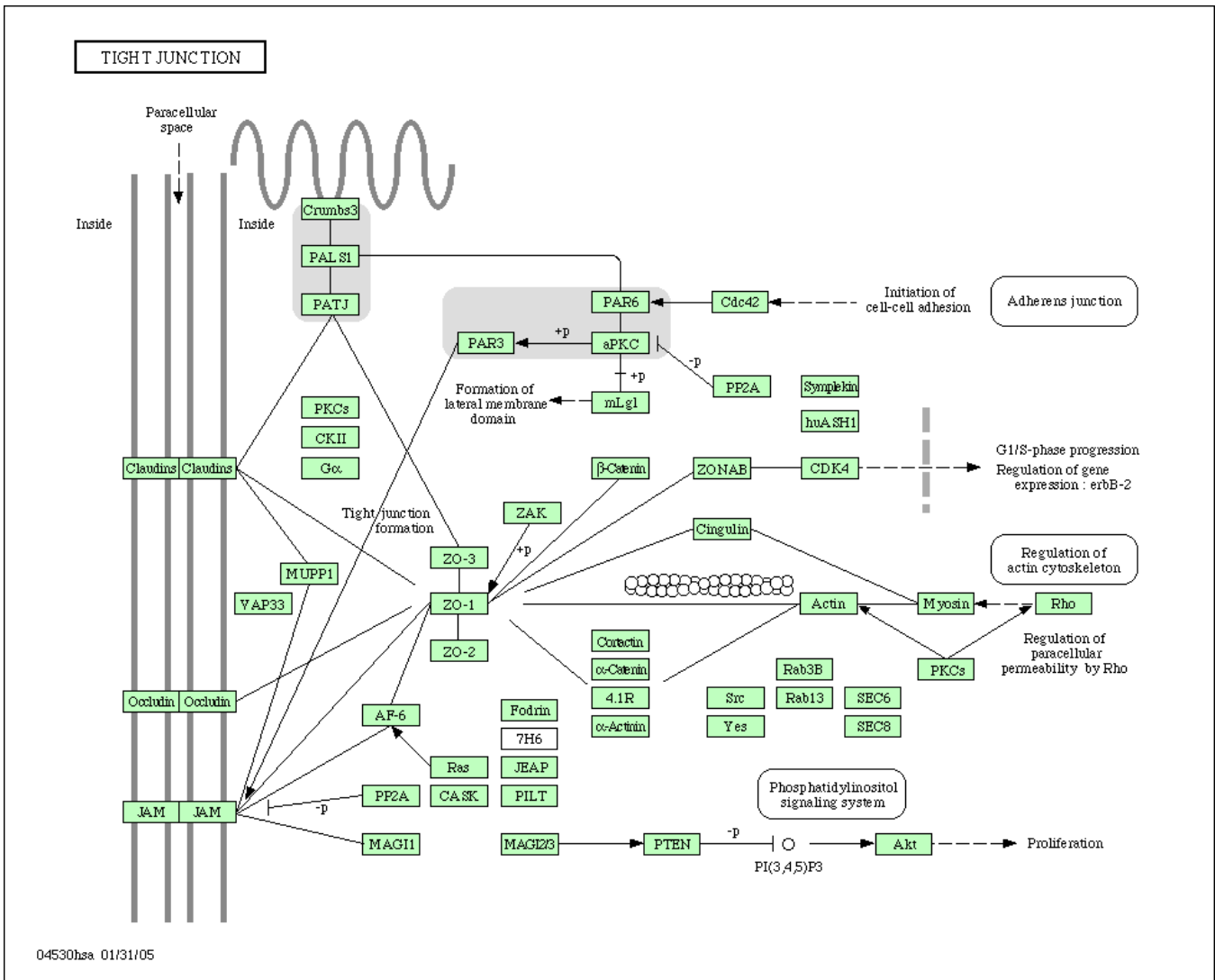
Figure 27.: Schematic representation of the protein complexes at tight junctions. Figures are taken from the Kyoto Encyclopedia of Genes and Genomes (KEGG) (http://www.genome.jp/kegg/)

molecules and ions through the cellular interstitium. Therefore, they provide control of the molecular traffic since the molecules have to pass through the tissue either by diffusion or active transport. A good example for that function is the important role tight junctions play in the blood-brain barrier. As part of the body's normal activity, tight junctions selectively open and close in response to various signals both inside and outside of cells. This allows the passage of large molecules or even entire cells across the tight junction barrier. Epithelia can be classed as "tight" or "leaky" depending on the ability of the tight junctions to prevent water and macro-molecular movement. Tight junctions consist of transmembrane proteins such as occluding, claudin and JAMs[107] which are anchored in the membranes of two adjacent cells and interact with each other to hold the cells together. Tight junction membrane proteins interact with scaffold proteins (e.g., ZO-1[108] and its isoforms ZO-2 and ZO-3) and connect them with the cytoskeleton or with various proteins involved in signal transduction and the transcriptional control tight junction functions. A great number of tight-junction associated proteins contain PDZ domains and act as scaffolds to recruit other proteins to the tight junction. Therefore, beside its structural function as a junction and a diffusion barrier, tight junctions can be seen also as multiprotein complexes, which function as signalling platforms (for review read Shin et al., 2006) (Figure 27).

PDZ domain containing proteins targeted by high-risk HPV E6

Today, 10 PDZ domain containing proteins have been described to interact with the C-terminal class I consensus binding motif of high risk HPV E6. These proteins include hDLG and hScrib, which are human tumour suppressor proteins homologous to the *D. melanogaster* proteins DLG and Scrib, respectively (Kiyono et al., 1997; Lee et al., 1997; Nakagawa & Huibregtse, 2000), MUPP1 (Lee et al., 2000), Tip-1[109] (Hampson et al., 2004), Tip-2[110] (Favre-Bonvin et al., 2005), PATJ[111] (Latorre et al., 2005), CAL[112] (Jeong et al., 2006) and the members of the MAGUK

---

[107] JAM: Junctional Adhesion Molecule

[108] ZO-1: Zonula Occludens-1

[109] Tip-1: Tax interacting protein 1

[110] Tip-2: Tax interacting protein 2; aka GIPC (GAIP Interacting Protein C-terminus; GAIP=Galpha-Interacting Protein)

[111] PATJ: PALS1-Associated Tight Junction protein (PALS1=Protein Associated with Lin Seven-1)

[112] CAL: CFTR Associated Ligand (CFTR=Cystic Fibrosis Transmembrane Regulator)
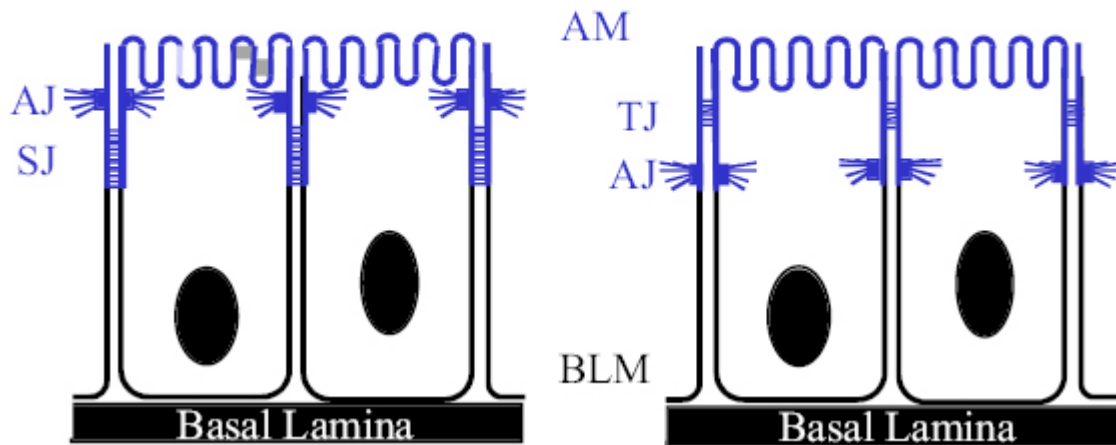
Figure 28.: Schematic representation of the membrane domains and junctions of polarised epithelial cells in invertebrates and vertebrates. The apical membrane (AM) domain of the cells depicted in blue contains microvilli facing the lumen or exterior of the cell. The apico-lateral membrane contains specialised cell junctions, including on one side the adherens junction (AJ) and septate junction (SJ) in invertebrates and on the other side the tight junction (TJ) and adherens junction in vertebrates. In invertebrates DLG localises to the SJ in *D. melanogaster* and at the adherens junction in *C. elegans*. hDLG localise to the tight junctions in vertebrates (Caruana, 2002). BLM: Basolateral Membrane

familiy MAGI-1, MAGI-2 and MAGI-3 (Dobrosotskaya et al., 1997; Glausinger et al., 2000; Thomas et al., 2002).

hDLG: The first PDZ domain containing protein which has been described to interact with high-risk E6 was hDLG (Kiyono et al., 1997; Lee et al., 1997). This protein is a member of the MAGUK family of proteins (Figure 23). MAGUKs are in general multifunctional proteins localised at the membrane-cytoskeleton interface of cell-cell junctions where they function as structural and signalling proteins (Anderson, 1996; Craven & Bredt, 1998). In *D. melanogaster,* DLG is localised at septate junctions, which are the homologues of the human tight junctions (Figure 28). It has multiple functions, including the regulation of cell proliferation as well as the maintenance of septate junctions and apicobasal cell polarity (Woods & Bryant, 1991; Woods et al., 1996). Mutation of the *dlg* gene results in loosening of cell-cell contact, neoplastic proliferation of imaginal disc epithelial cells and a prolonged larval period followed by death (Lue et al., 1994; Woods & Bryant, 1991). Homozygous loss of the *dlg* gene leads to overgrowth of the imaginal discs and loss or severe reductions of the septate junctions in both imaginal and larval epithelia. Neuromuscular junctions are also reduced in larvae with homozygous loss of *dlg*, indicating that DLG plays a structural role in the formation of both synaptic and septate junctions (Lahey et al., 1994; Woods et al.; 1996). In mutants with "weak" alleles, morphologically normal junctions form but cell proliferation proceeds unchecked. Lethal tumours develop probably due to defects in growth–inhibiting signals normally propagated at the junction (Bryant et al., 1993). This separation of phenotypes suggests both a structural role for DLG in junction formation and its requirement for DLG in cell signalling (Anderson, 1996).

hDLG, the human homologue of the Drosophila tumour suppressor DLG is expressed in different cell types including eptihelia, where it is localised at regions of cell-cell contact contributing to the formation of adherens junctions. It has been described to regulate cell-cell adhesion, apicobasal polarity and proliferation in epithelial tissues (Lue et al., 1994; Wodarz et al., 2000). In fact, the correct function of epithelial cells is correlated with their apico-basal polarity which is also often coupled to proliferation control (Humbert et al., 2003). Overexpression of hDLG blocks cell-cycle progression from G0/G1 to S-phase in mouse fibroblasts. Another

study described a murine DLG truncation mutant that resulted in growth retardation *in utero*, cranofacial abnormalities and perinatal lethality during murine development (Caruana & Bernstein, 2001). hDlg has been localised at adherens junctions of epithelial cells where it is recruited by E-cadherin (Ide et al., 1999; Reuver & Garner, 1998). It has been described to interact through different domains with different proteins, including Shaker-type K+ channels (Kim et al., 1995), cytoskeletal proteins 4.1 (Lue et al., 1994; Marfatia et al., 1996), adenovirus E4-ORF1 (Lee et al., 1997), HTLV-1[113] Tax protein (Suzuki et al., 1999) and the APC tumour suppressor protein (Ishidate et al., 2000; Matsumine et al., 1996), which is mutated in the majority of colon cancers (Kinzler & Vogelstein, 1996).

Indeed, complex formation between hDLG and APC was reported to block cell cycle progression (Ishidate et al., 2000). HPV E6 targets hDLG for ubiquitine-mediated degradation (Gardiol et al., 1999). However there are contradictory results if the degradation is E6AP dependent or not and in consequence mediated by other ubiquitin-ligases (Matsumoto et al., 2006; Sterlinko & Banks, 2004). Since hDLG is ubiquitinated and degraded by the proteasome in non HPV transformed cells, it has been proposed, that presence of E6 amplifies the phenomenon and deregulates the normal, physiological turnover of hDLG, thus deregulating its activity *in vivo* (Gardiol et al., 1999; Mantovani et al., 2001).

In fact, epithelial cells organise as layers of cells with an apico-basal polarisation which is maintained by tight junctions. The interaction between E6 and hDLG might be necessary at a defined point during the viral life cycle, in order to disrupt cell junctions and to abolish cell polarity, thereby altering the normal maturation and inducing proliferation of the infected keratinocytes.

Differences according to the aggressivity of the cancer and the metastatic potential of tumours according to the HPV types infecting the cells have been reported and might be correlated to the ability to bind and mediate degradation of PDZ domain containing proteins. For example, cervical tumours associated to HPV 18 have been reported to be more invasive and to present a higher risk of recurrence than those associated to HPV 16 (Burnett et al., 1992; Kurman et al., 1988; Zhang et al., 1995). In agreement with this observation it has been described, that HPV 18 E6 interacts with a higher affinity with hDLG than does HPV 16 E6 (Pim et al., 2000) but

---

[113] HTLV-1: Human T-cell Leukemia Virus type 1

on the other hand HPV 18 E6 has been reported to degrade p53 less efficient than HPV 16 E6 (Scheffner et al., 1990). This might highlight the importance of disrupting cell-cell adhesion by degrading PDZ domain containing scaffold molecules such as hDLG. The higher affinity of HPV 18 for PDZ2 of hDLG might be explained by the ETQV consensus binding motif, which seems to be more adapted to insert into the pockets of the binding groove on PDZ2 of hDLG than does the ETQL consensus motif of HPV 16 E6. The hydrophobic pocket might not be deep enough to bury a large L residue, while a V might be fitting. In consequence HPV 18 E6 degrades more efficiently hDLG (For more details about the peptide PDZ-binding see chapter: "Structural basis of PDZ domains and ligand recognition") (Gardiol et al., 1999; Pim et al., 2000).

Another interesting observation is that low-risk HPV E6 proteins fail to interact with and to degrade PDZ domain containing proteins such as hDLG because of the absence of a C-terminal PDZ domain binding consensus motif (Kiyono et al., 1997; Lee et al., 1997; Gardiol et al., 1999; Pim et al., 2000). If providing them with a high-risk HPV C-terminal consensus binding motif they bind and degrade hDLG (Pim et al., 2000). In contrast, cutaneous HPV E6 proteins are unable to degrade hDlg even when supplying them with a high-risk HPV PDZ-binding motif (Pim et al., 2002). This suggests that the mucosal HPV E6 proteins, whether derived from high or low-risk HPV types, can similarly interact with the cellular proteolytic machinery, whereas the high-risk cutaneous HPV E6 proteins appear to significantly differ in this aspect.

hScrib: hScrib[114] is the human homologue of the *D. melanogaster* tumour suppressor protein scribble. In *D. melanogaster* Scribble localises at the septate junctions, a structure which is equivalent to the vertebrate tight junctions (Figure 28), and serves as an apico-basal polarity determinant in epithelial cells. Mutations in the *scribble* gene causes disruption of the cell polarity and leads to the overgrowth of epithelial cells in the imaginal discs, follicle and "brain" (Bilder et al., 2000a; Greaves, 2000; Wodarz, 2000). hScrib is a membrane associated protein, containing sixteen LRRs[115] and four PDZ domains (Figure 23). It is expressed and localised at epithelial tight junctions and adherens junctions, where it colocalises with E-cadherin

---

[114] hScrib: human homologue of the scribble tumour suppressor protein (*D. melanogaster*)
[115] LRR: Leucine Rich Repeat

(Nakagawa et al., 2004). It cooperates with DLG and Lgl[116] to control both formation of cell junctions and inhibition of epithelial cell growth, possibly through controlling the localisation of growth factor receptors and signalling molecules (Bilder & Perrimon, 2000; Bilder et al., 2000; Dow et al., 2003).

It is a substrate for ubiquitination by the E6-E6AP complex *in vitro* and for proteasome degradation mediated by high-risk E6 *in vivo* (Nakagawa & Huibregtse, 2000). Murine Scrib has been described as the mutated protein in the mouse circletail mutant. These mice die perinatally from neural tube defects and show defects in tissue patterning in the embryonic cochlea (Montcouquiol et al., 2003; Murdoch et al., 2003). This indicates that this protein plays an important role in development as well in the regulation of planar cell polarity (Dow et al., 2003). In endocervical columnar epithelium hScrib has been located to the basolateral cell layers. In SILs expression of hScrib diminishes with progression of the disease from L-SILs to invasive cancers. In H-SILs and invasive cancers hScrib levels were very weak or almost negative, which correlates with disruption of cell-cell junctions and mislocalisation of junctional proteins in invasive cancers (Nakagawa et al., 2004). In addition recent studies demonstrate interaction of hScrib with the tumour suppressor APC. Knock down experiments of hScrib showed disruption of the proper localisation of APC at the adherens junctions. Disruption of the interaction of hScrib with APC generated a loss in its negative regulatory ability of cell-cycle progression from G1 to S phase, underscoring that hScrib negatively controls cell-cycle progression by associating with APC (Takizawa et al., 2006). It has been shown that there are differences in the relative efficiency of different E6 proteins to target hScrib. HPV 16 seems to degrade hScrib more efficiently than does HPV 18 E6, whereas HPV 18 E6 generally degrades other PDZ domain containing proteins, as for example hDLG, more efficiently, as described in the paragraph afore. This was put in relation with its more adapted ETQV consensus binding motif and correlates with its higher agressivity (Thomas et al., 2005; Pim et al., 2000).

MUPP1: HPV 18 E6 has also been shown to bind to MUPP1, a large multi PDZ domain containing scaffold protein with putative role in signal transduction. MUPP1 is a PDZ only protein containing 13 PDZ domains thus being the protein bearing the

---

[116] Lgl: Lethal giant larvae; a neoplastic tumour suppressor in *D. melanogaster*

largest number of PDZ domains yet reported (Figure 23). It has the capacity to assemble cellular proteins into a multitude of signalling complexes. It has been described to be present both in the cytoplasm and in discrete regions of cell-cell contacts such as tight junctions (Lee et al., 2000; Ullmer et al., 1998). Indeed, MUPP1 interacts selectively through its 10th PDZ domain with the cytoplasmic portion of the 5-HT2C serotonin receptor (Ullmer et al., 1998; Becamel et al., 2001) and with the unphosphorylated c-Kit tyrosine kinase receptor (Mancini et al., 2000). The increasing number of plausible links between alteration of the localisation of PDZ proteins and cancer, let suspect, that MUPP1 might be negatively involved in cellular proliferation and to be a potential tumour suppressor protein (Gonzalez-Mariscal et al., 2003; Lee et al., 2000). HPV 18 E6 has been described to interact with MUPP1 via its C-terminal consensus binding motif and to reduce MUPP1 half-life *in vitro* and *in vivo*, implicating that MUPP1 function is abrogated in HPV 18 infected cells (Lee et al., 2000; Mantovani & Banks, 2001). Destruction of MUPP1 by E6 could thereby interfere with the assembly of signalling complexes at the epithelial cell membranes. Whether degradation occurs via ubiquitin-mediated, proteasome dependent proteolysis or whether E6 uses other degradation pathways is not totally clear (Lee et al., 2000). MUPP1 is also bound by other viral oncoproteins such as the Ad 9 E4-ORF1[117], which abolishes its activity not by degradation but by sequestering it to the cytoplasmic bodies (Lee et al., 2000; Glausinger et al., 2000). Additional results showing that low-risk HPV E6 proteins neither bind MUPP1 nor degrade it, might be a supplementary evidence that binding of high-risk E6 proteins to MUPP1 may contribute to HPV-induced carcinogenesis (Lee et al., 2000).

MUPP1 has been described as a paralogue of PATJ. PATJ contains 10 PDZ domains and concentrates at the tight junctions of epithelial cells where it is involved in regulating their integrity by clustering proteins into junctional complexes (Lemmers et al., 2002; Roh et al., 2002). PATJ is in fact a key component of the CRB3-PALS-PATJ[118] polarity complex which is required for tight junction establishment and proper apicobasal polarity in epithelial cells (Shin et al., 2005). The fact that MUPP1 also binds PALS1 let suggest that it shares common functions with PATJ (Roh & Margolis, 2003).

---

[117] Ad 9 E4-ORF1: E4 Open Reading Frame of the Adenovirus type 9
[118] CRB3: Crumbs homologue 3

Tip-1: Another PDZ mediated interaction partner of high-risk HPV E6 proteins is the Tip-1 protein. It is an unusual PDZ protein as it contains only one PDZ domain (Figure 23). It therefore represents a departure from the classic PDZ protein clustering scheme since these proteins usually rely on multiple PDZ domains in order to scaffold membrane, cytoskeletal and signalling proteins. As a consequence Tip-1 cannot cluster any proteins and has been described as acting as a scaffold antagonist, thereby acting following a novel mechanism. Tip-1 directly associates with the PDZ binding motif of Kir2.3 and competes for interaction with mLin-7, a scaffolding protein, thereby hindering a channel to form at the cell surface (Alewine et al., 2006). In addition, Tip-1 has been reported to bind to the HTLV-1 Tax protein and to have a possible role in RhoA signalling (Reynaud et al., 2000). Tip-1 has also been described to interact with L-glutaminase and β-catenin (Kanamori et al., 2003; Olalla et al., 2001). Interaction of high-risk HPV E6 oncoprotein has been reported to increase the motility of C33A tumour cells, which behaved less cohesive than control cells (Hampson et al., 2004). However, Tip-1 is not degraded by high-risk HPV E6 proteins and it has been proposed that it may be a gain of function target of E6 in the control of tumour cell formation and migration (Hampson et al., 2004).

Tip-2: Tip-2 has initially been identified as a PDZ protein which interacts with HTLV-1 Tax (Rousset et al., 1998) and the GTPase-activating protein for Gal GAIP (De Vries et al., 1998). It has been shown by yeast two hybrid and co-immunoprecipitation experiments that HPV 18 E6 interacts with the PDZ domain of Tip-2. E6 and E7 silencing revealed that E6 mediates continuous degradation of Tip-2 (Favre-Bonvin et al., 2005). However, as already denoted for other PDZ domain containing proteins, Tip-2 is not entirely eliminated from the cell, which might be explained by the presence of limited amounts of E6 or the partial inactivation of its C-terminal binding motif via phosphorylation by PKA (Kuhne et al., 2000). In addition, it remains to be confirmed that this partial degradation is mediated via E6AP or another ubiquitine-ligase. Degradation of Tip-2 by high-risk HPV E6 oncoprotein might contribute to loss of cell-cell contact and malignant

hyperproliferation. Tip-2 participates in the regulation of TGF-β[119] signalling and the effect of E6 on Tip-2 raise the possibility of a positive feedback loop in which degradation of Tip-2 renders cells less sensitive to TGF-β, thereby increasing the expression of E6, since TGF-β downregulates the expression of E6 and E7 *in vivo* (Baldwin et al., 2004; Favre-Bonvin et al., 2004).

CAL: The CFTR associated ligand, CAL, is a 454 amino acid protein composed of two coiled-coiled domains and one PDZ domain (Figure 23) by which it dimerises and interacts with other cellular partners. Overexpressed CAL diminishes the half-life of mature CFTR by enhancing its lysosomal degradation (Cheng et al., 2002; Cheng et al., 2005). In addition, CAL is associated to and targets a multitude of different cellular proteins among which EGF receptors, CALEB/NGC (Hassel et al., 2003) and a chloride channel protein CLC-3B (Gentzsch et al., 2003). It has been proposed that CAL takes part in intracellular trafficking of these proteins (Charest et al., 2001; Chiang et al., 2001; Neudauer et al., 2001). Cal has been identified as a Golgi-associated interaction partner of high-risk HPV E6 oncoproteins. It interacts with E6 via its PDZ domain leading to enhanced E6AP mediated proteasomal degradation. However, in absence of E6, the coiled-coil region is degraded more efficiently than the PDZ domain, suggesting a natural turnover of CAL by the ubiquitin-mediated proteasomal degradation. Whether E6AP alone directly binds to and ubiquitinates CAL *in vivo* remains unclear. HPV 16 binds with higher affinity and degrades more efficiently CAL than HPV 18. A similar binding behaviour was described for hScrib (Thomas et al., 2005) and might be explained by a cavity in the PDZ binding groove which is sufficiently deep to accommodate the C-terminal L residue from the HPV 16 PDZ consensus binding motif. However, a V to L mutation in HPV 18 E6 does not completely restore the binding affinity suggesting that other residues beside the consensus binding motif are implicated in ligand binding affinity (Jeong et al., 2006).

MAGI: Last but not least, E6 has been described to interact with MAGI-1[120] (Dobrosotskaya et al., 1997; Ide et al., 1999) and with the close related proteins

---

[119] TGF: Transforming Growth Factor

[120] MAGI-1 identifies per definition the rat protein. Its human counterpart is called Brain Angiogenesis Inhibitor 1 (BAI1)-associated protein (BAP1) or at least hMAGI-1. However in this manuscript I will refer to MAGI-1 as the human protein, for simplicity (Mino et al., 2000).
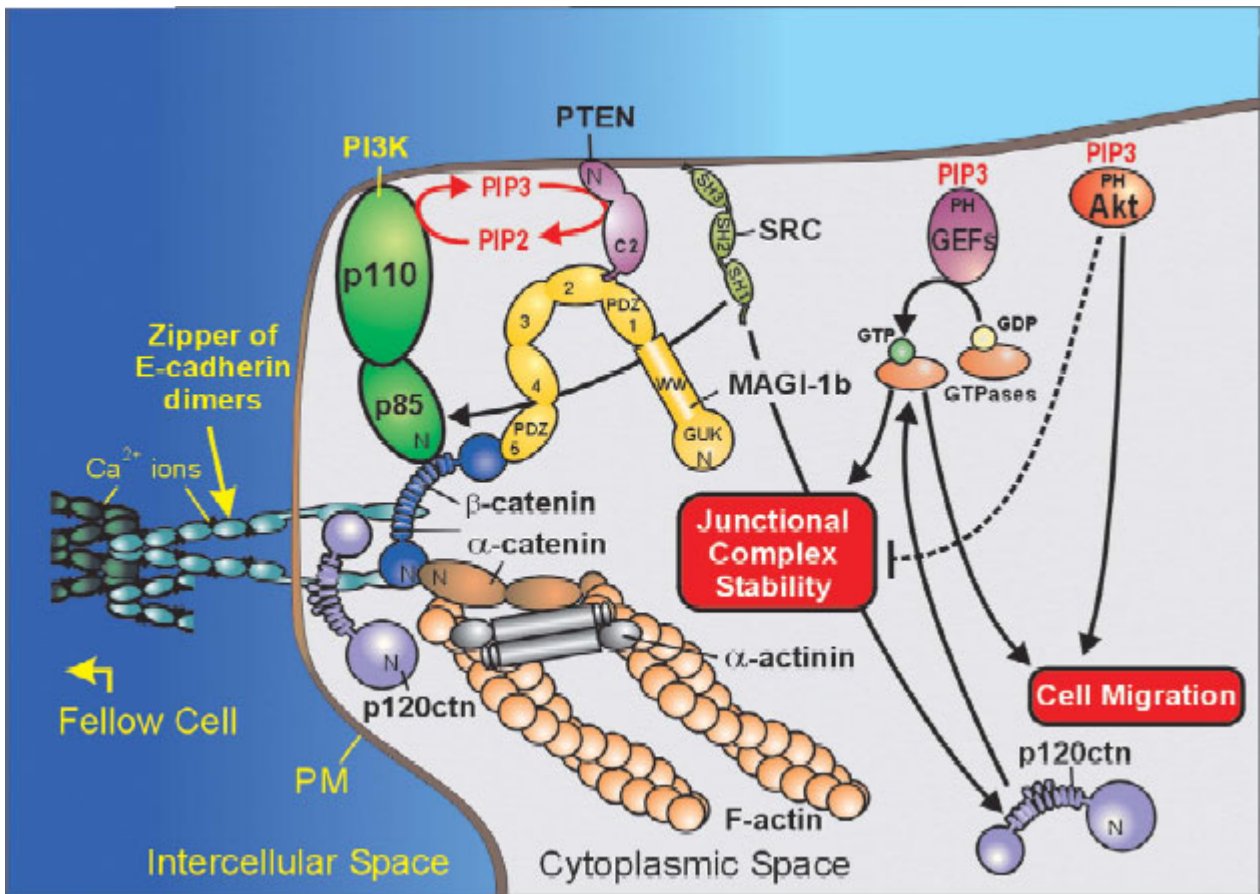
Figure 29.: Schematic representation of stabilisation of junctional complexes containing E-cadherin by MAGI-1B/PTEN signalosome. PTEN is targeted to the plasma membrane and its PDZ domain anchors it at adherens junctions via interaction with MAGI-1B, which itself binds â-catenin. Activated Src recruits PI3K to the plasma membrane, where it interacts with â-catenin. The antagonistic activites of PTEN and PI3K allow subtle and local regulation of second messenger PIP₃ and the recruitment of AKT and GTP exchange factors for the Rho family of GTPases. AKT activity is necessary and sufficient for Src-induced cell migration. PM = plasma membrane (Kotelevets et al., 2005).

MAGI-2[121] and MAGI-3. These proteins belong to the MAGUK family of proteins and are frequently found on the cytoplasmic side of the plasma membrane at regions of cell-cell adhesion in epithelial cells and neurons. MAGUKs in general have been described as scaffolding proteins which assemble multimolecular protein complexes at functionally relevant subcellular sites in polarised epithelia and neurons (Laura et al., 2002). The first MAGI was identified in 1997 as a protein which interacts with K-RAS (Dobrosotskaya et al., 1997) while the second and the third isoforms were identified later in 2002 as proteins which interact with the PTEN[122] tumour suppressor phosphatase (Figure 29) (Wu et al., 2000a; Wu et al., 2000b). MAGI-1 and MAGI-3 are widely expressed in various adult tissues, whereas MAGI-2 is mainly expressed in the brain (Hirao et al., 1998; Shoji et al., 2000; Wood et al., 1998). MAGI-1 and MAGI-2 RNAs are alternatively spliced (Dobrosotskaya et al., 1997; Hirao et al., 2000; Patrie et al., 2001).

Studies indicate that MAGI-1 is alternatively spliced on the carboxy terminus, yielding three distinct forms, referred to as MAGI-1A (+16 residues), MAGI-1B (+48 residues) and MAGI-1C (+251 residues), bearing each additional residues (indicated in brackets) at the C-terminus of the PDZ5 (Figure 30) (Dobrosotskaya et al., 1997). Two additional alternatively spliced domains were identified, referred to as α and β domains (Figure 30) (Laura et al., 2002). MAGI-1C appears to be expressed in tissues with the highest expect of epithelial cells, such as colon, kidney, lung, liver and pancreas. In addition, MAGI-1C bears the longest C-terminus which is highly charged and contains three segments which conform to the consensus sequence for a bipartite NLS[123]. Cellular localisation experiments in MDCK[124] cells showed a major nuclear localisation (Dobrosotskaya et al., 1997) suggesting that it is capable to interact with transcription complexes (Laura et al., 2002). The B-form has been described to be mainly expressed in brain and heart tissue, suggesting that these epithelial-poor tissues may contain a functionally different form. The A-form appears

---

[121] MAGI-2 is the neural isoform of MAGI-1 and is also known as S-SCAM (Synaptic Scaffold Molecule), atrophin-interacting protein and activin receptor-interacting protein (Hirao et al., 1998; Shoji et al., 2000; Wood et al., 1998).

[122] PTEN: Phosphatase and tensin homologue. A phosphatase which is located on chromosome 10q23. It has been described as a tumour suppressor. It is linked with cell regulation and apoptosis and has been shown to be mutated in several human malignancies. Mutations in the PTEN gene are documented in cancers of the breast, prostate, endometrium, ovary, colon, melanoma, glioblastoma and lymphoma. Animal models have shown that the loss of just one copy of the PTEN gene is enough to interrupt cell signalling and begin the process of uncontrolled cell growth. However, the significance of PTEN alterations in carcinogenesis is controversial since aberrant transcripts of PTEN have also been identified in normal non-cancerous tissues (Smith & Ashworth, 1998; Wang & Chang, 1999).

[123] NLS: Nuclear Localisation Signal
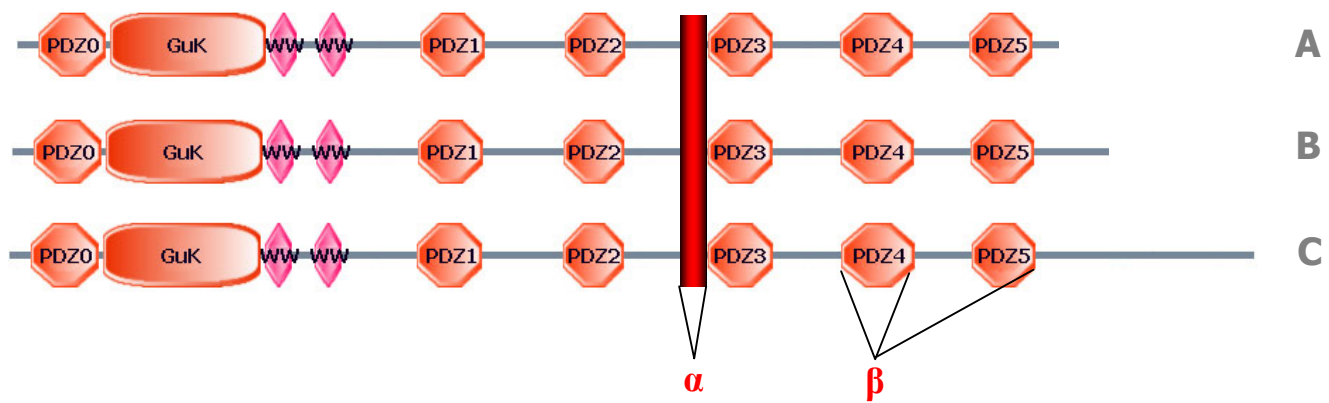
[124] MDCK: Madin-Darby Canine Kidney

Figure 30.: Schematic representation of the different splice variants of MAGI-1. Domain structure of MAGI-1A, MAGI-1B and MAGI-1C. The known alternatively spliced domains called á and â are indicated in red (Figure adapted from Laura et al., 2002).

to be expressed in brain and pancreas (Laura et al., 2002). The β-form generates MAGI-1 proteins which are deleted for their PDZ4 or both PDZ4 and PDZ5 functionality (Laura et al., 2002). Loss of the β domain would result in a protein which is incapable of binding to ligands of the PDZ4 domain. The α-splice variant seems to be essentially expressed in brain tissue (Laura et al., 2002). In summary, MAGI-1 shows complex alternative splicing which, by analogy to other MAGUKs, is likely to be important in modulating its functions spatially in distinct tissues or temporally during development.

MAGI-1 has been described to be targeted to the tight junctions, independently of its alternative splicing forms, suggesting that splicing plays other roles in the epithelial function. However, the MAGI-1B isoform has also been identified as a scaffolding protein in epithelial adherens junctions (Kotelevets et al., 2005).

In adherens junctions and probably in tight junctions MAGI-1 forms a complex with β-catenin through its PDZ5 (Dobrosotskaya & James, 2000; Kotelevets et al., 2005), the expression of which is deregulated in many human cancers (Polakis, 1999).

E6 proteins from high-risk, but not low-risk HPV, types were shown to bind to MAGI-1 (Dobrosotskaya et al., 1997; Ide et al., 1999), via the PDZ1 domain (out of six) (Thomas et al., 2001). Interaction with high-risk E6 reduces steady-state levels and half-life of MAGI-1 (Glausinger et al., 2000).

MAGI-1B has been described to bind selectively via PDZ2 to PTEN and via PDZ5 to β-catenin. The latter interaction is responsible for anchoring MAGI-1B at the adherens junction, where it could be colocalised in addition with E-cadherin under physiological conditions. Interestingly, MAGI-1B, as well as the closely related proteins MAGI-2 and MAGI-3, are involved in the regulation of the PTEN tumour suppressor, a component of the Akt kinase signalling pathway that promotes cell survival and proliferation (Kotelevets et al., 2005; Marte & Downward, 1997; Wu et al., 2000a; Wu et al., 2000b). Therefore it is likely that MAGI degradation, promoted by high risk HPV E6 might also affect Akt signalling and thereby inhibit apoptosis independently of targeting p53, thus representing an alternative mitogenic activity of E6.

Invasion assays showed evidence that the recruitment of MAGI-1B to junctional complexes plays a critical role for the control of cell growth and invasiveness (Kotelevets et al., 2005). In addition MAGI-1B has been reported to co-localise through interaction with JAM4 and α-actinin-4 at the tight junctions to form junctional complexes (Patrie et al., 2002; Hirabayashi et al., 2003).
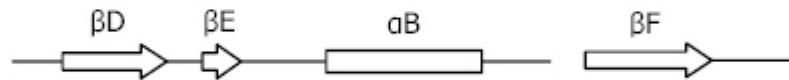
Furthermore, the protein mNet[125] has been proposed to be a natural ligand of the PDZ1 of MAGI-1. This protein is a Rho family nucleotide exchange factor, which is implicated in nuclear signalling events. It might be possible that E6 either competes for MAGI-1 PDZ1 binding and thus influences nuclear signalisation events, or that E6 mediated proteasomal degradation of MAGI-1 is the perturbating event. These are only hypothesises and have to be studied more in detail.

Taken together, high-risk HPV E6 oncoproteins target a class of PDZ domain containing proteins thereby affecting cell contact and polarity. This loss of control in immortalised cells can finally lead to an invasive and metastatic phenotype (Mantovani & Banks, 2001). However E6 mutated in the PDZ binding motif does not disrupt the integrity of epithelial tight junctions (Nakagawa & Huibregtse, 2000). This feature of high-risk E6 has been described as one of the important oncogenic activities besides p53 degradation to induce neoplasia and malignant progression towards cancer (Nguyen et al., 2003; Simonson et al., 2005). However, experiments concerning the localisation of high-risk HPV E6 revealed either a mainly nuclear localisation (Kanda et al., 1991; Masson et al., 2003; Schwalbach et al., 2000; Sherman & Schlegel et al., 1996) or an even distribution in the cytoplasm and the nucleus (Guccione et al., 2002), according to the cell line used. Most of the PDZ domain containing proteins are mainly localised in multimolecular complexes at cell-cell junctional complexes. However some of the PDZ domains targeted by high-risk E6 display also cytoplasmic as well as nuclear localisation, like for example hDLG or some splice variants of MAGI-1 (Dobrosotskaya et al., 1997; Massimi et al., 2004). Obviously there are different pools of membrane bound or soluble PDZ domain containing proteins. Until now it remains unclear whether E6 degrades the PDZ proteins located at cell junctions and thus deregulates cell contact inhibition and cellular polarity or if this is an indirect effect of degrading soluble or nuclear PDZ proteins implicated in signalling pathways. A recent study showed evidence that E6

---

[125] mNET1: Mouse homologue of a human guanine nucleotide exchange factor for the Rho family of small GTPases

Figure 31.: Sequence alignment of selected PDZ domains. The secondary structure elements of the PDZ3 of PSD95 are represented in a box and exemplary shown as arrows (â-sheet), bars (á-helix) and lines (connecting loops). Hydrophobic conserved residues stabilising the hydrophobic core of the protein are coloured magenta. Residues contacting the peptide in the PSD95 PDZ3-peptide complex as well as their counterparts in the homologous domains are underlined in gray. Amino acid numbering corresponds to that of PSD95. The sequences are as follows: hPSD95_3 (PDZ3, residues 309-393 of human PSD95 (Cho et al., 1992)), MAGI1_1 (PDZ1, residues 467-554 of human MAGI-1c (Dobrosotskaya et al., 1997)), hScribb_1, hScribb_2, hScribb_3 (PDZ1, PDZ2 and PDZ3, residues 724-814, 858-949 and 1000-1092 of human Scribble, respectively (Bilder & Perrimon, 2000)), hDLG_3 (PDZ3, residues 462-546 of human DLG (Woods & Bryant, 1991)), Sap97 (PDZ3, residues 461-545 of Sap97 (Muller et al., 1995)), ZO-1, ZO-2 (residues 408-491 and 93-176 of ZO-1 and ZO-2, respectively (Willot et al., 1993; Jesaitis & Goodenough, 1994)), hDLG_1 (PDZ1, residues 221-310 of human DLG (Woods & Bryant, 1991)), hPSD95_1 (PDZ1, residues 61-151 of human PSD95 (Cho et al., 1992)), hDLG_2 (PDZ2, residues 318-413 of human DLG (Woods & Bryant, 1991)), hPSD95_2 (PDZ2, residues 156-246 of human PSD95 (Cho et al., 1992)), nNOS (residues 12-100 of nNOS (Bredt et al., 1991)), hNERF_1 (PDZ1, residues 10-92 of human NHERF (Weinman et al., 1995)), ZASP (residues 1-83 of human ZASP (Faulkner et al., 1999)), p55 (PDZ1, residues 67-15, of p55 (Ruff et al., 1991)).

might preferentially degrade the soluble pools of PDZ proteins and the insoluble pools only in a lesser degree (Massimi et al., 2004).

Structural basis of PDZ domains and ligand recognition

PDZ domains are small globular domains of about 80-100 amino acids. They usually comprise eight segments of secondary structure, six β-strands (βA-βF) and two α-helices (αA-αB), which fold into a six-stranded β-sandwich, with the βD strand participating in both sheets (Figure 31, 32A). The N-terminus and the C-terminus of PDZ domains are close in space in the folded structure. This is a common feature shared by peptide binding domains, such as SH2 or PTB domains, which has been conserved through the course of evolution. It allows easy incorporation into large modular multi-domain proteins ensuring the proper folding of the domain and the peptide binding unit (Harris & Lim, 2001). Structural analysis of the complex reveals that many amino acids are conserved for functional purposes (Figure 31). Other amino acids are clearly conserved for structural purposes as for example the hydrophobic residues which contribute to the hydrophobic core (Figure 31). Conservation of these structural residues indicates that the overall fold will be the same across the different PDZ domains. PDZ domains recognise and bind short C-terminal peptide motifs, which encompass usually 5 amino acids. The peptide binds into an extended groove between the βB strand and the αB helix. It adopts an antiparallel orientation to βB within the PDZ, which is referred to as β-strand addition (Harrison, 1996) (Figure 32A). Therefore, the main chain carbonyl and amide groups of the peptide ligand establish hydrogen bonds comparable to those usually observed in a classical antiparallel β-sheet conformation. The structure of a PDZ domain does not change significantly upon ligand binding. The crystal structures of peptide free and complexed PSD-95 PDZ3 are almost identical with a RMSD of 0.9 Å between the α-carbon atoms (Doyle et al., 1996). The βA-βB connecting loop contains the sequence GLGF, after which the PDZ domains were usually named (Cho et al., 1992). These amino acids play an important function in binding the C-terminal carboxylate group of the peptide. This loop is usually referred to as the carboxylate-binding loop. A molecular surface representation of the complex (Figure 32B) reveals two important general features of the binding site: 1) the peptide binds into a groove on the surface; 2) the C-terminal residue dips into a cavity. As mentioned afore, PDZ
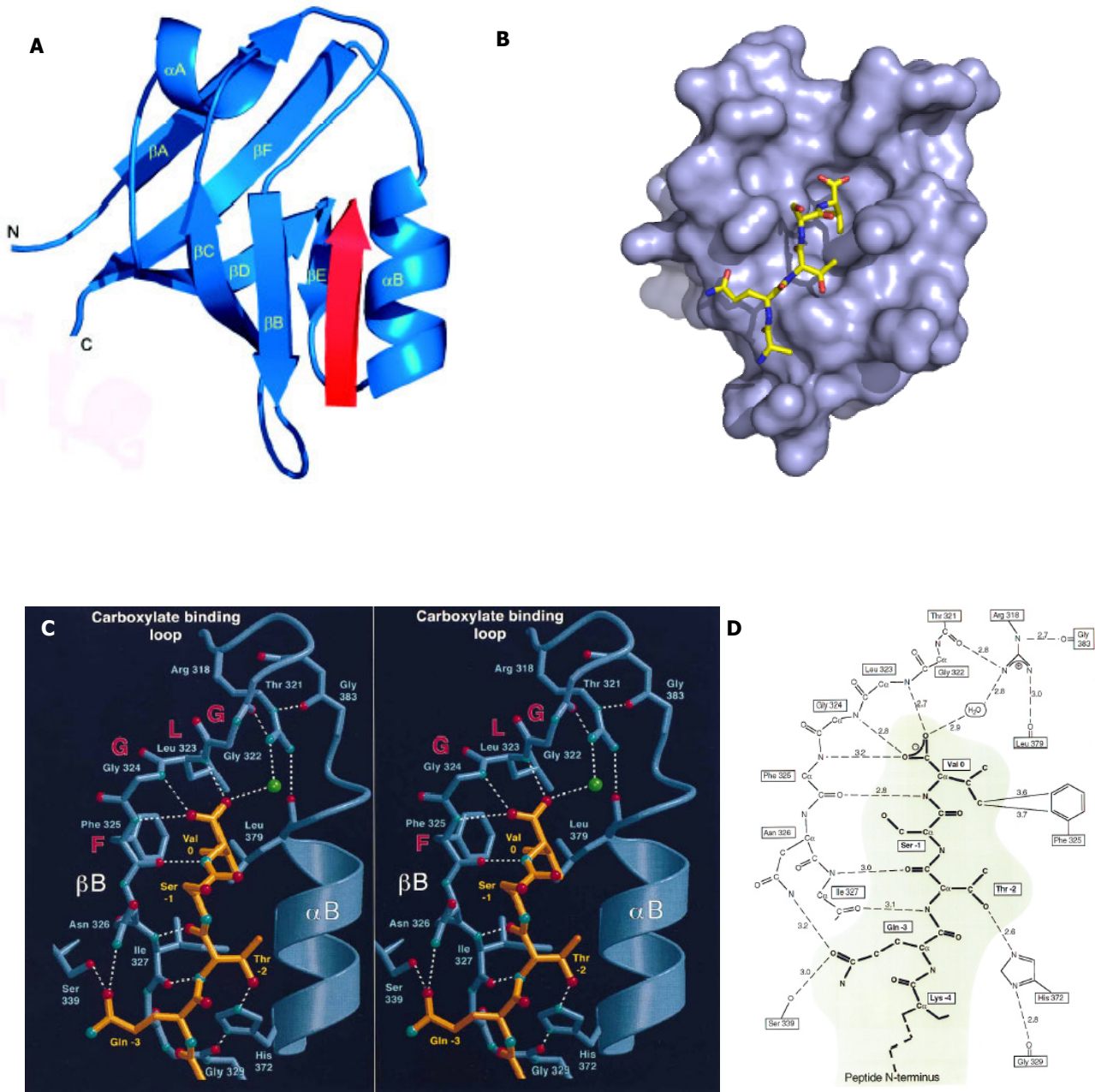
Figure 32.: Structural basis of the PDZ fold and C-terminal peptide interaction. A/ Ribbon representation of the third PDZ domain of PSD95 (blue) with a KQTSV peptide (red) forming an antiparallel â-sheet with the âD strand. Numbering of â-strands and á-helices is shown. (Doyle et al., 1996; PDB code 1be9). **B/** Solvent accessible surface representation of the PSD95 PDZ3 (light blue) complexed to the same peptide than in A (yellow) illustrating the peptide binding groove. **C/** Stereo view and magnification of the peptide bound to the binding groove. Hydrogen bonds are represented as dotted lines. Oxygen atoms are shown in red and nitrogen atoms in blue, the ordered water molecule linking the carboxylate group to R318 in green (Doyle et al. 1996). **D/** Schematic view of all contacts identified in the PSD95 PDZ3-peptide complex. Hydrogen bonds are represented as dashed lines and the two closest atom-to-atom distances are drawn as solid black lines.

domains bind to C-terminal sequence motifs, which comprise four to five residues. Some cases have been described where the binding specificity extends beyond these residues (Cai et al., 2002; Dobrosotskaya et al., 2001; Niethammer et al., 1998). According to the standard nomenclature, residues within the C-terminal peptidic consensus sequence are numbered as P_0 for the C-terminal residue and P_{-1}, P_{-2}, etc… for subsequent N-terminal residues until P_{-n} for the nth residue (Doyle et al. 1996; Kang et al., 2003). S_{-n} identifies the corresponding binding pocket of the PDZ domain (Kang et al., 2003). Peptide library screens combined with sequence analysis allowed to reveal some specificities of certain PDZ domains (Songyang et al., 1997; Schultz et al., 1998). These studies showed that residues P_0 and P_{-2} are important for ligand recognition and allowed to divide the peptide at least into three major classes based on consensus motifs implicating the P_0 and P_{-2} position. Class I PDZ domains recognise a X[S/T]Xɸ-COOH[126], class II PDZ domains a XɸXɸ-COOH and class III PDZ domains a [D/E]XW[C/S]-COOH consensus motif (Table 4). Additional consensus binding motifs could be identified, like the class IV motifs comprising a XXΨ[D/E]-COOH[127] consensus. However the latter was only identified by phage display and has to be confirmed *in vivo* (Vaccaro & Dente, 2002).

Canonical binding of a consensus binding motif to a PDZ domain will now be described using the crystal structure of PSD95 PDZ3 complexed to a class I peptide originating from the protein CRIPT[128]. As mentioned already, the C-terminal peptide binds into an extended groove on the surface of the PDZ domain. This groove is formed by three structural elements: 1) the carboxylate binding loop situated between βA and the βB strand, 2) the βB strand and 3) the αB helix (Figure 31, 32A). The βB strand forms hydrogen bonds between the amide nitrogen of V_0 and the carbonyl of F325, between the carbonyl oxygen of T_{-2} and the amide nitrogen I327 and between the amide nitrogen of T_{-2} and the carbonyl oxygen of I327. These interactions stabilise the peptide in its extended conformation in the binding groove and increases global affinity but do not account for specificity (Figure 32A,C). Specificity is conferred by several interactions at distinct positions:

---

[126] Residues are typed in one letter code, with X standing for any residue, ɸ for a hydrophobic residue and COOH for the free carboxy terminus

[127] Ψ stands for a polar residue

[128] CRIPT: Cysteine Rich Interactor of PDZ Three (Niethammer et al., 1998)

Class	C-terminus of the partner protein	Interacting partner	Example of PDZ domain
Class I X[S/T]Xφ	-EDTV -ESDV -TTRV -TSVF -ESLV -QSAV -DSSL -DTRL -QTRL -SSTL -PTRL	-Shaker K^+ channel -NMDA receptor subunits NR2 A/B -neuroligin -ILR-5α -PMCA4b -voltage-gated Na+ channel -protein kinase C-α -β_2-adrenergic receptor -CFTR -GKAP -metabotropic glutamate receptor subunit mGluR5 -GRK6A	-PSD95 (PDZ2) -PSD95 (PDZ2) -PSD95 (PDZ3) -syntenin (PDZ1, 2) -PSD95 (PDZ1, 2, 3) -syntropin -PICK -NHERF (PDZ1) -NHERF (PDZ1) -Shank/ProSAP -Shank/ProSAP
Class II XφXφ	-SVKI -EYFI -EYYV -EFYA -YYKV -DVPV	-AMPA receptor subunit GluR2 -glycophorin C -neurexin -syndecan-2 -ephrin B1 -receptor tyrosine kinase ErbB2	-PICK2, GRIP (PDZ5) -erythrocyte p55 -CASK -CASK, syntenin (PDZ2) -PICK1, GRIP (PDZ6), syntenin (PDZ2) -erbin
Class III XXXC	-DHWC -YXC	-N-type Ca^{2+} channel -L6 antigen	-Mint1 -SITAC

Table 4.: Classification of PDZ domains based on the C-terminal sequence of their binding partners.

First, the carboxylate binding loop comprising residues 318 to 324 of PSD95 PDZ3 and more precisely the GLGF motif (residues 322 to 325) form an environment of amide nitrogens which bind with the C-terminal carboxylate group of the peptide. In addition, R318 interacts with the carboxylate via a highly ordered water molecule (Figure 32C,D). Thus one important feature for peptide recognition by the PDZ domain is the stabilisation of the C-terminal carboxylate group. This is obvious when looking at the sequence alignment where a positive charge at position 318 as well as the GLGF-motif in the carboxylate binding loop are highly conserved among PDZ domains (Figure 31). In addition, this tight binding of the carboxylate group by the carboxylate binding loop pulls the V (P_0) residue in a hydrophobic pocket built by L323, F325, I327 and L379 (Figures 32B,C). This explains the high degree of conservation of a hydrophobic residue at the P_0 position across all classes of C-terminal peptides. Among the PDZ domains these hydrophobic residues building up the hydrophobic pocket are highly conserved and it is suggested, that variation of the residues forming the pocket and the C-terminal residue of the peptide could add selectivity by varying the volume of the pocket according to the steric volume of the C-terminal residue.

Second, T (P_{-2}) forms a hydrogen bond with its hydroxyl oxygen and the N-3 nitrogen of H372 (Figure 32C,D). An additional hydrogen bond between the N-1 nitrogen of H372 with the carbonyl oxygen of G329 stabilises this interaction and positions the side chain of H372 well. This strongly suggest that the P_{-2} position in a C-terminal binding motif is critical for interaction. This is furthermore supported by the fact that this conserved histidine residue in class I PDZ interactions is replaced by a leucine or methionine residue in the class II PDZs to interact with the hydrophobic residue in the P_{-2} position and by a charged residue in the class III PDZs, suggesting, that it is not only implicated into binding but also in some way into specificity of the interaction.

Third Q (P_{-3}) forms a hydrogen bond via its side chain oxygen with N326 from βB and with S339 from βC thus adding an additional constraint for specific peptide recognition (Figure 32C,D).

In this specific interaction no functional important interaction could be detected for the residue at the P_{-1} position.
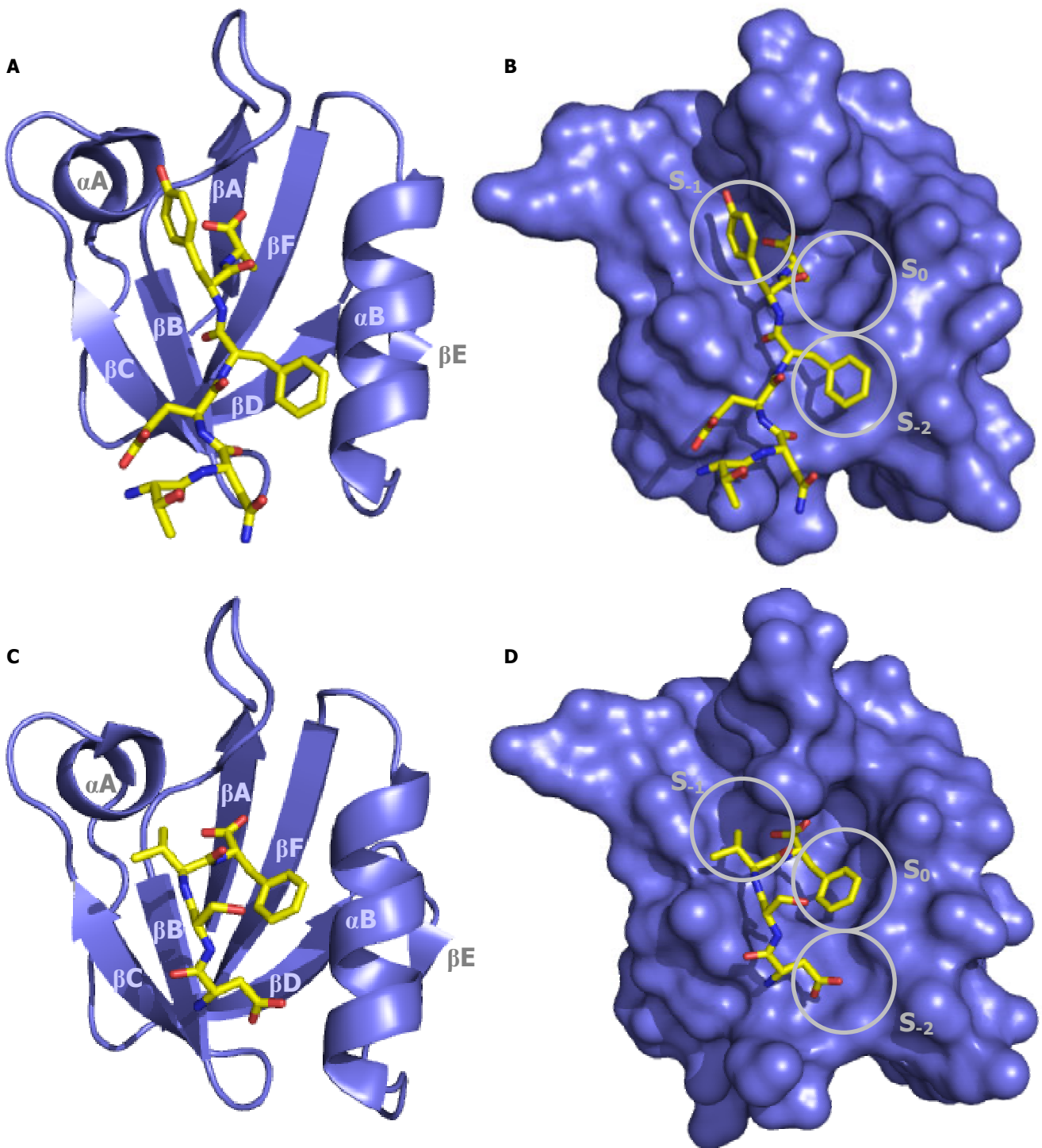
Figure 33.: Structural basis of peptide recognition according to the „pocket model" taking syntenin PDZ2 as an example. A/ Ribbon representation of syntenin PDZ2 bound to the syndecan-4 peptide (NEFYA) (PDB code 1OBY). **B/** Solvent accessible surface representation of the same complex as in A illustrating the peptide binding groove. The three binding pockets are circled and labelled S_0, S_{-1} and S_{-2}. The side chains of Y (P_{-1}) and F (P_{-2}) occupy the S_{-1} and S_{-2} pockets, whereas A (P_0) only occupies a part of S_0. **C/** Ribbon representation of syntenin PDZ2 bound to the IL5R peptide (DSVF) (PDB code 1OBZ). **D/** Solvent accessible surface representation of the same complex as in C illustrating the peptide binding groove. The three binding pockets are circled and labelled S_0, S_{-1} and S_{-2}. The side chains of F (P_0) and V (P_{-1}) occupy the S_0 and S_{-1} pockets, whereas S (P_{-2}) only occupy a part of S_{-2}. Figures were adapted from (Kang et al., 2003), and made using PyMol (DeLano Scientific).

However, this classical binding model implicating residues at the P_0 and P_{-2} and possibly at the P_{-3} position and the classification of C-terminal consensus binding motifs according to the chemical properties of the amino acids at these positions seems oversimplified, and not suited for the proper description of an increasing number of PDZ mediated interactions which do not follow this canonical type of recognition. The class III and IV peptides for example seem to recognise P_{-1} instead of P_{-2}. In addition some PDZ domains are able to recognise more than one class of C-terminal consensus binding motifs. One example should be cited explicitly, which is the interaction of the PDZ2 of syntenin with the class II C-terminal peptide TNEFYA of syndecan-4 and with the class I peptide ETLEDSVF of the interleukin 5 receptor α subunit (IL5Rα) (Figure 33). The syndecan-4 peptide was expected to engage a regular class II interaction by β-extension with A (P_0) and F (P_{-2}) interacting in a canonical way at the S_0 and S_{-2} pockets, similar to the above mentioned binding pattern for PSD95 PDZ3 and the C-terminal CRIPT peptide. However structural analysis of this complex revealed that the small A (P_0) did not occupy totally the S_0 pocket and that in contrast there was an additional interaction involving Y (P_{-1}), lodged in the S_{-1} pocket (Figure 33A,B) (Kang et al., 2003). Interestingly, other syntenin PDZ2 binding proteins such as neurexin I (class II) or neuroglian (class I) display also a Y (P_{-1}) residue. This indicates that the P_{-1} position may confer additional determinant for binding affinity to the initially described canonical PDZ binding model, thus giving a greater flexibility and enabling the interaction with peptides originating from different classes (Grootjans et al., 2000; Koroll et al., 2001). Indeed, the IL5Rα peptide was originally reported as a class I interaction but structural analysis revealed considerable differences. First the H residue at the beginning of the αB helix, which is conserved in most class I interactions is missing in syntenin PDZ2. Furthermore S (P_{-2}) does only indirectly interact with I212 by a hydrogen bond through a water molecule and thus does not interact directly with the S_{-2} pocket as for a classical class I interaction (Figure 33C,D). On the other hand V (P_{-1}) fits into the S_{-1} pocket. These two examples were meant to illustrate the flexibility of PDZ domains with respect to their ligand peptides. A more general model was proposed in contrast to the rigid classification of PDZ domains which is essentially based on the residues at the P_0 and P_{-2} positions. Kang et al. propose a combinatorial model based on binding pockets on the PDZ domain including the P_{-1}
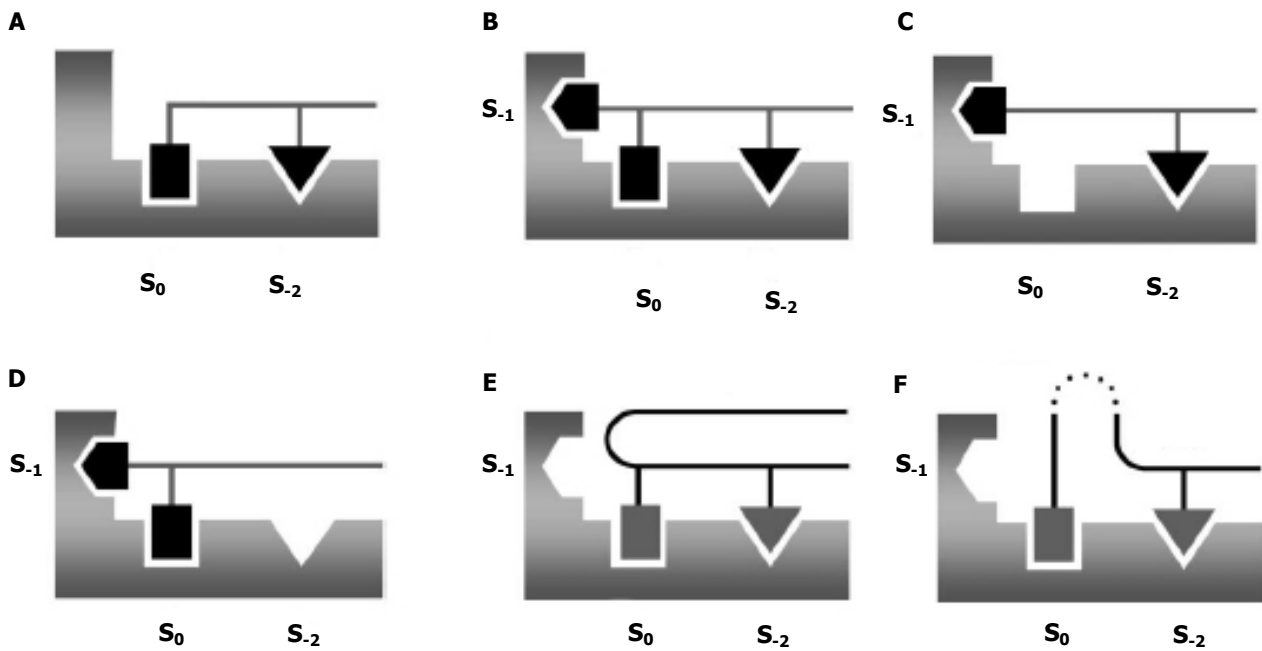
Figure 34.: Schematic representation of different possible PDZ interactions based on the „pocket model". **A/** Canonical PDZ binding model, only depending on the P_0 and P_{-2} position of the C-terminal peptide. **B/** PDZ binding model implicating all three C-terminal residues. **C/** Binding model where the C-terminal binding depends on the fitting of the P_{-1} and P_{-2} residues into the corresponding pockets. **D/** Binding model where the interaction depends on the binding of P_0 and P_{-1} at S_0 and S_{-1}, respectively. **E/** Interaction of an internal binding motif presented as a â-finger and implicating the positions P_0 and P_{-2}, like in a canonical interaction. This interaction mode is seen in the binding of a syntropin PDZ with nNOS (Hillier et al., 1999). **F/** Interaction of internal residues at pocket S_{-2} while C-terminal residues binds at S_0. This interaction mode is seen in the binding of syntenin PDZ2-PDZ2 interactions (Kang et al., 2003). Figure adapted from (Kang et al., 2003).

and in some cases the P_{-3} position (Kang et al., 2003). This model is more flexible and would allow to explain dual specificities for some PDZ domains. Figure 34 depicts the different possible binding modes to PDZ domains according to this combinatorial model.

In addition to modulable binding pockets concerning residues in the positions P_0 to P_{-2}, there is also growing evidence for the involvement of additional residues upstream of the last four C-terminal residues, which seem to be implicated into the recognition process and ligand binding affinity, adding even more flexibility to peptide recognition mechanism than the initial canonical model does. Specificity of the P_{-3} residue which often interacts across the βB strand with residues on the βC strand has been reported (Doyle et al., 1996). Some PDZ domains which have long βB strands or βB-βC loops, seem to have further interactions to their targets at this region (Birrane et al., 2003; Cai et al., 2002; Dobrosotskaya et al., 2001; Kozlov et al., 2002; Walma et al., 2002). As a concluding remark one can say, that indeed PDZ domains bind to C-terminal peptides which can be classified into different classes according to consensus binding sequences including the last four C-terminal residues. This classification and the initial canonical binding mechanism are very useful for a first evaluation and rapid classification of a given PDZ-peptide interaction. However, in reality PDZ domain-peptide interactions are much more complex and show a high degree of flexibility concerning modular binding pockets in the binding groove of the PDZ domain and by including additional residues on the peptide upstream to the "consensus motif". Therefore, predictions of a given interaction remain difficult and error prone. For detailed investigations of individual PDZ-peptide complexes structural analysis remains the only valuable tool for understanding the mechanism at an atomic scale.

Comparison with PTB domains

There is another group of protein-peptide interaction domains, the PTB[129] domains, which share local similarities in ligand binding. PTB domains are small domains of about 100-150 amino acids. They can be found in the insulin receptor substrates 1 and 2 or in the adaptor protein Shc[130] (Kavanaugh & Williams, 1994).

---

[129] PTB: Phosphotyrosine Binding domain, also known as PID (Protein Interaction Domain)
[130] Shc: Src homology 2 domain containing protein
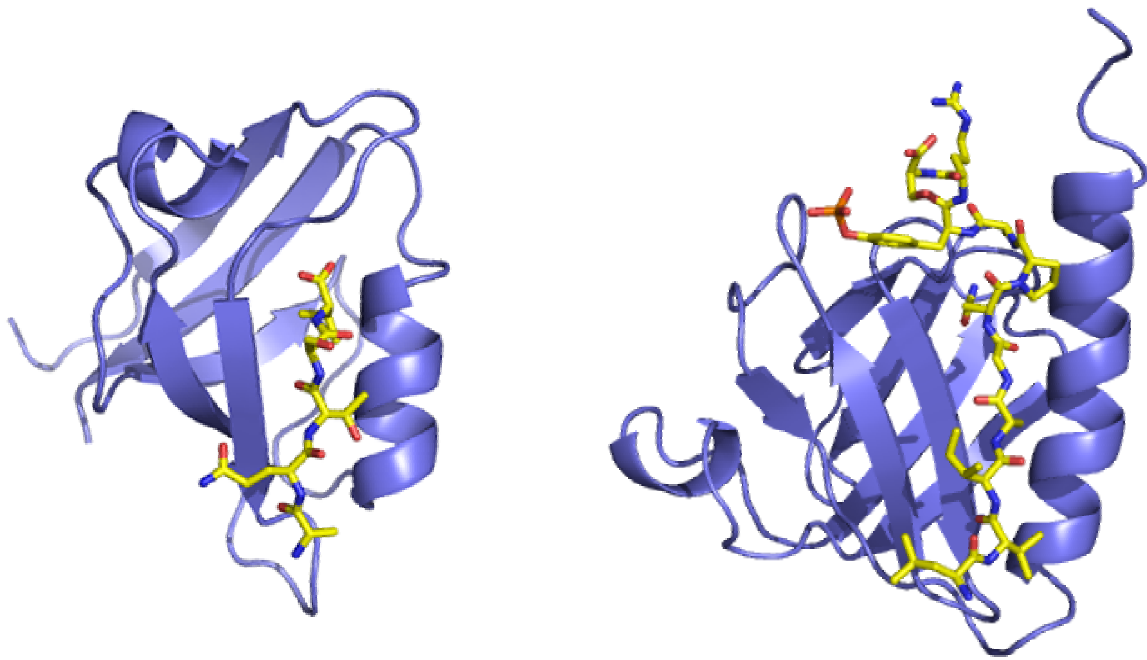
Figure 35.: Schematic comparison of PDZ and PTB domains. PDZ3 of PSD95 binding the C-terminal peptide from CRIPT is shown as a cartoon representation to the left (Doyle et al., 1996; PDB code: 1be9). PTB domain of the insuline receptor 1 substrate binding a peptide from the interleucin-4 receptor is shown as a cartoon representation to the right (Zhou et al., 1996; PDB code: 1IRS). The bound peptides are coloured in yellow and represented as sticks. Both peptides bind in a groove between a â-strand and a á-helix by â extension. The segment bound to the PTB domain adopts a strand turn conformation. The phosphytyrosine appears to be the last ordered residue in most of the PTB-peptide interactions and needs not to be at the C-terminus of the bound peptide. The peptide bound to the PDZ domain is a simple C-terminal strand.

PTB domains show a seven-stranded β-sandwich framework which is capped by C-terminal helices (Eck et al., 1996; Zhou et al., 1995; Zhou et al., 1996). PTB domains recognise peptides with a phosphotyrosine at the end of a NPXpY sequence, which in contrast to PDZ domains, has not to present necessarily a free C-terminus (Wolf et al., 1995). As in PDZ domains the complexes form via antiparallel β-extension of the peptide with one of the β-sheets of the PTB domain. On the other side, also similar to PDZ domains, the peptide packs against an α-helix (Figure 35). However, the NPXpY motif is much more extensive than the simple carboxylate group at the C-terminus of PDZ bound peptides and the polypeptide chain can continue beyond the phosphotyrosine. Usually, PTB NPXpY motifs are presented as a β-turn which is anchored by a complex network of hydrogen bonds in the PTB binding groove. Differential specificities seem to depend on the presence of pockets for hydrophobic residues at specific positions (DiNitto & Lambright, 2006; Harrison, 1996).

Regulation of PDZ-peptide interactions by phosphorylation

In C-terminal PDZ binding motifs there is often a phosphorylable residue at position P_{-2} or P_{-1} such as T, Y or S (Table 4). For example the S (P_{-2}) at the C-terminus of Kir2.3[131] is situated within a consensus sequence for PKA. Phosphorylation of S (P_{-2}) by PKA abolishes Kir2.3 interaction with PDZ domains of PSD95 (Cohen et al., 1996). Phosphorylation of S (P_{-3}) of the C-terminus of the AMPA[132] receptor subunit GluR2 by PKC prevents binding of GluR2 to the PDZ domain of GRIP (Matsuda et al., 1999). Phosphorylation of T (P_{-2}) of the C-terminus of high-risk HPV E6 oncoprotein by PKA shows inhibition of binding of E6 to the PDZ domains of hDLG and its degradation (Kuhne et al., 2000). Therefore, phosphorylation of residues near the C-terminus is likely to be a general mechanism to regulate PDZ interactions. These modifications may be either carried out by common protein kinases, such as PKA or PKC, or by orphan kinases which are specialised in phosphorylation of C-terminal sequences.

---

[131] Kir2.3: inward rectifier K$^+$ channel
[132] AMPA: α-Amino-5-hydroxy-3-Methyl-4-isoxazole Propionic Acid

Affinity of PDZ-peptide interactions in general

PDZ-peptide interactions fit to a simple one-to-one binding model as represented by equation 1),

$$1) \quad A + B \underset{k_{off}}{\overset{k_{on}}{\rightleftarrows}} AB$$
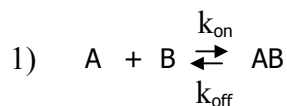
where A and B are each one of the free interaction partners, AB the complex and k_{on} and k_{off} the kinetic association and dissociation rate constants[133], respectively. The affinity of a complex can be expressed as the dissociation rate at equilibrium, which is nothing else than the ratio k_{off}/k_{on}. K_D[134] measurements of various PDZs interacting with C-terminal peptides have shown a wide range going from 10^{-8} M to 10^{-6} M. Binding affinities of 42 nM and 154 nM were measured by Biacore[135] for mDLG[136] PDZ2 and PTP1E/PTPbas-3 PDZ3 with their peptidic ligands, respectively (Songyang et al., 1997). By comparison, apparent affinity of 10^{-8} M was obtained by ELISA for the interaction of SAP102 PDZ2 with the C-terminal peptide of its natural ligand (Muller et al., 1996). Affinity of the interaction of PSD95 PDZ3 with the C-terminal peptide of CRIPT was measured by fluorescence and estimated to be 1 μM (Niethammer et al., 1998). This variability could partly be explained by the real differences according optimal binding affinity between various PDZs and also by the different affinities of peptide ligands for one PDZ domain. Another source of variation could be the diversity of *in vitro* binding assays used to determine K_Ds of PDZ domains. However, besides these precise examples more general studies showed that in general K_D values are in the low micromolar range (1-20 μM), suggesting that transient interactions may facilitate reversible assembly and thus a certain flexibility in ligand binding (Funke et al., 2005).

---

[133] kon: kinetic association rate constant [$M^{-1}s^{-1}$]; koff: kinetic dissociation rate constant [s^{-1}]

[134] K_D: dissociation constant at equilibrium of a complex between two entities, which is measured in molar [M].

[135] Biacore: The Biacore technology is a label-free surface plasmon resonance (SPR) based technology for studying biomolecular interactions in real time.

[136] mDLG: murine Disc Large

Introduction to the research project

In this scientific context my Ph.D. thesis aimed generally speaking in studying the structure-function relationship of the HPV 16 E6 oncoprotein (16E6). More precisely I studied the structural behaviour and the binding specificity of the interaction between 16E6 and the PDZ1 domain of MAGI-1 as well as the kinetic interaction behaviour between MAGI-1 PDZ1 and peptides originating from a subset of high-risk HPV E6 oncoproteins. In a first time, I participated in the setup of a production protocol for unlabelled as well as isotopically labelled HPV E6-C domain. These samples allowed Dr. Yves Nominé to perform a biophysical investigation on this domain, to assign the NMR proton frequencies and to calculate the 3-D structure. In a second time, I cloned MAGI-1 PDZ1 and set up a production protocol for producing unlabelled and isotopically labelled PDZ1 domain. In close collaboration with Dr. Andrew Atkinson, Prof. Bruno Kieffer as well as with my Ph.D. supervisor Dr. Gilles Travé, I assigned the NMR proton frequencies of the free PDZ domain. NMR titration experiments using isotopically labelled MAGI-1 PDZ1 and unlabelled E6-C or C-terminal peptides originating from E6-C allowed to identify all essential amino-acids implicated into ligand binding affinity and specificity.

In a parallel approach to the structural study of the interaction I performed a kinetic study of the interaction. To do so I collaborated with the Dr. Katia Zanier to set up and optimise experimental conditions for a peptide-protein interaction assay using the Biacore technology. I further optimised this method with the Dr. Yves Nominé in order to make it "medium throughput" compatible. Complementary to that, I set up a new method, baptised "holdup assay". The holdup assay is based on chromatographic retention and allows to study protein-protein interactions at equilibrium, giving thus access to qualitative binding data as well as to equilibrium dissociation constants. The great advantage of this method is that it requires only very basic laboratory equipment and that the experiments are easy and fast to perform.

I applied these two methods and studied the kinetic interaction behaviour of MAGI-1 PDZ1 against a subset of C-terminal peptides originating from different high-risk HPV viruses. I measured the KD values and together with the NMR experiments it was possible to identify key residues in the minimal C-terminal binding motif, which are important for ligand binding affinity and specificity.

RESULTS

The results presented in this section are organised in four parts. The first part concerns the production of monodisperse and active HPV16 E6-C domain suitable for *in vitro* assays and NMR studies. The second part will present in detail two methods for analysing protein-protein interactions. One method is based on protein-protein mediated selective chromatographic retention and provides a fast and easy tool for assaying or screening interactions at equilibrium. It is especially suited for low affinity complexes. The other method is based on the Biacore technology and allows medium throughput screening of protein-protein interactions, giving access to the kinetic binding parameters. The third part describes the production of MAGI-1 PDZ1 as well as the structural and kinetic analysis of the interaction of high-risk HPV E6 oncoproteins with this PDZ domain and more precisely, on a structural basis, the binding specificity of the HPV16 E6-C domain with MAGI-1 PDZ1 by NMR.

Part I

Expression, isotopic labelling and purification of HPV16 E6-C as an active and folded entity.

The strategies adopted to set up standard protocols for producing unlabelled or isotopically labelled samples of 16E6-C are displayed in 2 publications:

-Publication 1:

Domain substructure of HPV E6 oncoprotein: biophysical characterization of the E6 C-terminal DNA-binding domain. Nomine Y, Charbonnier S, Ristriani T, Stier G, Masson M, Cavusoglu N, Van Dorsselaer A, Weiss E, Kieffer B, Trave G. (2003) *Biochemistry* **42**(17):4909-4917

-Publication 2:

1H and 15N resonance assignment, secondary structure and dynamic behaviour of the C-terminal domain of human papillomavirus oncoprotein E6. Nomine Y, Charbonnier S, Miguet L, Potier N, Van Dorsselaer A, Atkinson RA, Trave G, Kieffer B. (2005) *J Biomol NMR* **31**(2):129-141.

Publication 1:

Nomine Y, <u>Charbonnier S</u>, Ristriani T, Stier G, Masson M, Cavusoglu N, Van Dorsselaer A, Weiss E, Kieffer B, Trave G. (2003) Domain substructure of HPV E6 oncoprotein: biophysical characterization of the E6 C-terminal DNA-binding domain. *Biochemistry* **42**(17):4909-4917

In this publication we describe in detail the expression and purification protocol of E6-C (4C/4S), a stabilising clone of the C-terminal zinc binding domain of HPV16 E6. Previous research on optimising the production of full length E6 oncoprotein in bacteria showed that expression levels and protein quality were enhanced when the protein was fused to a carrier protein named MBP. However, the commercial pMal vector system allowing to produce MBP-fusion proteins, yields medium expression levels due to a pTac promoter, which revealed limiting for subsequent NMR measurements. In addition, protein expression below 27°C is not possible. The breakthrough for producing 16E6-C (4C/4S) came when we switched to a pET-derived vector system, called pETM-vector series, which had been developed at the European Molecular Biology Laboratory (EMBL, Heidelberg, Germany). These vectors present a series of improvements, which allowed bypassing the bottleneck of protein production. First of all, they provide a variety of different carrier proteins among these, the requested MBP carrier protein, under the control of a strong viral T7 promoter. This promoter allowed me to screen for optimal expression conditions and to decrease expression temperature from 27°C to 22°C, which revealed optimal for protein quality. Moreover, the strong T7 promoter yielded protein expression levels which were sufficient for subsequent NMR experiments. Second, a peptidic linker sequence between the carrier protein and the protein of interest, bearing a TEV protease cleavage site, is provided by this vector series. This cleavage site is exceptionally long (7 amino acids), and therefore proteolytic cleavage is very precise so that nearly no protein degradation is observed even over long incubation times. In addition, TEV protease is even active in the presence of low amount of commercial protease inhibitor cocktails, which allows proteolytic removal of the MBP moiety at an intermediate stage of purification, lowering the risk of proteolysis due to contaminating proteases. Finally, the pETM vector kit is provided with an expression vector for recombinant TEV protease, making this system a very cheap alternative to other commercial systems. All these factors together allowed me to produce high levels of recombinant 16E6-C domain. I set up a standard purification protocol described in detail in publication 1. The resulting high concentrated samples of monomeric and active 16E6-C allowed the first author to perform the biophysical characterisation of this protein domain.

Publication 2:

Nomine Y, Charbonnier S, Miguet L, Potier N, Van Dorsselaer A, Atkinson RA, Trave G, Kieffer B. 1H and 15N resonance assignment, secondary structure and dynamic behaviour of the C-terminal domain of human papillomavirus oncoprotein E6. (2005) *J Biomol NMR* **31**(2):129-141.

The standard protocol for producing 16E6-C (4C/4S) allowed structural analysis of the protein domain by NMR. However, homonuclear spectra revealed too complex when using unlabelled protein samples. ^{15}N-labelling of 16E6-C (4C/4S) was required to increase the resolution of the NMR spectra and therefore to make frequency assignment possible, which is the basis of all subsequent investigation and for structure calculations. For this purpose, bacterial expression cultures are grown in so called M9 minimal medium which provides only a ^{15}N labelled nitrogen source. Compared to rich media, like LB medium for example, M9 medium presents the advantage of the absolute control of medium composition. On the other hand, expression levels decrease as the medium is lacking many nutrients found in rich media. In case of 16E6-C (4C/4S) expression, protein quality decreased drastically and yields of folded protein were not sufficient for heteronuclear NMR experiments. I performed a systematic screening of expression conditions, monitoring temperature, expression time and zinc concentration in the medium, since E6-C is a zinc binding domain. The screening revealed that expression at 22°C during 6h is optimal. In addition, it appeared that zinc concentration is an important parameter and I observed that a concentration between 6–25 µM was optimal. Lower or higher concentration increased the amount of aggregated protein. These data allowed me to set up a standard production protocol for ^{15}N-labelled 16E6-C which is described in detail in publication 2.

It had been suggested until then, that E6-C was a zinc binding domain with a 1:1 binding stoichiometry. I took part to the experiments concerning the measurement of zinc content in 16E6-C (4C/4S). Using a mass spectrometry analysis, we could show that 16E6-C (4C/4S) binds one zinc atom per domain.

Domain Substructure of HPV E6 Oncoprotein: Biophysical Characterization of the E6 C-Terminal DNA-Binding Domain

Yves Nominé, **Sebastian Charbonnier,** Tutik Ristriani, Gunter Stier, Murielle Masson, Nukhet Cavusoglu, Alain Van Dorsselaer, Étienne Weiss, Bruno Kieffer, et Gilles Travé

Biochemistry, 2003, Vol. 42, Pages 4909 -4917

^{1}H and ^{15}N resonance assignment, secondary structure and dynamic behaviour of the C-terminal domain of human papillomavirus oncoprotein E6

Yves Nominé[a], Sebastian Charbonnier[a], Laurent Miguet[b], Noelle Potier[b], Alain Van Dorsselaer[b], R. Andrew Atkinson[c], Gilles Travé[a,*] & Bruno Kieffer[c,*]

[a]*Laboratoire d'Immunotechnologie, CNRS UMR 7100, École Supérieure de Biotechnologie de Strasbourg, 67400 Illkirch, France;* [b]*Laboratoire de Spectrométrie de Masse Bio-Organique, CNRS UMR 7509, Faculté de Chimie, 67087 Strasbourg, France;* [c] *Laboratoire de RMN, CNRS UMR 7104, École Supérieure de Biotechnologie de Strasbourg, 67400 Illkirch, France*

Abstract

E6 is a viral oncoprotein implicated in cervical cancers, produced by human papillomaviruses (HPVs). E6 contains two putative zinc-binding domains of about 75 residues each. The difficulty in producing recombinant E6 has long hindered the obtention of structural data. Recently, we described the expression and purification of E6-C 4C/4S, a stable, folded mutant of the C-terminal domain of HPV16 E6. Here, we have produced ^{15}N-labelled samples of E6-C 4C/4S for structural studies by NMR. We have assigned most ^{1}H and ^{15}N resonances and identified the elements of secondary structure of the domain. The domain displays an original α/β topology with roughly equal proportions of α-helix and β-sheet. The PDZ-binding region of E6, located at the extreme C-terminus of the domain, is in a random conformation. Mass spectrometry demonstrated the presence of one zinc ion per protein molecule. Kinetics of replacement of zinc by cadmium followed by ^{1}H,^{15}N-HSQC experiments revealed specific frequency changes for the zinc-binding cysteines and their immediate neighbours. NMR spectra were affected by severe line-broadening effects which seriously hindered the assignment work. Investigation of these effects by ^{15}N relaxation experiments showed that they are due to heterogeneous dynamic behaviour with μs–ms time scale motions occurring in localised regions of the monomeric domain.

Introduction

E6 is one of two oncoproteins produced by 'high-risk' genital human papillomaviruses (HPVs) responsible for cervical cancers (zur Hausen, 1991). E6 interacts with more than 20 cellular proteins (Dell and Gaston, 2001; Du et al., 2002; Filippova et al., 2002; Iftner et al., 2002; Kumar

et al., 2002), including tumour suppressor p53 (Werness et al., 1990), cellular ubiquitin ligase E6-AP (Huibregtse et al., 1993), and a family of proteins containing PDZ domains (Thomas et al., 2002). The interaction of E6 with its target proteins often leads to their cellular degradation. For example, E6-mediated degradation of p53 occurs via formation of a trimeric complex involving E6, p53 and the cellular ubiquitin ligase E6-AP (Scheffner et al., 1993). E6 is also a transcription modulator. In particular, it activates the

---

transcription of the gene encoding the retrotranscriptase activity of human telomerase (Gewin and Galloway, 2001; Oh et al., 2001; Veldman et al., 2001). Finally, E6 is a DNA-binding protein which specifically recognises four-way DNA junctions (Ristriani et al., 2000).

Although the first DNA sequences encoding E6 were described in 1985 (Schwarz et al., 1985), no structural data are yet available for the wild-type E6 protein. Two forms of HPV16 E6 are produced *in vivo*, with lengths of 151 and 158 residues, depending on which methionine is used as a start codon (Androphy et al., 1987). Sequence alignments of E6 proteins from numerous HPV subtypes suggested the presence of two zinc-binding motifs (Cole and Danos, 1987) (Figure 1). However, biochemical studies that have been reported so far do not agree on the number of zinc ions bound to the protein. A first study reported that full-length E6 lacking nine C-terminal residues binds two zinc ions whereas the N-terminal domain of E6 binds only one zinc ion (Lipari et al., 2001). However, it has since been suggested that the biologically active form of full-length E6 may contain only one zinc ion per molecule (Degenkolbe et al., 2003).

Recently, we designed and produced a stable, folded and monomeric form of the C-terminal domain of HPV16 E6, referred to henceforth as E6-C 4C/4S, in which four non-conserved cysteines were mutated into serines (Nominé et al., 2003). This domain encompasses residues 80–151 of the shorter transcript of full-length HPV16 E6. It interacts specifically with four-way DNA junctions (Ristriani et al., 2001; Nominé et al., 2003) and contains many residues important for p53 degradation (Crook et al., 1991; Pim et al.,

1994; Nakagawa et al., 1995; Liu et al., 1999; Ristriani et al., 2002). Furthermore, E6-C 4C/4S contains the C-terminal sequence responsible for PDZ-domain recognition (Pim et al., 2002). Here, we describe the production of ^{15}N-labelled E6-C 4C/4S, the near complete assignment of ^{1}H and ^{15}N resonances and the subsequent identification of the elements of secondary structure. We have established the stoichiometry of zinc binding and identified the four chelating residues. ^{15}N relaxation rates have been measured to characterise the heterogeneous nature of the backbone dynamics of E6-C 4C/4S.

Material and methods

M9 medium

1 l of M9 ^{15}N minimal medium contained 7.7 g $Na_2HPO_4 \cdot 2H_2O$, 3 g KH_2PO_4, 0.5 g NaCl, 0.5 g $^{15}NH_4Cl$, 8.3 mg $FeCl_3 \cdot 6H_2O$, 0.8 mg $ZnCl_2$, 0.1 mg $CuCl_2$, 0.1 mg $CoCl_2 \cdot 6 H_2O$, 0.1 mg H_3BO_3, 11.1 mg $MnCl_2 \cdot 2H_2O$, 0.04% glucose, 120 mg $MgSO_4$, 36.6 mg $CaCl_2$, 1 mg biotin, 1 mg thiamine. Under these standard conditions, the zinc concentration was 6 μM. This concentration was varied from 0 to 200 μM, depending on the specific experiment.

Screening of conditions for optimal folding

The E6-C 4C/4S construct was subcloned within the pETM-10 vector allowing expression with a N-terminal HisTag. The DNA construct was electroporated into BL21 DE3 *E. coli* cells. About
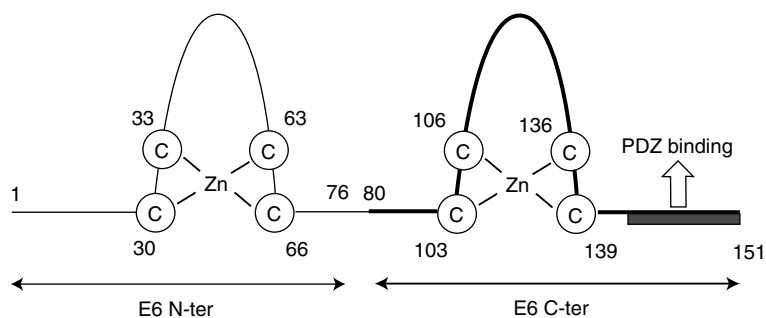
Figure 1. Schematic representation of the HPV E6 sequence with putative zinc binding sites deduced from the conservation of cysteine residues. The C-terminal domain encompasses residues 80–151 which are highlighted using a bold line.

1/200th of electroporated bacteria were transferred into M9 ^{15}N minimal medium containing 15 µg/ml kanamycin and were incubated for 3 days at 25 °C and 250 rpm. The pre-culture was diluted 40-fold in fresh medium and grown at 37 °C until OD (600 nm) reached 0.6–0.7. Cells were harvested by centrifugation at $2300 \times g$ and 25 °C during 10 min and transferred to fresh medium containing five possible zinc concentrations (0, 6, 25, 100 and 200 µM). IPTG (0.5 mM) was added for induction and the cultures were grown further at three possible temperatures (37 °C, 27 °C and 22 °C) for respectively 3, 6 or 18 h. Samples of 1 ml of expression culture corresponding to each combination of zinc concentration and expression temperature were pelleted at $1800 \times g$ in a benchtop centrifuge at 4 °C, re-suspended in 500 µl of buffer A (Tris-HCl 50 mM, NaCl 150 mM, DTT 2 mM, pH 6.8 containing 2.5 µg/ml DNase I, 2.5 µg/ml RNase I and eq. 1/100th tablet of complete EDTA-free protease inhibitor cocktail (Roche) in a 1.5 ml Eppendorf tube and sonicated with the 3 mM probe of a Vibracell 72412 sonicator (Bioblock Scientific) performing 2 runs of 15 s with a pulse frequency of 1 s pulse for 1 s pause and an amplitude of 6%. Fifty µl of these sonicated extracts were centrifuged at $16000 \times g$ for 15 m. The supernatant was taken up in 50 µl sample buffer and the pellet in 50 µl sample buffer and 50 µl H_2O. 20 µl samples of pellet and supernatant were tested by Tris–Tricine PAGE and the ratio of soluble to insoluble protein was determined with Quantity One™ quantification software (BIORAD) and validated by eye.

Production of ^{15}N-labelled E6-C 4C/4S

BL21 DE3 *E. coli* cells freshly electroporated with the MBP-E6-C 4C/4S expression construct (Nominé et al., 2003) were transferred directly into M9 ^{15}N minimal medium containing 15 µg/ml kanamycin and 6 µM ZnSO$_4$. The pre-culture and culture protocols were identical to those described for the screening. Cells from 2 l of expression culture were harvested by centrifugation at $2300 \times g$ and 4 °C during 20 min. To minimise oxidation problems, all buffers were degassed using a water vacuum pump and bubbled extensively with argon. The pellet was re-suspended in 150 ml of buffer A containing 5% glycerol, 1 µg/ml DNase I, 1 µg/ml RNase I

and three tablets of complete EDTA-free protease inhibitor cocktail. Cells were broken by sonication on ice, then centrifuged at $18000 \times g$ at 6 °C for 30 min. The supernatant was filtered (Millipore 0.22 µm) and loaded onto a 80 ml column of amylose resin (New England Biolabs) pre-equilibrated with buffer A. The column was washed stepwise with 1 volume, 3 volumes and 12 volumes of buffer A containing respectively 100%, 50%, and 12% of initial anti-protease concentration. Remarkably, MBP-E6-C 4C/4S eluted as a relatively pure form by leaking from the column in the late washing steps. The protein was incubated for 12–24 h at 6 °C with an appropriate concentration of recombinant TEV protease until full separation of E6-C 4C/4S from the MBP tag was achieved. The TEV cleavage site results in two additional residues (Gly–Ala) on the N-terminus of the construct, prior to the methionine residue. The digestion product was concentrated to a 1 ml volume using a 5 kDa-centriprep concentration device (Amicon) and applied on a Hiload 16/60 superdex 75 gel-filtration column (Amersham Biosciences) pre-equilibrated with buffer A. Pure monomeric E6-C 4C/4S peptide eluted as a single peak at the volume expected for a monomer according to the calibration of the column. The buffer was adjusted to 20 mM Tris-HCl, 50 mM NaCl, 1 mM DTT, pH 6.8 by performing dilution/concentration steps using a 15 ml Ultrafree Biomax 5K NMWL Membrane (Millipore). The final concentration was raised to 1.0–1.2 mM.

Mass spectrometry

Analyses were performed on a ESI-TOF mass spectrometer (LCT, Micromass U.K.). The sample buffer was exchanged by dilution/concentration steps against a 50 mM ammonium acetate buffer at pH 6.8, then diluted to 10 pmol/µl and continuously infused into the ESI ion source at a flow rate of 8 µl/min through a Harvard syringe pump. Parameters were optimised such that non-covalent interactions survive the ionisation desorption process. The accelerating voltage was set to 80 V and the pressure to 3 mbar. ESI-MS data were acquired in the positive ion mode, in the range 500–4000 m/z. The instrument was calibrated using multiply charged ions produced by a separate injection of horse heart myoglobin

diluted to 2 pmol/μl in a 1:1 water/acetonitrile mixture (v/v), acidified with 1% of formic acid.

NMR spectroscopy

Homonuclear ^{1}H spectra were acquired at 5 °C, 15 °C and 25 °C on a Bruker DRX600 spectrometer whereas heteronuclear experiments were recorded at 15 °C on a Bruker DRX500 spectrometer equipped with a z-gradient triple resonance cryoprobe. In all spectra, the water signal was suppressed using the WATERGATE sequence (Piotto et al., 1992). Proton assignments were based on homonuclear and ^{15}N-edited NOESY and TOCSY spectra. Heteronuclear ^{15}N-edited spectra were recorded with mixing times of 200 and 60 ms for the NOESY and TOCSY spectra, respectively. Data were processed using either NMRPipe (Delaglio et al., 1995) or XWINNMR (Bruker) and analysed with XEASY (Bartels et al., 1995). Dihedral ϕ angles were obtained from $^3J_{NH-H\alpha}$ scalar coupling constants determined from the ratio of diagonal to cross-peak intensities in a HNHA spectrum (Vuister and Bax, 1993). The ^{15}N resonance assignments of His118 and His126 rings and subsequently their protonation states were determined using a long-range $^1H,^{15}N$-HSQC spectrum (Pelton et al., 1993) in which the ^{15}N spectral width was set to 200 ppm. Two ^{15}N relaxation datasets were recorded for sample concentrations of 1.2 and 0.3 mM. Longitudinal (R_1) and transverse (R_2) ^{15}N relaxation rates together with heteronuclear $^1H-^{15}N$ NOE values were measured using proton-detected heteronuclear experiments (Kay et al., 1989). [15]N R_1 rates were obtained from 9 experiments with relaxation delays of 8, 40, 101, 243, 405, 568, 810, 1216 and 1824 ms and [15]N R_2 rates from 12 experiments with delays of 16, 32, 48, 72, 96, 119, 143, 183, 223, 263, 303 and 350 ms. During the R_2 relaxation delay, [15]N 180° pulses with a field strength of 2.0 kHz were applied every 1.2 ms ($\nu_{CPMG} = 0.83$ kHz) and ^1H 180° pulses were applied every 7.2 ms. For error estimation, the spectra with relaxation delays of 101 ms ([15]N R_1) and 72 ms ([15]N R_2) were recorded twice. In these sets of experiments, the recycle delay was 2 s. In ^{15}N, R_1 and heteronuclear $^1H-^{15}N$ NOE experiments, the protons were saturated using a train of 180° pulses separated by 4 ms. In the heteronuclear $^1H-^{15}N$ NOE experiment, the proton

magnetisation was saturated during 4 s to achieve the steady state. Data sets were typically recorded as 100×1024 complex points with spectral widths of 6.6 and 1.75 kHz in F1 and F2, respectively. The total acquisition time for one complete set of relaxation data measurements (R_1, R_2, $^1H-^{15}N$ NOE) was 2 days. Peak intensities were measured using Felix 2.10 (Accelrys Inc.) and exponential decay rates were obtained from a non-linear two-parameter least-squares fit using the Levenberg–Marquardt algorithm implemented in Matlab (The MathWorks Inc.). A set of 200 Monte Carlo simulations were used to estimate the statistical error on the parameters. Residues with ^{15}N resonances affected by exchange contributions were identified by recording relaxation-compensated CPMG dispersion experiments with a constant transverse relaxation time of 40 ms (Loria et al., 1999) but different values of the spacing delay (τ_{CP}) in the spin–echo pulse train. The extreme τ_{CP} delays were set to 20 and 0.2 ms corresponding to CPMG field strengths of 50 and 5000 Hz, respectively. The ^{15}N refocussing pulse length was set to 80 μs. Each 2D spectrum was recorded as 128×1024 complex points matrix with 24 scans per FID and a relaxation delay of 1.5 s resulting in a total acquisition time of 3 h per dispersion point. The presence of chemical exchange on the amide proton was derived from the ratios of peak intensities measured on a $^1H,^{15}N$-HSQC experiment and a $^1H,^{15}N$-CPMG-HSQC experiment in which both $^1J_{HN}$ transfer delays are replaced by a train of 16 refocussing pulses on both ^{15}N and ^1H nuclei (Mulder et al., 1996). Translational diffusion experiments were recorded using the LED pulse scheme as described previously (Nominé et al., 2003). To identify the zinc-binding residues, we substituted the initially bound zinc by cadmium by adding to the sample a solution of cadmium–EDTA, up to a final ratio of 3:1 (Cd:Zn), and following the displacement of amide resonances on $^1H,^{15}N$-HSQC spectra.

Results

Production of ^{15}N-labelled E6-C 4C/4S

Preliminary studies allowed us to express the C-terminal domain of E6 as a folded, monomeric

peptide (Nominé et al., 2003). The expression strategy was based on a fusion of the sequence of the C-terminal domain of E6 with maltose-binding protein (MBP) together with a systematic conversion of non-conserved cysteine residues into serines. Sequential assignment using homonuclear methods proved nonetheless to be extremely difficult due to spectral overlap and heterogeneous line-broadening. Initial attempts to produce a ^{15}N-labelled protein using M9 ^{15}N minimal media produced insoluble E6-C 4C/4S after proteolytic removal of the MBP carrier. We have shown previously that expression under improper conditions can lead to misfolded proteins, solubilised by the MBP carrier, forming high molecular weight soluble aggregates (Nominé et al., 2001). To avoid this solubilisation artefact, we performed a screen of expression conditions using a N-terminal HisTag fusion of E6-C 4C/4S. Three temperatures (22 °C, 27 °C and 37 °C) and five zinc concentrations (0, 6, 25, 100 and 200 μM) were probed and the solubility of the expressed material was analysed by pellet/supernatant assays. At all temperatures, we observed minimal solubility (30% of expressed material) in the absence of zinc, maximal solubility at low to intermediate zinc concentrations (6–25 μM), and lower solubility (40% of expressed material) at higher zinc concentrations (100–200 μM). The highest solubility (up to 70% of expressed material) was obtained for expression at 22 °C in the presence of 6 μM zinc. A minimum concentration of zinc in the expression medium is therefore required for proper folding of the domain, but an excess is detrimental and promotes aggregation, and this trend is reinforced upon raising the expression temperature. These conditions were subsequently applied for large scale production of the E6-C 4C/4S using the MBP version of the construct which provides higher purity and folding homogeneity as compared to the His-tagged version. Production of soluble ^{15}N-labelled E6C 4C/4S was successfully performed by expressing the His-MBP-E6C 4C/4S fusion at 22 °C using 2 l of M9 ^{15}N minimal medium with a zinc content of 6 μM. Purification and concentration of the product led to a ca. 1 mM protein sample that was suitable for NMR analysis.

^1H and ^{15}N resonance assignments, secondary structure and topology of E6-C 4C/4S

The ^1H,^{15}N-HSQC spectrum of ^{15}N-labelled E6-C 4C/4S exhibits good dispersion of the ^1H and ^{15}N resonances, as expected for a folded protein (Figure 2). However, both the line-shapes and the intensities of the cross-peaks display great variability, indicative of differential line-broadening along the peptide chain. This feature is striking when one compares, for example, the line-shapes of Gly130 and Gly134, the former displaying a much larger line-width in the ^{15}N dimension. Since this behaviour may originate from dynamic oligomerisation of the protein (Pfuhl et al., 1999), the oligomeric state was probed by measuring the translational diffusion coefficient (D_t) at two concentrations (0.3 and 1.2 mM) using a modified LED experiment (Dingley et al., 1995). The measured values were compatible with that expected for a monomer and were independent of the concentration, indicating that the observed line-broadening is due to μs–ms time-scale motions within the peptide chain (data not shown). Despite this line-broadening phenomenon, the quality of the ^{15}N-edited NOESY and TOCSY spectra (Figure 2) allowed near-complete assignment of the backbone and side-chain ^1H and ^{15}N nuclei. No spin system could be found for Thr86, probably due to severe line-broadening. The resonances of Asn127 show significant deviation from expected 'random coil' values (Hα: 4.28; Hβ: 1.83/-0.30; Hδ: 5.75/3.86 and Nδ: 107.0 ppm) for a reason which remains to be elucidated. To assign the histidine side-chains unambiguously we used long-range ^1H,^{15}N-HSQC spectra (Pelton et al., 1993). Analysis of $^{15}N\delta$ and $^{15}N\varepsilon$ chemical shifts indicated that the Nδ atom of His118 is protonated while His126 is in equilibrium between a fully and Nδ-protonated form, giving an estimated pKa value of 6.8 for this residue. ^1H and ^{15}N chemical shifts have been deposited in the BioMagResBank (Accession code 6407).

NOEs observed in homonuclear and ^{15}N-edited NOESY experiments allow the definition of three helical segments (Figure 3). For these segments, characteristic $H\alpha_i$-HN_{i+3} NOEs are supported by both low values of $^3J_{HN-H\alpha}$ coupling constants, deduced from the HNHA spectrum,
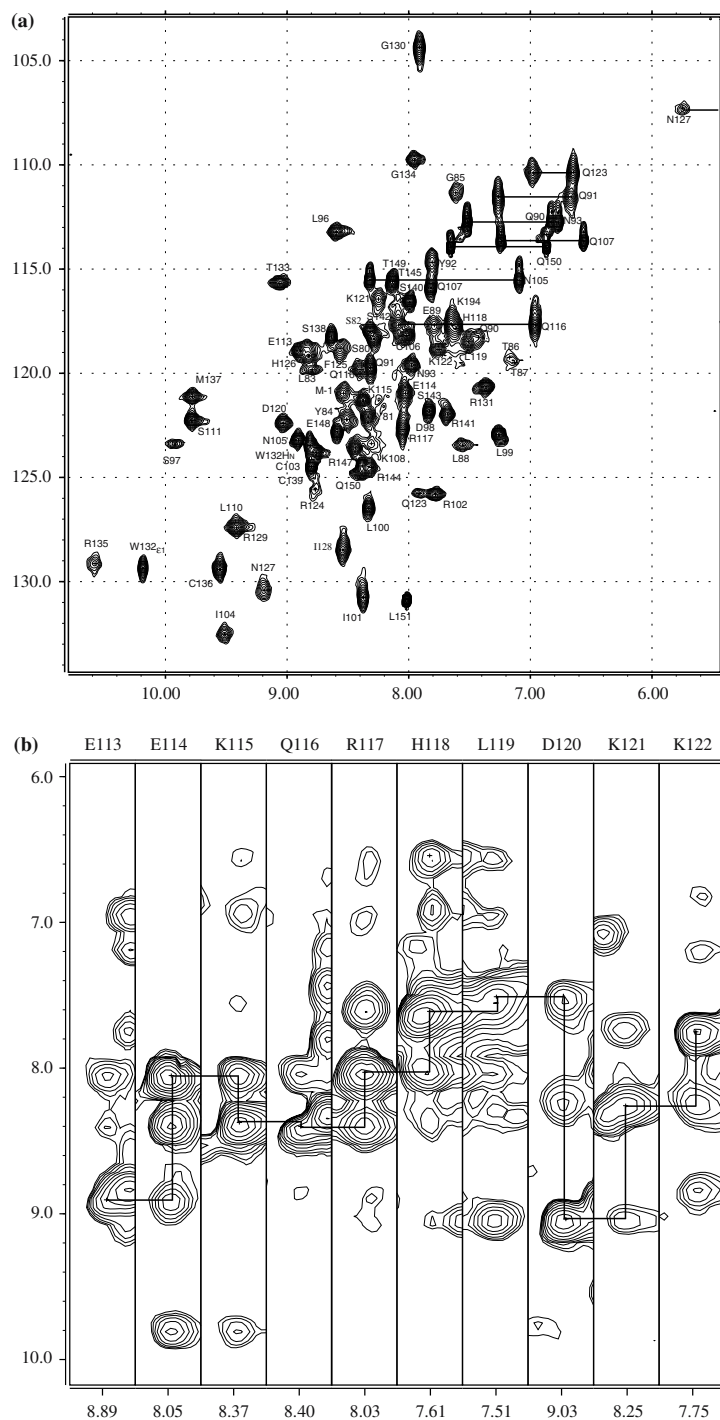
Figure 2. Representative spectra of the ^{15}N-labelled E6-C 4C/4S domain recorded at pH 6.8 and 15 °C. (a) 2D ^{1}H,^{15}N-HSQC spectrum. Cross-peaks are numbered according to the full-length HPV16 E6 (151-residue form). Side-chain NH_2 resonances of asparagine and glutamine residues are connected by horizontal bars, excepted for Asn127, for which only one proton appears in the region of the spectrum shown. (b) Regions of 3D ^{15}N-edited NOESY-HSQC spectrum (mixing time 200 ms) showing sequential connectivities for residues between Glu113 and Lys122.

and the chemical shift index (Wishart et al, 1992). In addition, three short β-strands were identified, encompassing residues 81–83, 125–127 and 131–133. Several inter-strand NOEs between amide protons belonging to residues Leu83–Phe125, His126–Thr133 and Ile128–Arg131 were observed together with NOEs between Hα protons of residues Asn127–Trp132 and Ser82–His126, indicating the presence of an anti-parallel β-sheet structure. This definition of secondary structure elements is in good agreement with circular dichroism results which estimated the α-helical and β-sheet content to be 22 ± 1% and 22 ± 3%, respectively (Nominé et al., 2003). 'Random coil' values for Hα chemical shifts and a lack of medium-range NOEs suggest that the last 10 residues are unfolded.

Zinc binding properties of E6-C 4C/4S

The zinc content of the E6-C 4C/4S domain was probed by ESI mass spectrometry. The difference in mass of 63.7 Da between experiments carried out under native and denaturing conditions (Table 1) is consistent with the binding of a single zinc ion. To further characterise the zinc-binding properties of E6-C 4C/4S and to identify those residues directly involved in zinc binding, a zinc/cadmium substitution experiment was monitored by $^1H,^{15}N$-HSQC measurements. Chemical shift perturbations upon replacement of zinc by cadmium were localised in two regions encompassing residues Cys103–Gln107 and Cys136–Arg141, identifying cysteines 103, 106, 136 and 139 as the residues involved in zinc coordination. The perturbation was found to be systematically higher for the amide group of the second cysteine of each pair, which could be explained by with the involvement of this amide proton in hydrogen bond with the sulphur atom of the preceding cysteine residue (Summers et al., 1990; Allen et al., 1997). The possibility of zinc coordination by the two histidine residues, His118 and His126, can be discounted by the lack of chemical shift perturbation of both backbone and side-chain ^1H and ^{15}N nuclei upon cadmium substitution.

The identification of the zinc-binding residues together with the assignment of inter β-strand NOEs provided enough constraints to establish the topology of E6-C 4C/4S (Figure 4). The three short β-strands form an anti-parallel β-sheet which, together with two α-helices between strands β1 and β2, define the limits of two 5-residue loops (L1 and L3) and a longer, 18-residue loop (L2) encompassing the first two zinc-coordinating cysteine residues. The C-terminal α-helix contains the third and fourth cysteine residues involved in zinc coordination and precedes an unfolded C-terminal tail of 10 residues.

Dynamic behaviour of E6-C 4C/4S

Structural analysis of E6-C 4C/4S was severely hindered by the presence of broadened resonances in NMR spectra. We therefore investigated the dynamic behaviour of E6-C 4C/4S via a set of ^{15}N relaxation experiments (Figure 5), with the aim of identifying possible sources of increased transverse relaxation rates. The profile of ^1H-^{15}N NOE values along the peptide sequence indicates that the peptide backbone is fairly rigid on the ps–ns time-scale from residue Ser82 to Ser140, with the exception of Lys108, localised within loop L2. Reduced values of the heteronuclear NOE for the 10 C-terminal residues reflect dynamic disorder for this portion of the peptide chain. The profile of ^{15}N transverse relaxation rates exhibits much larger variations along the sequence, indicative of conformational exchange contributions to these rates. This behaviour made the extraction of a global correlation time from the relaxation data rather difficult and hinders further interpretation of these rates in terms of molecular motions. In order to identify those residues not affected by exchange contributions and to further characterise these intermediate time-scale motions, dispersion experiments were conducted. A fast and convenient way of mapping exchange contributions to the transverse relaxation rates of amide protons is to compare peak intensities measured in $^1H,^{15}N$-HSQC and $^1H,^{15}N$-CPMG-HSQC spectra (van Tilborg et al., 2000; Palmer et al., 2001). The profile of the ratio of peak intensities shows a steady increase followed by a decrease for residues located in loops L1 and L3 (Figure 6). Other residues also display increased values such as Ser97 and Ile104, the latter located between the first two zinc-coordinating cysteine residues, and Leu110, sandwiched between two proline residues. From this profile, it is readily apparent that the line-broadening phenomena affecting the

136
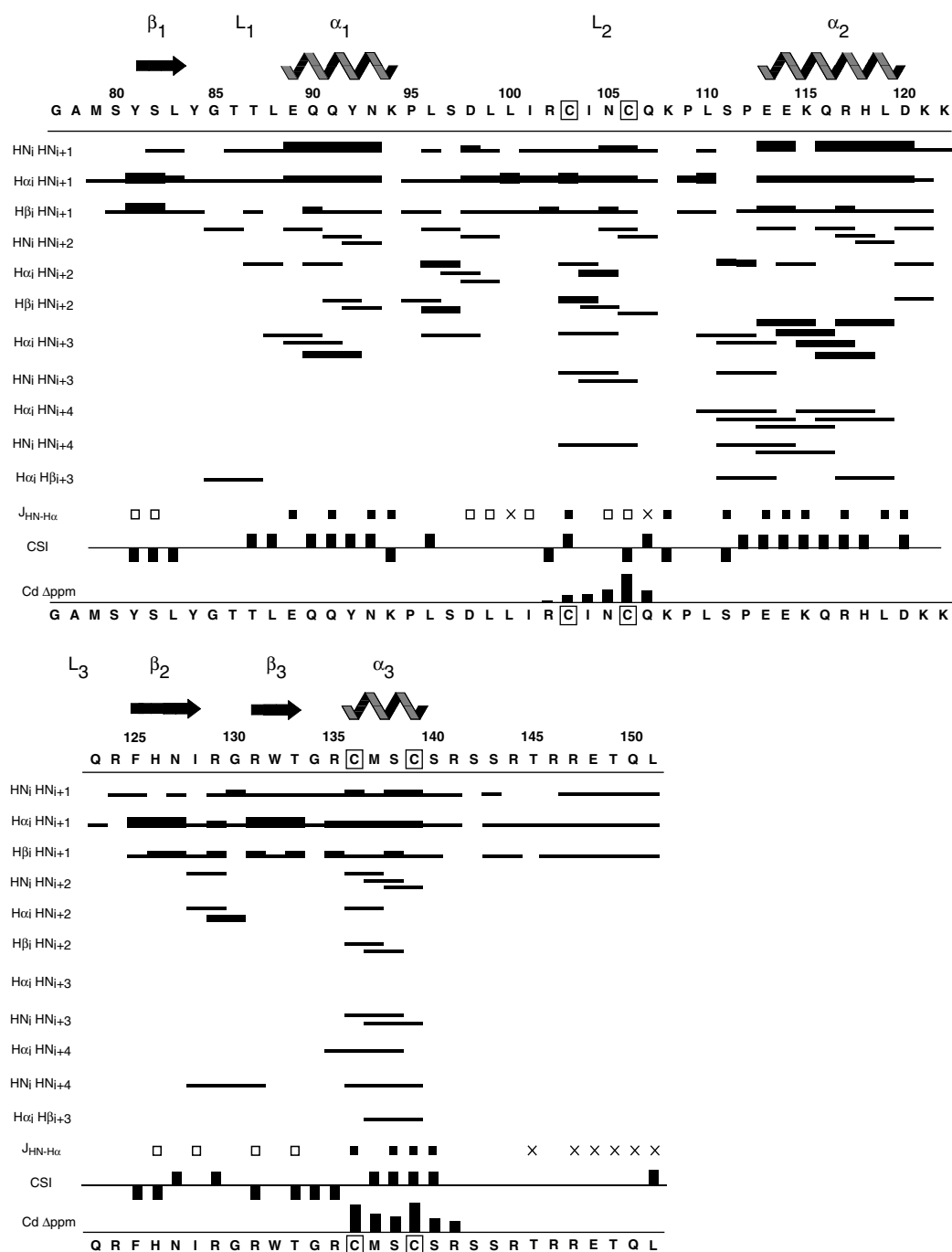
Figure 3. Secondary structure of E6-C 4C/4S deduced from short- and medium-range NOEs, $^3J_{HN–H\alpha}$ coupling constants and Hα chemical shift index (Wishart et al., 1992). The first two residues (77–78) originate from the TEV protease cleavage site. The presence of a NOE cross peak is indicated by a line with a thickness proportional to its intensity. $^3J_{HN–H\alpha}$ coupling constants greater than 6.5 Hz and lower than 7.5 Hz are indicated by closed and open squares, respectively. The lack of coupling information (due to overlapped or unobserved peaks) is indicated by a cross. CSI represents the H$_\alpha$ chemical shift index where 'random coil' values are taken from (Merutka et al., 1995). A vertical bar indicates an absolute difference between experimental and 'random coil' values greater than 0.2 ppm. The line Cd ?ppm reports the composite chemical shift differences for ^{1}H-^{15}N cross-peaks recorded for the zinc- and the cadmium- loaded proteins calculated using: $\Delta ppm = [\,5(\delta^H(Cd^{2+}) - \delta^H(Zn^{2+}))^2 + (\delta^N(Cd^{2+}) - \delta^N(Zn^{2+}))^2\,]^{1/2}$.

Table 1. Mass deduced from ESI mass spectrometry experiments The expected mass deduced from the sequence of E6-C 4C/4S is 9010.1 Da (assuming uniform ^{15}N-labelling)

	E6-C 4C/4S (Zn^{2+})	E6-C 4C/4S (Cd^{2+})
native	9073.4 ± 0.6 Da	9120.0 ± 1.0 Da
denaturated	9009.7 ± 0.1 Da	9010.1 ± 0.2 Da

NMR spectra of E6-C 4C/4S probably originate from several different motions. The profile of peak intensity ratios obtained from the extreme values of the delay between ^{15}N refocussing pulses ($\tau_{cp} = 20$ ms and $\tau_{cp} = 0.2$ ms) in a series of ^{15}N relaxation dispersion experiments (Figure 6) is very similar to that obtained from peak intensities measured in ^{1}H,^{15}N-HSQC and ^{1}H, ^{15}N-CPMG-HSQC spectra. Most residues affected by exchange contributions are localised in loops L1 and L3, β-strands β1 and β2 and α-helices α1 and α2 (see Figure 4). An analysis of these relaxation profiles allowed the identification of a small number of residues (97–103 excluding 101) whose ^{15}N transverse relaxation rates are not affected by exchange contributions. The R2/R1 ratio yielded estimates of the global rotational correlation time of 8.7 ± 1.5 ns and 7.8 ± 1.7 ns at high (1.2 mM) and low (0.3 mM) concentrations, respectively, compatible with the

domain behaving as a monomer (expected correlation time: 7.6 ns), and in agreement with translational diffusion data obtained for unlabelled E6-C 4C/4S.

Discussion

Over the past 20 years, a large number of publications has established the key role played by oncoprotein E6 in HPV-induced carcinogenesis (Rapp and Chen, 1998; Finzer et al., 2002). E6 activity is based on its ability to target numerous cellular targets (Mantovani and Banks, 2001). Thus E6 has been identified as an important subject for structural studies aimed at designing potential inhibitors of its various interactions. Such studies have long been hindered however, by the difficulty of producing well-behaved samples of either full-length E6 or its subdomains. In a recent publication, we reported the production of the C-terminal domain of E6 (E6-C) in a pure, monomeric form amenable to NMR studies (Nominé et al., 2003). Analysis of preliminary NMR spectra showed that several resonances of this domain were affected by significant line-broadening so that the production of an ^{15}N-labelled sample was required. However, the use of M9 ^{15}N minimal medium in place of LB
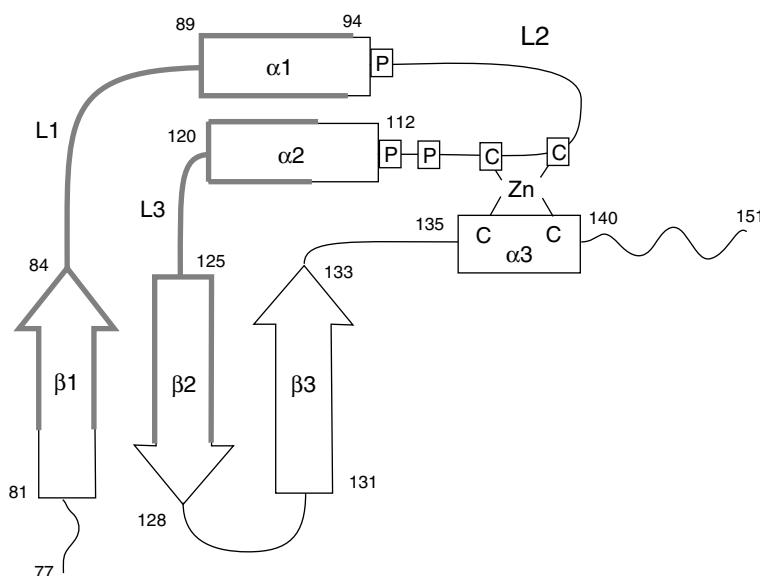
Figure 4. Topology of E6-C 4C/4S domain. Elements of secondary structures are shown as arrows and rectangles for β-strands and α-helices respectively. Regions of the peptide affected by chemical exchange are highlighted using a bold grey line.
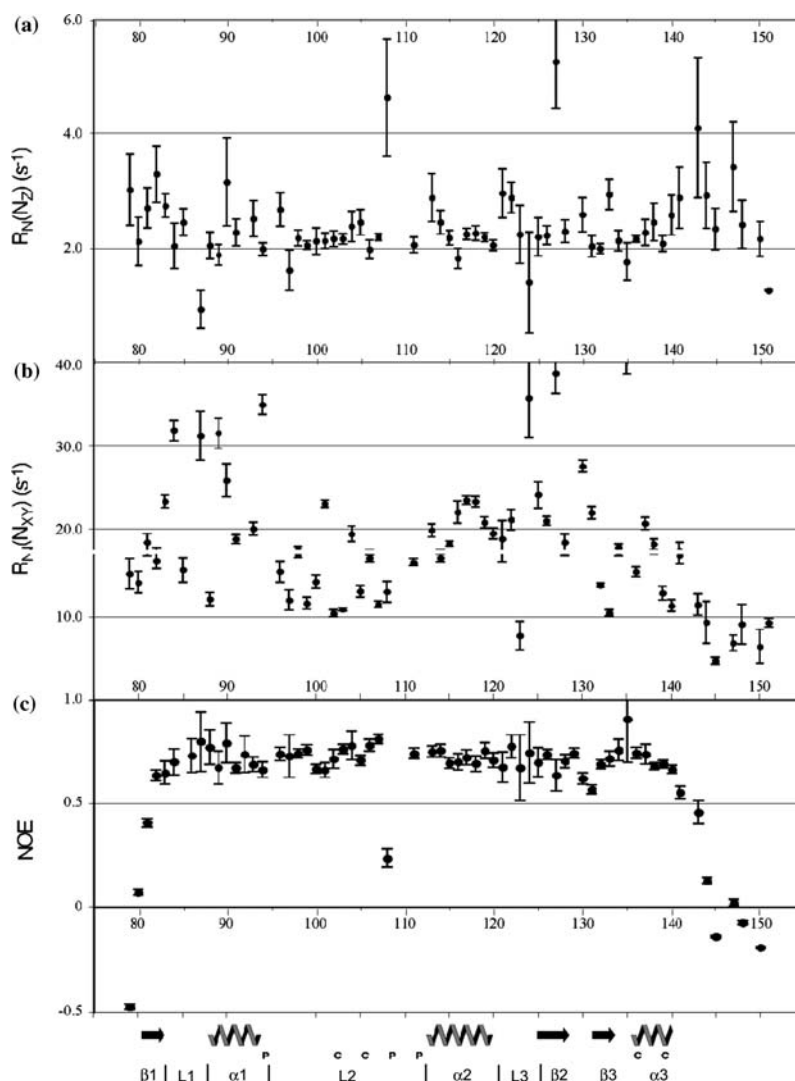
Figure 5. Longitudinal (a) and transverse (b) ^{15}N relaxation rates and steady state ^{1}H-^{15}N NOE values (c) for E6-C 4C/4S measured at 15 °C.

led unexpectedly to the production of misfolded E6-C 4C/4S samples. Only by careful screening of the expression conditions could the origin of this behaviour be identified: optimal solubility required fine tuning of the zinc content within the M9 ^{15}N expression medium, and a lower temperature of expression. This optimisation of zinc content and expression temperature by screening proved critical for the obtention of folded ^{15}N-labelled E6-C 4C/4S and subsequent assignment of resonances of the domain.

The strict requirement for zinc in the expression medium provided the first evidence of the

zinc binding to E6-C 4C/4S. Mass spectrometry showed that E6-C 4C/4S indeed contains a single zinc atom. Finally, zinc–cadmium substitution experiments identified the four zinc-binding cysteines of the domain. These data provide sound experimental confirmation of the earlier prediction of a zinc-binding site in this region of E6 based on sequence analysis of all HPV strains (Cole and Danos, 1987). The analysis of the secondary structure and the topology of the E6-C domain (Figure 4) suggests a structural role for the zinc ion which contributes to the stabilisation of a short C-terminal helix (α3) by anchoring it
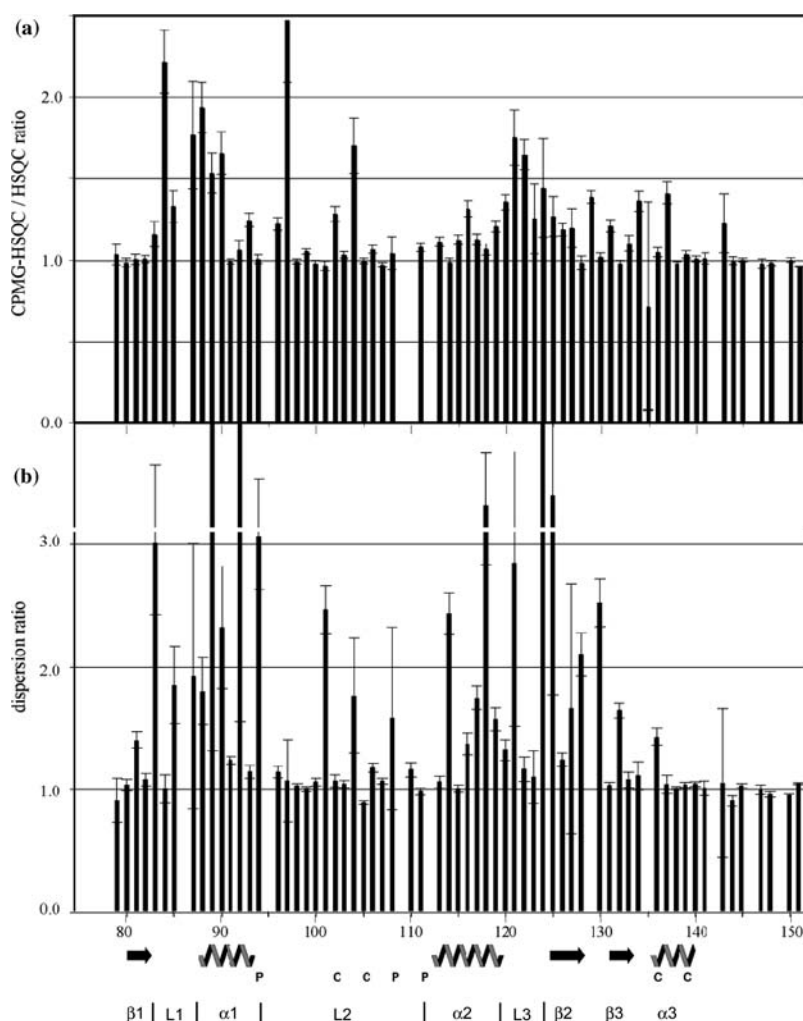
Figure 6. Identification of intermediate exchange contributions to transverse relaxation. (a) HSQC-CPMG/HSQC ratio as a function of sequence. (b) Ratio between relaxation-compensated CPMG dispersion experiments recorded with inter-pulse delays of 20 and 0.2 ms. In both cases, statistical errors were estimated from the standard deviation of the noise.

to the large L2 loop connecting helices α1 and α2.

Analysis of medium range NOEs (Figure 3) shows that the minimal folding unit begins with Tyr81 initiating the first β strand and ends at Arg141. These boundaries for the C-terminal domain of E6 are consistent with the domain definition based on results from mild proteolysis experiments (Lipari et al., 2001; Nominé et al., 2003). The strong sequence similarity between the N- and C-terminal domains suggests that the N-terminal domain should adopt the same topology. Both the lack of medium-range NOE and ^{15}N relaxation studies show that the last 10 C-terminal residues (Ser142–Leu151) are not

structured and rather flexible on the ps–ns time-scale. This region encompasses the PDZ domain-binding site consisting of the last four residues (sequence ETQL) and is likely to fold upon binding to its partner (Cowburn, 1997; Hung and Sheng, 2002).

A specific feature of E6-C 4C/4S is the presence of motions on the μs–ms time scale in several regions of the peptide (Figure 6). The L2 loop contains a number of isolated residues affected by exchange broadening which may arise from cis–trans isomerisation of the peptide bonds of proline residues in this loop. However, most residues affected by exchange broadening are clustered within two continuous sequence

elements (β1–L1–α1) and (α2–L3–β2) (see grey elements in Figure 4). This suggests that these two regions sense a common conformational exchange phenomenon which may indicate a spatial proximity of these two regions within the folded domain. A precise characterisation of time-scales (τ_{ex}), frequency differences and the populations of the exchanging states would require an extensive set of relaxation dispersion experiments (Palmer et al., 2001). Preliminary attempts to fit the relaxation dispersion dataset recorded at 500 MHz with the general equation of a two-state exchange system (Carver and Richards, 1972) led to an estimated time range of 0.2–3 ms and unequal populations of the states (data not shown). This indicated that the equilibrium is on the slow-intermediate time-scale ($\tau_{ex}.\Delta v > 1$) (where Δv is the chemical shift difference between the two states) and is consistent with our repeated observations that increased temperature induced the concerted disappearance of a subset of peaks in NMR spectra of E6-C 4C/4S. Indeed, an increase in temperature is expected to reduce the value of τ_{ex} according to Arrhenius' law and to shift $\tau_{ex}.\Delta v$ towards 1, leading to enhanced line-broadening (Mandel et al., 1996). Moreover, it was observed that the exchange contribution was independent of sample concentration, which excludes the possibility that this effect is due to fractional self-association of E6-C 4C/4S. Therefore, the occurrence of slow-intermediate exchange in specific regions of the primary sequence of E6-C 4C/4S may rather be related to a transition from a folded state toward a less populated locally unfolded state resulting from the lack of stabilising interactions, normally provided by the N-terminal domain. This would suggest that the two loops are involved in the domain interface, an hypothesis which is reinforced by the fact that two hydrophobic residues within L1 loop (Tyr84 and Leu88) are conserved among all HPV strains. These residues may thus prove to be potentially valuable mutation targets for stabilising the domain. Following this detailed characterisation of the dynamic behaviour of the C-terminal domain of E6, we are now in a position to interpret structural information in NMR spectra of E6-C 4C/4S and expect soon to be able to report on its three-dimensional structure.

References

Allen, F.H., Bird, C.M., Rowland, R.S. and Raithby, P.R. (1997) *Acta Crystallogr. Sect. B-Struct. Commun.*, **53**, 696–701.

Androphy, E.J., Hubbert, N.L., Schiller, J.T. and Lowy, D.R. (1987) *EMBO J.*, **6**, 989–992.

Bartels, C., Xia, T.H., Billeter, M., Güntert, P. and Wüthrich, K. (1995) *J.Biomol. NMR*, **5**, 1–10.

Carver, J.P. and Richards, R.E. (1972) *J. Magn. Reson.*, **6**, 89–105.

Cole, S.T. and Danos, O. (1987) *J. Mol. Biol.*, **193**, 599–608.

Cowburn, D. (1997) *Curr. Opin. Struct. Biol.*, **7**, 835–838.

Crook, T., Tidy, J.A. and Vousden, K.H. (1991) *Cell*, **67**, 547–556.

Degenkolbe, R., Gilligan, P., Gupta, S. and Bernard, H.U. (2003) *Biochemistry*, **42**, 3868–3873.

Delaglio, F., Grzesiek, S., Vuister, G.W., Zhu, G., Pfeifer, J. and Bax, A. (1995) *J. Biomol. NMR*, **6**, 277–293.

Dell, G. and Gaston, K. (2001) *Cell. Mol. Life Sci.*, **58**, 1923–1942.

Dingley, A.J., Mackay, J.P., Chapman, B.E., Morris, M.B., Kuchel, P.W., Hambly, B.D. and King, G.F. (1995) *J. Biomol. NMR*, **6**, 321–328.

Du, M., Fan, X., Hong, E. and Chen, J. (2002) *Biochem. Biophys. Res. Commun.*, **296**, 962–969.

Filippova, M., Song, H., Connolly, J., Dermody, T. and Duerksen-Hughes, P. (2002) *J. Biol. Chem.*, **277**, 21730–21739.

Finzer, P., Aguilar-Lemarroy, A., Rösl, F., Rapp, L., Chen, J.J., Mantovani, F. and Banks, L. (2002) *Cancer Lett.*, **188**, 15–24.

Gewin, L. and Galloway, D.A. (2001) *J. Virol.*, **75**, 7198–7201.

Huibregtse, J.M., Scheffner, M. and Howley, P.M. (1993) *Mol. Cell. Biol.*, **13**, 775–784.

Hung, A.Y. and Sheng, M. (2002) *J. Biol. Chem.*, **277**, 5699–5702.

Iftner, T., Elbel, M., Schopp, B., Hiller, T., Loizou, J., Caldecott, K. and Stubenrauch, F. (2002) *EMBO J.*, **21**, 4741–4748.

Kay, L.E., Torchia, D.A. and Bax, A. (1989) *Biochemistry*, **28**, 8972–8979.

Kumar, A., Zhao, Y., Meng, G., Zeng, M. Srinivasan, S., Delmolino, L., Gao Q., Dimri, G., Weber, G., Wazer, D., Band, H. and Band, V. (2002) *Mol. Cell. Biol.*, **22**, 5801–5812.

Lipari, F., McGibbon, G.A., Wardrop, E. and Cordingley, M.G. (2001) *Biochemistry*, **40**, 1196–1204.

Liu, Y., Chen J.J., Gao Q., Dalal S., Hong Y., Mansur C.P., Band V. and Androphy, E.J. (1999) *J. Virol.*, **73**, 7297–7307.

Loria, J.P., Rance M. and Palmer, A.G. (1999) *J. Am. Chem. Soc.*, **121**, 2331–2332.

Mandel, A.M., Akke M. and Palmer, A.G. (1996) *Biochemistry*, **35**, 16009–16023.

Mantovani, F. and Banks, L. (2001) *Oncogene*, **20**, 7874–7887.

Merutka, G., Dyson, H.J. and Wright, P.E. (1995) *J. Biomol. NMR*, **5**, 14–24.

Mulder, F.A., Spronk, C.A. E.M., Slijper, M., Kaptein, R. and Boelens, R. (1996) *J. Biomol. NMR*, **8**, 223–228.

Nakagawa, S., Watanabe, S., Yoshikawa, H., Taketani, Y., Yoshiike, K. and Kanda, T. (1995) *Virology*, **212**, 535–542.

Nominé, Y., Charbonnier, S., Ristriani, T., Stier, G., Masson, M., Cavusoglu, N., Van Dorsselaer, A., Weiss, E., Kieffer, B. and Travé, G. (2003) *Biochemistry*, **42**, 4909–4917.

Nominé, Y., Ristriani, T., Laurent, C., Lefèvre, J.F., Weiss, E. and Travé, G. (2001) *Protein. Expr. Purif.*, **23**, 22–32.

Oh, S.T., Kyo, S. and Laimins, L.A. (2001) *J. Virol.*, **75**, 5559–5566.

Palmer, 3rd, A.G., Kroenke, C.D. and Loria, J.P. (2001) *Meth. Enzymol.*, **339**, 204–238.

Pelton, J.G., Torchia, D.A., Meadow, N.D. and Roseman, S. (1993) *Protein Sci.*, **2**, 543–558.

Pfuhl, M., Chen, H.A., Kristensen, S.M. and Driscoll, P.C. (1999) *J Biomol NMR*, **14**, 307–320.

Pim, D., Storey, A., Thomas, M., Massimi, P. and Banks, L. (1994) *Oncogene*, **9**, 1869–1876.

Pim, D., Thomas, M. and Banks, L. (2002) *Oncogene*, **21**, 8140–8148.

Piotto, M., Saudek, V. and Sklenár, V. (1992) *J. Biomol. NMR*, **2**, 661–665.

Rapp, L. and Chen, J.J. (1998) *Biochim. Biophys. Acta*, **1378**, 1–19.

Ristriani, T., Masson, M., Nominé, Y., Laurent, C., Lefèvre, J.F., Weiss, E. and Travé, G. (2000) *J. Mol. Biol.*, **296**, 1189–1203.

Ristriani, T., Nominé, Y., Laurent, C., Weiss, E. and Travé, G. (2002) *Protein Expr. Purif.*, **26**, 357–367.

Ristriani, T., Nominé, Y., Masson, M., Weiss, E. and Travé, G. (2001) *J. Mol. Biol.*, **305**, 729–739.

Scheffner, M., Huibregtse, J.M., Vierstra, R.D. and Howley, P.M. (1993) *Cell*, **75**, 495–505.

Schwarz, E., Freese, U.K., Gissmann, L., Mayer, W., Roggenbuck, B., Sremlau, A. and zur Hausen, H. (1985) *Nature*, **314**, 111–114.

Summers, M.F., South, T.L., Kim, B. and Hare, D.R. (1990) *Biochemistry*, **29**, 329–340.

Thomas, M., Laura, R., Hepner, K., Guccione, E., Sawyers, C., Lasky, L. and Banks, L. (2002) *Oncogene*, **21**, 5088–5096.

van Tilborg, P.J., Czisch, M., Mulder, F.A., Folkers, G.E., Bonvin, A.M., Nair, M., Boelens, R. and Kaptein, R. (2000) *Biochemistry*, **39**, 8747–8757.

Veldman, T., Horikawa, I., Barrett, J.C. and Schlegel, R. (2001) *J. Virol.*, **75**, 4467–4472.

Vuister, G.W. and Bax, A. (1993) *J. Am. Chem. Soc.*, **115**, 7778–7782.

Werness, B.A., Levine, A.J. and Howley, P.M. (1990) *Science*, **248**, 76–79.

Wishart, D.S., Sykes, B.D. and Richards, F.M. (1992) *Biochemistry*, **31**, 1647–1651.

zur Hausen, H. (1991) *Virology*, **184**, 9–13.

Part II

Protein-protein interactions: Experimental setup of the holdup assay and optimisation of the experimental conditions for investigating peptide-protein interactions by Biacore.

Two methods to study protein-protein interactions are presented in the two following publications:

-Publication 3:

Capturing protein-protein complexes at equilibrium: The holdup comparative chromatographic retention assay. Charbonnier S, Zanier K, Masson M, Trave G. (2006) *Protein Expr Purif* **50**(1) :89-101

-Publication 4:

Kinetic analysis of the interactions of human papillomavirus E6 oncoproteins with the ubiquitin ligase E6AP using surface plasmon resonance. Zanier K, Charbonnier S, Baltzinger M, Nomine Y, Altschuh D, Trave G. (2005) *J Mol Biol* **349**(2):401-412

Publication 3:

Charbonnier S, Zanier K, Masson M, Travé G. Capturing protein-protein complexes at equilibrium: The holdup comparative chromatographic retention assay. (2006) *Protein Expr Purif* **50**(1):89-101

Several methods have been described for studying protein-protein interactions, among them the pulldown assay or the Biacore technology. The pulldown assay is mainly a qualitative interaction assay and detection of low affinity or fast dissociating protein complexes is often not possible. In addition, protein concentration in pulldown assays is very low and revelation of the experimental result is often mediated by western blotting, which increases the risk of artefacts. Finally the washing steps in classical pulldown approaches might displace binding equilibria and therefore render precise kinetic data extraction impossible. Concerning the Biacore technology, one limiting factor of this technology is the lack of precise data when studying protein interactions with low K_D values beyond 10 µM. In addition it requires sophisticated and expensive hardware. To address these limits, we designed a method, called "holdup assay", which allows studying low affinity or fast dissociating protein-protein complexes at equilibrium. It is suited for qualitative medium throughput screening purposes. On the other hand it does not dramatically influence the binding equilibrium of the studied interaction. Therefore it gives access to all protein species of the interaction and allows accurate estimation of the K_D value. The assay is mainly designed for working at µM concentration with purified recombinant proteins. We showed that this assay can also be adapted to study protein interactions by using proteins transiently expressed in eucaryotic cells.

Publication 4:

Zanier K, <u>Charbonnier S</u>, Baltzinger M, Nomine Y, Altschuh D, Trave G. Kinetic analysis of the interactions of human papillomavirus E6 oncoproteins with the ubiquitin ligase E6AP using surface plasmon resonance. (2005) *J Mol Biol* **349**(2):401-412

Biacore technology represents a very powerful tool for studying protein-protein interactions. It is a technology based on surface plasmon resonance (SPR) and allows monitoring of protein-protein interactions in real time. Briefly, a first molecule is immobilized on a sensorchip surface, which is in contact with a flow of buffer. The second molecule is injected into this buffer flow and the interaction can be monitored in real time. SPR is very sensitive and gives access to the kinetic association and dissociation and permits to describe very precisely binding kinetics and affinities of protein complexes. However protein-protein interactions display mainly medium to low affinities. Working in this affinity range requires injection of highly concentrated proteins. The major drawback is that often proteins start altering their binding behaviour due to aggregation or precipitation. As SPR is very sensitive, even slight changes in protein quality or binding behaviour alter the experimental output and totally falsify the results. In the present publication, we optimised experimental conditions for testing the interaction of the E6 oncoprotein with GST-coupled peptides. We used a "GST-capture kit". Briefly, an anti GST antibody is bound covalently on the sensorchip surface, enabling immobilisation of a GST-peptide. E6 is injected and the kinetic binding parameters are recorded. The great power of this method consists in the fact that the antibody GST-fusion interaction can be disrupted and the chip surface therefore regenerated in order to measure a new interaction. A sensorchip surface can thereby be used up to 50 times. We optimised the quality of the protein sample preparations, worked out the buffer conditions and the experimental setup on the Biacore apparatus in order to measure accurately the affinity of this peptide/protein complex. This setup was then used by the first author to study more in detail the kinetic binding behaviour of different E6 oncoproteins with peptides originating from the E6AP ubiquitin ligase.

Capturing protein–protein complexes at equilibrium: The holdup comparative chromatographic retention assay

Sebastian Charbonnier, Katia Zanier, Murielle Masson et Gilles Travé

Kinetic Analysis of the Interactions of Human Papillomavirus E6 Oncoproteins with the Ubiquitin Ligase E6AP Using Surface Plasmon Resonance

Katia Zanier, **Sebastian Charbonnier,** Mireille Baltzinger, Yves Nominé, Danièle Altschuh and Gilles Travé

Part III

Analysis of the dissociation konstants of high-risk HPV E6 C-terminal peptides to MAGI-1 PDZ1 and structural investigation of the binding specificity of HPV16 E6 to MAGI-1 PDZ1.

My work focussing on the study of the interaction of HPV 16 E6 to the cellular protein MAGI-1 is presented in publication 5 and 6 as well as in the supplementary data section:

-Publication 5:

Investigating binding specificity and kinetics of the interaction of HPV E6 oncoproteins with the PDZ1 domain of MAGI-1. Charbonnier S, Nominé Y, Fournane S, Masson M, Stier G, Altschuh D, Kieffer B, Atkinson RA, Travé G. *[manuscript in preparation]*

-Publication 6:

Assignment of the PDZ1 domain of MAGI-1 using QUASI. Charbonnier S, Coutouly MA, Kieffer B, Trave G, Atkinson RA., (13)C, (15)N and (1)H Resonance (2006) *J Biomol NMR* **36** Suppl 5:33

Publication 5:

Charbonnier S, Nominé Y, Fournane S, Masson M, Stier G, Altschuh D, Kieffer B, Atkinson RA, Travé G. Investigating binding specificity and kinetics of the interaction of HPV E6 oncoproteins with the PDZ1 domain of MAGI-1. *[manuscript in preparation]*

Publication 6:

Charbonnier S, Coutouly MA, Kieffer B, Trave G, Atkinson RA., (13)C, (15)N and (1)H (2005) Resonance Assignment of the PDZ1 domain of MAGI-1 using QUASI. *J Biomol NMR* **36** Suppl 5:33

The high-risk mucosal HPV E6 mediated degradation of the cellular PDZ domain-containing proteins is at least one of the major factors for HPV-mediated cancerogenesis. MAGI-1 is one of these targets and interacts with E6 via its PDZ domain n°1. Structural investigations concerning the binding mechanism revealed a conserved binding mechanism implicating a C-terminal peptide, which binds into a groove located in the PDZ domain. However, this conserved binding mechanism is not sufficient to describe the binding specificity and we show that additional residues are involved in binding. Although many functional investigations have been performed concerning the interaction of high risk mucosal E6 oncoproteins to PDZ containing proteins, no structural data have been collected concerning the binding specificity of E6 to PDZ domains.

In the present publication, I investigated the binding behaviour of HPV16 E6-C domain to the PDZ1 of MAGI-1 by means of NMR and analysed the binding kinetics of high risk HPV E6 oncoproteins to MAGI-1 PDZ1. For a little bit more detailed information about the NMR techniques used, especially concerning the ^{1}H-^{15}N HSQC experiments, please confer to the NMR chapter in the materials and methods section.

I describe in detail how I obtained a pure recombinant and isotopically labelled MAGI-1 PDZ1 domain, suitable for NMR experiments. We assigned most of ^{1}H , ^{13}C and ^{15}N frequencies of this domain. NMR titration experiments allowed identifying a C-terminal region of the E6-C domain, which exhibits all necessary elements for binding specificity. I decided to investigate the binding behaviour of different high-risk HPV E6 C-terminal peptides to the PDZ1 domain. In order to perform that, we optimised the afore described Biacore method for analysing protein-protein interactions, making it faster and more accurate. Finally, we determined the K_D values of the different high-risk HPV C-terminal peptides, and sequence analysis in function of binding affinity allowed us to identify residues that might be important for ligand binding specificity.

Investigating binding specificity and kinetics of the interaction of HPV E6 oncoproteins with the PDZ1 domain of MAGI1.

Sebastian Carbonnier[1], Yves Nominé[1], Sadek Fournane[1], Murielle Masson[1], Gunter Stier[4], Danièle Altschuh[2], Bruno Kieffer[3], R.Andrew Atkinson[3], Gilles Travé[1+]

SUMMARY

High-risk human papillomaviruses can cause anogenital and skin cancers. Two early viral genes code for the multifunctional oncoproteins E6 and E7, which interact with several cellular proteins and drive cells to malignant progression. Recent findings suggest that the E6 protein contribute to HPV16 oncogenic potential via at least two degradation pathways: degradation of the tumor suppressor p53 and of PDZ-domain containing proteins. Human MAGI-1 is a multi-PDZ protein located at cellular junctions, which is implicated in multi-protein complex assembly and cell-cell communication. Mucosal high-risk HPV E6 proteins carry a class I PDZ consensus binding motif at their C-terminal extremity and target MAGI-1 for proteasome-mediated degradation via its PDZ1 domain. A monodisperse and active MAGI-1 PDZ1 domain was produced by varying the domain boundaries and optimising expression and purification conditions. ^{1}H-^{15}N HSQC titrations with PDZ1, E6-C and E6 C-terminal peptides allowed the last 12 C-terminal residues of E6 to be identified as the minimal binding sequence. Kinetic analysis of the interaction between MAGI-1 PDZ1 and eight different HPV E6 C-terminal peptides reported to bind MAGI-1 PDZ1, revealed equilibrium dissociation constants in the low to medium micromolar range. Analysis of the sequences of these peptides as a function of binding affinity allowed a preferential negative charge in the –3 and positive charges in the –4 to –6 positions of the C-terminal peptides to be identified as important for binding specificity.

(1) Équipe Oncoprotéines, UMR CNRS 7175-LC1, École Supérieure de Biotechnologie de Strasbourg, Boulevard Sébastien Brandt, BP 10413, 67412 Illkirch Cedex, France.
(2) Équipe Biocapteurs, UMR CNRS 7175-LC1, École supérieure de Biotechnologie de Strasbourg, Boulevard Sébastien Brandt, BP10413, 67412 Illkirch Cedex, France.
(3) Biomolecular NMR Group, UMR CNRS 7104, IGBMC, ESBS, Boulevard Sébastien Brandt, 67400 Illkirch, France.
(+) Corresponding author
tel: 0033 3 90 24 47 20, fax: 0033 3 90 24 47 70, e-mail: trave@esbs.u-strasbg.fr

INTRODUCTION

Human papillomaviruses (HPVs) are small double-stranded DNA viruses which can be divided into 17 families, comprising more than 100 different types (1). They can be classified into low-risk and high-risk types according to their ability to induce cancer, with 99% of cervical cancers being associated with infection by high-risk mucosal HPV types (2). HPV16, 18, 31, 33, 45 and 51 are the most common types identified in squamous cell carcinoma of the uterine cervix (3-5) with more than 50% of cases worldwide being associated with HPV16 (6). However, progression into cancer is rare, yet cervical cancer represents 6% of female diseases (7) and is the third cause of cancer-related death in women worldwide (8). High-risk genital HPV-related carcinogenesis has been linked to the expression of two early viral oncoproteins, E6 and E7(9). E7 interacts with and partly degrades a variety of cellular proteins, notably Rb, p107, p130, E2F/Cyclin A complex and Cyclin E, thus disrupting regulation of the cell cycle. E6 also interacts with and degrades more than 20 cellular proteins, including the tumor suppressor p53, E6AP, E6BP, Bak, Paxillin, IFR3 and PDZ domain-containing proteins. Thus it contributes to cell immortalization and transformation (10). It has been shown that the HPV16 E6 gene alone is sufficient to induce carcinomas in transgenic mice (11), and that the ability of HPV16 E6 to bind PDZ domains is required for E6-mediated immortalization of mammalian cells (12), p53-independent transformation of rodent cells (13) and induction of epithelial hyperplasia *in vivo* (14). At least two distinct activities contribute to the oncogenic potential of E6, namely degradation of p53 and degradation of PDZ domain-containing proteins (15). The HPV16 E6 oncoprotein has been shown to target membrane-associated guanylate kinases (MAGUK) Dlg (13,16-18), hScrib (19) and MAGUK with Inverted domains-1 (MAGI-1) (20), as well as the non-MAGUK protein MUPP-1 (21), to proteosomal degradation via PDZ domain-mediated interactions.

The common function of PDZ domains consists of binding and bringing into proximity components of the cellular signalling pathway (22-24). PDZ domains are 90-100 amino acid polypeptides, which adopt a stable fold, consisting of six β-strands forming a β-barrel, capped by two α-helices (25,26). The domains have a binding groove situated between the second β-strand and the second α-helix (25,26), recognizing short, four amino-acid C-terminal motifs which fall into four classes according to their consensus sequence (27,28). According to the standard nomenclature, residues within this consensus sequence are numbered 0 for the C-terminal residue and –1, -2, etc. for subsequent N-terminal residues (25). HPV16 E6 sequence contains the class I consensus binding motif X-T/S-X-V/L at

its C-terminus and is able to bind to the PDZ1 domain of the MAGI-1 and drive it to degradation (29,30).

MAGI-1 is a tight junction associated protein, which bears one guanylate kinase-like, two WW and six PDZ domains, named PDZ0-PDZ5 (31,32). Three splice variants have been described, MAGI-1a and MAGI-1b, which are both localized to cell-cell junctions and MAGI-1c which is localized to the nucleus due to a bipartite nuclear localization signal (NLS). In the kidney, MAGI-1 is enriched in glomerular podocytes, where it is linked via α-actinin-4 and synaptopodin to the actin-cytoskeleton of podocytes and is involved in the molecular architecture of the slit diaphragm (33,34). MAGI-1b is localized at adherens junctions of colonic and kidney epithelial cells, where it is implicated as a scaffolding molecule in the formation of a cell junctional complex by directly binding the C-terminus of the PTEN tumor suppressor *via* its PDZ2 domain and anchoring it at the cell-cell junction by binding β-catenin via its PDZ5 domain. MAGI-1b is implicated in the PTEN-mediated control and suppression of cellular invasiveness (35). The binding specificity of the proteins to their cognate PDZ domain clearly plays an important role. However, over 440 putative PDZ domains have been identified in the human genome (36) and therefore the four classes of C-terminal consensus binding motifs are not sufficient to describe protein-PDZ binding specificity. It has been shown that a region of five to eight amino acids N-terminal to the consensus motif contains additional residues which confer additional specificity and selectivity, and that the amino acid composition in the PDZ binding groove could allow specific binding of different classes of PDZ-binding consensus motifs (37-40).

The ability of high-risk HPV E6 oncoprotein has been described as one essential activity for generating hyperplasias *in vivo* as well as for malignant progression into cancer. It has been shown, that E6 is more implicated into malignant progression than in promotion of cancer (15). This seems to correlate with the hypothesis that degradation of PDZ-domain containing proteins is important in the late stages of cancer development and contributes to the loss of cell polarity and inhibition contact, thus driving the cells towards invasive phenotypes (41). Detailed information of the interaction of high-risk HPV E6 oncoproteins with PDZ domains is therefore of interest in order to find new potential targets for therapeutic applications.

In this context we describe the optimised expression of MAGI-1 PDZ1. We analysed the interaction between MAGI-1 PDZ1 and HPV 16 E6-C. by NMR. We also investigated the kinetic parameters of the interaction of C-terminal peptides from high-risk and low-risk HPV types with MAGI-1 PDZ1 by Biacore. Combined analysis of the equilibrium

dissociation constants and the structural data allowed to identify the minimal C-terminal peptide conferring binding specificity of this precise PDZ-protein interaction.

EXPERIMENTAL PROCEDURES

cDNA constructs

The cDNA of the canonical MAGI-1 PDZ1 construct (PDZ1-0: residues 468-555 of human MAGI-1 [Thomas, 2001 #212], Swiss-Prot acc.: AF401656) was inserted into the NcoI/KpnI sites of the pETM-11, -30 and -41 expression vectors providing N-terminal His6, His6-GST and His6-MBP tags, respectively and a TEV protease cleavage site. The constructs of MAGI-1 PDZ1 called PDZ1-1, PDZ1-2, PDZ1-3, PDZ1-4 and PDZ1-5, comprising residues 468-566, 460-566, 468-580, 460-580 and 456-580 of human MAGI-1, respectively were cloned the same manner into the pETM-41 expression vector. Pairs of complementary DNA oligomers coding for the following peptides: none (GSNSGNGNS), HPV16e6ct (RSSRTRRETQL), HPV18e6ct (NRARQERLQRRRETQV), HPV31e6ct (WRRPRTETQV), HPV33e6ct (WRSRRRETAL), HPV45e6ct (DQARQERLRRRRETQV), HPV52e6ct (WRPRPVTQV), HPV58e6ct (WRPRRRQTQV), HPV11e6ct (WTTCMEDLLP), mNETct (QSGGKKETLV), hpTENct (EPFDEDQHTQITKV) and hbAr2ct (DSQGRNCSTNDSLL) were cloned into the NcoI/KpnI sites of the pETM-30 expression vector. All constructs produced by means of these pETM-vectors bear an additional peptidic GAMG sequence at the N-terminus. For the PDZ1-5 construct, the first G of this additional sequence was numbered as residue 1 for NMR studies.

Protein expression

Expression of unlabelled MBP-E6 (6C/6S) as well as unlabelled and ^{15}N-labelled MBP-E6C 4C/4S was carried out as described previously (42,43). All proteins were expressed in BL21 DE3 E. coli cells. MAGI-1 PDZ1-conctructs were grown in 25 ml LB cultures for monitoring protein quality and 200-300 ml of LB or M9 minimal medium, supplemented either with ^{15}NH$_4$Cl or with ^{15}NH$_4$Cl and ^{13}C-glucose, at 37°C until an OD$_{600}$ of 0.6 was reached. Cultures were then adjusted to 0.5 mM isopropyl-D-thio-ß-galactopyranoside (IPTG) and transferred to 27°C for 4 h in the case of LB or 6 h in the case of M9 cultures. GST-peptide expression cultures were grown in 25 ml LB cultures at 37°C until an OD$_{600}$ of 0.6 was reached, induced with IPTG and further grown for 2 h. Plasmid loss was suppressed by adding 15 µg/ml of kanamycin to the expression media. Expression cultures were harvested by centrifugation, the pellets shockfrozen in liquid nitrogen and stored at –20°C.

Purification procedures

All purification buffers were extensively degassed and bubbled with argon. Purification of MBP-E6 (6C/6S) as well as unlabelled and ^{15}N-labelled MBP-E6C (4C/4S) was performed as described previously (42,43). To purify PDZ1-3, PDZ1-4 or PDZ1-5, the bacterial expression pellets of unlabelled or ^{15}N-labelled MBP-PDZ1 constructs were sonicated in lysis buffer (50 mM Tris-HCl (pH 6.8), 400 mM NaCl, 1 mM DTT, 5 % glycerol, 1 µg/ml DNase I, 1 µg/ml RNase A and anti-protease cocktail (EDTA-free) (Roche)), cleared by centrifugation at 18,000 g and filtered (Millipore 0.22 µm). MBP-PDZ1-3, MBP-PDZ1-4, or MBP-PDZ1-5 extracts were loaded on an amylose column (New England Biolabs) pre-equilibrated with buffer A (50 mM Tris-HCl (pH 6.8), 400 mM NaCl, 1 mM DTT). Protein was eluted with buffer A supplemented with 10 mM maltose. The MBP-tag was removed by proteolytic cleavage with recombinant TEV-protease at 8°C overnight. To remove both the His6-MBP moieties and the His6-TEV-protease, the digestion mix was loaded 3 times on a 2 ml NiNTA resin (Qiagen) pre-equilibrated with buffer A and the resulting flow-through, containing essentially the PDZ1-3, PDZ1-4 or PDZ1-5 domain, was subjected to a final gel-filtration on a Hiload 16/60 Superdex 75 column (Amersham Biosciences) equilibrated in buffer A, resulting in pure monodisperse PDZ1-3, PDZ1-4 or PDZ1-5. For Biacore experiments, bacterial overexpression pellets of GST-peptides were suspended in 1 ml of buffer B (50 mM Tris-HCl, 400 mM NaCl, 1 mM DTT, pH 6.8) supplemented with 5% glycerol, 1% NP-40, 1 µg/ml DNase I, 1 µg/ml RNase A and antiprotease cocktail (EDTA-free) (Roche), sonicated on ice in 2 ml Eppendorf tubes and centrifuged at 16000 g and 4°C for 15 min. The resulting supernatants were stored at –80°C.

Synthetic peptides

Two peptides, e6ct(11) (RSSRTRRETQL) and e6ct(5) (RETQL), corresponding to the last 11 and last 5 C-terminal residues of HPV16-E6, respectively, were synthesized by P. Eberling (Chemical Peptide Synthesis Service, IGBMC, France). Peptides were resuspended in buffer A and passed on a NAP-5 desalting column pre-equilibrated in buffer A to remove acidic components. The peptides were checked by homonuclear 1D NMR experiments and their concentration after buffer exchange adjusted to 5 mM for e6ct(11) and to 2.5 mM for e6ct(5).

Protein quality monitoring

For the micropurification of recombinant His6-MBP fusion proteins of PDZ1-0, PDZ1-1, PDZ1-2, PDZ1-3, PDZ1-4 and PDZ1-5, bacterial overexpression pellets corresponding to 25 ml of expression culture, were sonicated in 1 ml of sonication buffer and cleared by centrifugation at 16,000 g. The supernatants were incubated for 30

min at 4°C in a 1.5 ml Eppendorf tube with 150 μl of Amylose Agarose resin (New England Biolabs) pre-equilibrated in buffer A and washed 6 times with 1 ml of buffer A. Resin was settled by centrifugation, the superfluous buffer discarded and fusion proteins eluted by adding 150 μl of buffer A containing 20 mM maltose for 10 min. The liquid was extracted from the resin by transferring the resin-liquid mixture to 0.5 ml Eppendorf tubes possessing a pin-hole, filled with glass wool and placed on top of 2 ml tubes and centrifuged at 1,000 g in a bench-top centrifuge, yielding pure His6-MBP-PDZ constructs (44). Aggregation ratio were measured as published earlier (45). Briefly, static light scattering and fluorescence were measured with a SPEX Fluorolog-2 spectrofluorimeter (SPEX Industries, Edison, NJ). Data were acquired with a photomultiplier with voltages fixed at 800 V. 2 ml of buffer A adjusted to 100-300 nM of protein sample were placed in a cuvette maintained at 20°C and excited by monochromatic light at 280 nm or 350nm. An emission spectrum was recorded in the range from 260-360 nm and a spectrum for buffer alone served as baseline. The maximal scattered light intensity at 350 nm and the maximal fluorescence light intensity at around 340 nm were recorded and the aggregation ratio calculated according to :

$$R_{agg} = I_{280} / I_{340}$$

For pellet-supernatant quality tests, 30 μl of micropurified His6-MBP-PDZ1 constructs were digested with recombinant TEV protease for 2-3 h at room temperature. The digestion mix was centrifuged at 16000 g and 4°C for 15 min. The resulting supernatant was transferred to a new tube and mixed with 30 μl of 2x sample buffer and the pellet resuspended in 30 μl of buffer A and mixed with 30 μl of 2x sample buffer. Samples were subjected to 15% SDS-PAGE and analyzed by Coomassie blue staining.

Biological activity monitoring by holdup assay

The previously described holdup method was adapted to test the E6-binding ability of the PDZ1-3, PDZ1-4 and PDZ1-5 constructs (46). Bacterial overexpression pellets of His6-MBP fusion proteins of PDZ1-3, PDZ1-4 and PDZ1-5 and the negative control, His6-MBP-E6, corresponding to 25 ml of expression culture, were sonicated in 1 ml of sonication buffer and cleared by centrifugation at 16000 g. Each supernatant was incubated for 30 min at 4°C in a 1.5 ml Eppendorf tube with 20 μl of Maltoheptaose Agarose (Sigma) pre-equilibrated with buffer A. The ligand-saturated resin beads were washed 4 times with 1 ml of buffer A. The resin-buffer mixture was adjusted to a final 1:5 ratio and 10 μl aliquots of each (corresponding to 2 μl of ligand-saturated resin beads) were aliquoted into two 500 μl Eppendorf tubes, labelled (-M) and (+M). The (-M) and (+M) tubes of each His6-MBP-ligand were

supplemented with 15 μl of a 17 μM pure and monodisperse E6 protein preparation in buffer B (50 mM Tris-HCl, 400 mM NaCl, 1 mM DTT) and incubated for 15 min on ice, interspaced with gentle vortexing. (-M) tubes were supplemented with 4 μl of buffer A, (+M) tubes with 4 μl of buffer A containing 100 mM maltose and incubated for an additional 10 min at 4°C, interspaced with gentle vortexing. The (-M) and (+M) resin-buffer mixtures were transferred to a 96-well filter-plate, the liquid phase was extracted by centrifugation, supplemented with 1 volume of 2x sample buffer and analyzed by SDS-PAGE and Coomassie staining.

Protein interaction test and determination of the equilibrium dissociation constant by Biacore

All experiments were performed on a Biacore 2000 apparatus at 25°C with the autosampler rack base cooled at 10°C. The running buffer was 50 mM Tris-HCl (pH 6.8), 200 mM NaCl, 1 mM DTT, 0.005% P-20. Surfaces were regenerated with 50 mM glycine (pH 2.2).
The binding behaviour of PDZ1-3, PDZ1-4 and PDZ1-5 was monitored using a protocol described previously (47). Briefly, experiments were performed on a single flow-cell of a CM5-chip surface on which 16,000-18,000 RU of anti GST-antibody were immobilized. GST-peptides were captured from soluble fractions of crude *E. coli* extracts diluted 1:10 in buffer C (50 mM Tris-HCl, pH 6.8, 400 mM NaCl, 1% NP-40, 1 mM DTT) and injected at a flow-rate of 10 μl/min during 5 min until saturation of the surface was reached. The interaction of the analyte with the negative control peptide, GST-none, and the test peptide, GST-HPV16e6ct, were measured on the same flow cell in two subsequent cycles. In the first cycle, the GST-none peptide was captured on the antibody surface, followed by a 10 min equilibration step in running buffer. E6 was then injected and the post injection phase recorded for 3 min. Finally the surface was regenerated by injecting two 1-min-pulses of regeneration solution. In the second cycle, the same procedure was repeated for the GST-HPV16e6ct peptide. The effective binding response was derived by subtracting the sensorgram obtained for the GST-none peptide from that of the GST-HPV16e6ct peptide.
Kinetic measurements of 11 different GST-peptides against PDZ1-5 were performed using a modified protocol where experiments were recorded using the four flow-cells of the sensorchip in parallel. First, the anti-GST-antibody coupling was modified by limiting the level of anti-GST antibody on the chip in order to improve repeatability of the data and reduce Biacore artifacts. The running buffer for antibody coupling was HBS buffer (10 mM HEPES, 150 mM NaCl, supplemented with 0.005% P-20) and a new CM5 sensor surface was washed with two 30 s pulses of 50 mM NaOH, 500 mM NaCl and one 30 s pulse of regeneration solution and then left for 1 h at a flow-rate of 20 μl/min in running buffer. Each flow-cell of the sensor chip surface was activated

independently for 10 min with a solution containing N-ethyl-N'-[3'-(diethylamino)-propyl]carbodiimide (EDC) and N-hydroxysuccinimide (NHS) at a flow-rate of 10 μl/min. Anti-GST-antibody was then fixed by manual injection of highly diluted anti-GST polyclonal antibody (8 μg/ml) in 10 mM sodium acetate coupling solution (pH 5.5) at a high flow-rate (40 μl/min) until 1500 RU was reached. The surface was deactivated by injecting ethanolamine (1M) for 7 min followed by four 1-min-pulses of regeneration solution at a flow-rate of 10 μl/min. Secondly, the running buffer for peptide-protein interaction tests was changed to buffer A supplemented with 0.005% P-20. Each flow-cell was saturated with GST-peptide in mono-channel mode, where flow-cell 1 usually contained the GST-none peptide (negative control). Soluble fractions of crude *E. coli* extracts diluted 1:10 in buffer C were injected during 6 min at a flow-rate of 10 μl/min. For kinetic measurements, the flow-rate was set to 20 μl/min and purified PDZ1-5 was injected at 8 different concentrations (0, 1, 2, 5, 15, 20, 30 and 60 μM) for 30 s followed by a 1-min-pulse of the synthetic e6ct(11) peptide concentrated at 25 μM using the "coinject" mode. This prevented rebinding of the PDZ1-5 domains on the surface and allowed a rapid return to a stable baseline. At least, two independent experiments were performed for each GST-peptide / PDZ1-5 interaction.

Data were analyzed using the BiaEvaluation 3.2 software (Biacore). Steady state analysis was performed by fitting equilibrium responses (R_{eq}) as a function of total analyte concentration with a simple 1:1 interaction binding isotherm model. The kinetic parameters reported here were obtained from fits with χ^2 values smaller than 5% of the globally fit R_{max} values.

NMR data acquisition

Samples for NMR experiments were prepared in 25 mM Tris-HCl (pH 6.8), 50 mM NaCl, 1 mM DTT, 10% D_2O. In order to investigate the binding between MAGI1-PDZ1 and E6-C, ^1H-^{15}N heteronuclear single quantum coherence (HSQC) spectra were acquired at 600 MHz and 295K on a Bruker DRX600 spectrometer equipped with a z-gradient triple-resonance cryoprobe. The water signal was suppressed using the WATERGATE sequence (48). Data were processed using NMRPipe (49) and analysed with NEASY (50).

RESULTS

Domain phasing of MAGI-I PDZ1

Sequence alignment of class I PDZ-domains of known structure (Fig. 1A) allowed the definition of a construct, MAGI-1 PDZ1-0$_{(468-555)}$, bearing the 6 β-strands and 2 α-helices of a canonical PDZ-domain (Fig. 1A/B). Expression and purification of PDZ1-0$_{(468-555)}$ gave mediocre results according to the pellet/supernatant test as well as the aggregation rate (R_{agg}) value (R_{agg} = 0.3), indicating

the presence of fusion protein multimers (Fig. 2). Attempts to purify this construct proved difficult and purified samples were unstable and unsuitable for NMR analysis (data not shown). In an endeavor to stabilize the domain, we adopted a domain phasing approach, adding N- and C-terminal elongations from the sequence of full-length MAGI-1 protein to the initial PDZ1-0 domain sequence. Five constructs, PDZ1-1, PDZ1-2, PDZ1-3, PDZ1-4 and PDZ1–5, were generated by adding combinations of a C-terminal proline-rich region, an N-terminal 8 amino acid stretch, a C-terminal 14 amino acid stretch ending at the next proline residue and a further N-terminal 4 amino acids containing a hydrophobic patch and a proline (Fig. 1B).

Analysis of protein solubility and monodispersity

PDZ1-1 and PDZ1-2 were revealed in the soluble fraction, after having performed a pellet/supernatant test but the bands were blurred and less intense than those of the three other PDZ1 constructs. This indicated possible proteolysis, multimerization and/or aggregation. Furthermore the high R_{agg} values for PDZ1-1 and PDZ1-2 clearly showed that these domains were forming soluble aggregates (Fig. 2) (44). In contrast, PDZ1-3, PDZ1-4 and PDZ1-5 displayed intense, clear-cut bands in the soluble fraction and very low R_{agg} values, indicating that these three constructs were produced soluble and monomeric (Fig. 2).

Analysis of biological activity using a "hold-up" assay

To evaluate the E6 binding, pure-pure "hold-up" assays (46) were performed with the soluble and monodisperse MBP-PDZ1-3, MBP-PDZ1-4 and MBP-PDZ1–5 proteins as ligands and with E6 as analyte. MBP-E6 served as a negative control ligand, as E6 is known not to bind to itself during the short incubation times of this method (data not shown). The presence of both MBP-PDZ1 and E6 were revealed in the (+M) fractions. The absence of E6 in the (-M) fractions, by chromatographic retention via the MBP-PDZ1 constructs, showed that the three PDZ1 constructs were biologically active insofar as E6-binding was concerned (Fig. 3). In the negative control, E6 was not retained on the resin and was visible at equal concentrations in the (-M) and (+M) fractions, excluding the possibility of non-specific binding to the MBP-moiety or to the resin.

Choice of an appropriate PDZ1 construct for structural analysis by NMR

^1H-^{15}N HSQC experiments in solution revealed differences in behaviour of the PDZ1-3, PDZ1-4 and PDZ1-5 constructs (Fig. 4A). The backbone amide proton frequencies of all three constructs were well dispersed over a broad frequency range, indicating that the polypeptides adopted a non-random coiled conformation. However, the spectrum of PDZ1-3 showed

considerable heterogeneity in peak line-widths and intensities. This could come from intra-molecular conformational exchange or, more probably, from self-association in solution at the concentrations required for NMR measurements. PDZ1-4 displayed a much improved spectrum in terms of heterogeneity of line-width, indicating a reduction in intra- or inter-molecular exchange. Nonetheless, a small number of cross-peaks are clearly broadened. On the contrary, the spectrum of PDZ1-5 was further improved with little peak broadening, reflecting a well-folded, monomeric polypeptide chain (Fig. 4A).

Consistent results were obtained upon analysis of the binding behaviour of the three PDZ1-domains to HPV16e6ct peptide using Biacore technology (Fig. 4B). PDZ1-3 showed a significantly different behaviour in binding to the E6 C-terminal peptide fused to GST (GST-HPV16e6ct) than either PDZ1-4 or PDZ1-5. We were not able to fit the experimental curves, obtained for PDZ1-3 with the expected Langmuir binding model. The shapes of the curves might indicate a biphasic binding behaviour. This again might suggest that PDZ1-3 was not present as a monomer but undergoes self-association, yielding a heterogeneous sample with more complex binding kinetics (51). Both PDZ1-4 and PDZ1-5 binding curves showed similar shapes (Fig. 4B). Monomeric and well-folded MAGI-1 PDZ1 displayed an association phase characterised by a fast association rate and two dissociation phases, with an initial fast dissociation rate and a second slower dissociation rate. This biphasic dissociation phase results probably from re-binding of the PDZ1 domain to the matrix in the post-injection phase. This phenomenon could be suppressed in later experiments by co-injecting a large excess of the E6 C-terminal peptide in the post-injection phase (Fig. 9A). Based on these results, the PDZ1-5 domain was chosen to further investigate the interaction with the HPV16 E6-C domain.

Identification of the residues of HPV16 E6-C implicated in MAGI-1 PDZ1 binding

[15]N-labelled HPV16 E6-C (E6-C*) was titrated by NMR with unlabelled PDZ1-5. ^{1}H-^{15}N HSQC spectra were recorded to identify residues of E6-C* for which the chemical shifts of backbone amide ^{1}H and ^{15}N nuclei were affected upon binding (Fig. 5A). Titration was performed until no additional change in chemical shift for E6-C* was observed. A set of 21 out of the 72 residues of E6-C* exhibits changes in chemical shifts. This observation allowed identification of a set of residues implicated in MAGI-1 PDZ1 binding, which could be grouped in 4 clusters on the E6-C structure (PDB ID : 2FK4; (43)).

(i) The most strongly perturbed chemical shifts were assigned (52) to the last 12 C-terminal amino acids (S140-L151), which were determined to be unstructured in solution. The C-terminal residues E148, T149, Q150 and L151, correspond to the C-terminal X-S/T-X-L/V class I consensus binding motif of PDZ domains (53) but additional strong perturbations of residues R144, T145, R147, indicate that they might play a direct or indirect role in the binding. Weaker chemical shifts changes were observed for S140 and R141. Residue S143 displays no significant chemical shift change and the amide proton frequencies of S142 and R146 are not visible on the HSQC spectrum (Fig. 5A,B).

(ii) The second cluster comprises residues I104, N105, I128, R135, C136 and C139, including side-chains that form the zinc-binding site. These residues are all located close to the C-terminus of E6-C* in the three-dimensional structure of the domain and might therefore be influenced directly by the binding of the PDZ domain to the C-terminus of E6-C* (Fig. 5A,B). The strong chemical shift change detected for R135 should be noted as it has been identified as a buried hydrophilic residue conserved among the E6-C domains of high-risk HPV (43). Conversely, I128 is also of note as it has been identified as a conserved exposed hydrophobic residue, which could be a critical determinant for PDZ domain binding specificity.

(iii) The third cluster comprises residues G85, L88, K121, R124 (Fig. 5A,B) located on the loop between the first β-strand and the first α-helix and on the loop connecting the second α-helix to the second β-strand (Fig. 5A,B). This cluster is situated on the opposite side of the domain and comprises mainly positively-charged residues.

(iv) The fourth cluster comprises two polar residues, S82 and N127, situated on the first and the second β-strands, respectively, and therefore on the bottom of the β-sheet of E6-C (Fig. 5A,B).

It is not clear yet whether the chemical shift changes of the third and fourth clusters are due to long range rearrangements upon PDZ-domain binding or if some of these residues play a role in the binding of MAGI1-PDZ1. Further chemical shift changes are observed for the residue S97 and for the buried residue F125, an essential part of the hydrophobic core of E6-C. This might confirm, that the chemical shift differences identified in the second cluster are more probably due to local structural rearrangements than to the change of chemical environment due to the local proximity of the PDZ domain upon binding.

Mapping of the minimal binding region of HPV16 E6-C to MAGI-PDZ1, which confers specificity

To confirm the extent of the set of residues of E6-C identified above as responsible for binding and defining specificity of binding, especially concerning the first cluster of C-terminal residues, ^{15}N-labelled MAGI1-PDZ1-5 (PDZ1-5*) was first titrated by NMR with unlabelled HPV16 E6-C until no more change in chemical shift differences of PDZ1-5* complexed to the E6-C was observed.. Among 40% of peaks visible on the ^{1}H-^{15}N HSQC spectra of PDZ are affected upon E6-C binding (Fig. 6). According to the sequential assignement of the free PDZ domain (46), and in contrast to the E6-C

domain, changes in chemical shift are observed throughout the entire PDZ domain upon E6-C binding.

PDZ1-5* was then titrated with peptides corresponding to the 11 C-terminal residues of 16E6-C (RSSRTRRETQL) and to the shortest C-terminal class I PDZ domain-binding consensus motif of E6-C (RETQL), named e6ct(11) and e6ct(5), respectively (Fig. 7A,B) until no more changes in chemical shifts of PDZ1-5* were observed. In both cases, the ^{1}H-^{15}N HSQC spectra show extensive changes of the amide proton and nitrogen frequencies (Fig. 7A,B) but superposition of the spectra with that of PDZ1-5* complexed to E6-C, (Fig. 7C), show that the spectrum of the PDZ1-5* – e6ct(5) differs considerably from the other two. Indeed, the spectra of PDZ1-5* – E6-C and PDZ1-5* – e6ct(11) are almost identical, with only minor chemical shift differences (Fig. 7C,D). This suggests that the residues of E6-C important for the binding to MAGI1-PDZ1 are located in the C-terminal 12 amino acids of the domain, but not exclusively in the last 5 residues as expected according to the class I PDZ domain-binding consensus motif.

Residues of PDZ1-5* which remain unchanged in their chemical shifts are G17, F19, L24, R25, D42, G54, C75, I88, S96, D98 and E128. None of these residues is located at key positions of the canonical binding groove of PDZ domains, which is located between the second β-strand and the second α-helix (Fig. 1A). Due to this localization, it can be expected that they are neither implicated into ligand binding nor in fold stabilization of the PDZ domain and therefore their chemical shifts remain unchanged.

As a result of the extensive chemical shift changes in the spectrum upon binding of PDZ1-5* to E6-C and E6-C C-terminal peptides, discrete clusters of residues of PDZ1-5* implicated in E6-C binding could not readily be identified, without assigning the resonances of the complexed form (to be published elsewhere).

Kinetic analysis of binding of C-terminal peptides of E6 to MAGI1-PDZ1-5$_{(456-580)}$

The finding that the C-terminal 12 amino acids of the HPV16-E6-C domain were sufficient for binding allowed the design of a peptide-based approach to analyze the binding kinetics of several high-risk HPV-E6 proteins and of other proteins reported to bind to MAGI1-PDZ1(54). Biacore technology was used to determine equilibrium dissociation constants (K_D) by applying a GST-peptide-based screening approach (47). An anti-GST antibody was preliminary immobilized on a CM5 sensor-chip surface. Then GST-peptides (ligands) were injected and bound to the chip via the antibody recognition. Finally the PDZ1-5 domain (analyte) was injected at 8 different concentrations and the relative signal measured by surface plasmon resonance. The previously published protocol was modified by an injection of the peptide ligand

(without any GST fusion) thereafter the injection of PDZ1-5 to avoid rebinding of the PDZ domain to the GST-peptide and therefore allow simultaneous work on all four flow-cells of the sensor-chip, providing a noticeable gain in experimental time and accuracy of the data. The peptides originating from HPV E6 proteins were: negative control GST-none, GST-HPV16e6ct, GST-HPV18e6ct, GST-HPV31e6ct, GST-HPV33e6ct, GST-HPV45e6ct, GST-HPV52e6ct, GST-HPV58e6ct, GST-HPV11e6ct. The peptides GST-mNETct, GST-hpTENct and GST-hbAr2ct were C-terminal peptides of human proteins which had been described as potentially binding MAGI-1 PDZ1. A representative Biacore data set of the positively tested GST-HPV16-e6ct peptide is represented in Fig. 8A as well as a representative data set of the negatively tested GST-HPV11e6ct peptide in Fig. 8B. The curves display a square profile, characteristic of a high association rate (k_{on}) and a high dissociation rate (k_{off}) which renders impossible the determination of k_{on} and k_{off} by global curve-fitting and thus the calculation of K_D. Instead, we performed a steady-state analysis by plotting the equilibrium responses (R_{eq}) as a function of the total injected analyte concentration and by fitting the resulting curve with a simple 1:1 binding isotherm model (Fig. 8C). Each experiment was at least recorded in duplicate. K_D values are in the low micromolar range and vary between 2.3 and 27.5 μM (Fig. 9A). The C-terminal peptides from E6 of high-risk HPVs with the highest affinities are those of HPV33, and HPV45. On the opposite, the C-terminal peptides from E6 of high-risk HPV 58 and 52 displayed the lowest affinities. The C-terminal peptide from E6 of low risk HPV11 shows no binding, as expected, and serves as a proper negative control. Of the non-HPV proteins, which were reported to bind (or potentially bind) MAGI-1 PDZ1 only the hmNET C-terminal peptide displayed binding with a mean K_D-value of 4.7 μM.

Correlation of the kinetic data according to the peptidic sequence

Analysis of the peptide sequences regarding to their experimental mean K_D values for binding to MAGI1-PDZ1 revealed that the C-terminal X-S/T-X-V/L class I PDZ domain consensus binding motif was necessary but not sufficient for binding. This consensus motif is required for binding MAGI-1 PDZ1, which is obvious by comparing all high risk HPV peptides to the HPV11e6ct peptide. The peptide from E6 of low risk HPV11 that lacks the C-terminal consensus motif did not bind at all to MAGI1-PDZ1 (Fig. 8C; Fig. 9). However this consensus motif is not sufficient for binding as suggested by analyzing the hpTENct and the hbAr2 peptides: both do bear the consensus motif, but fail to bind. Thus it seems obvious that attempts to correlate affinity and target sequence should include residues beyond this consensus motif.

Referring to the PDZ binding motif nomenclature, where the C-terminal residue is

numbered as 0 and subsequent residues towards the N-terminus as –1, -2, etc., the position 0 seems to be variable between a valine or a leucine residue, when the valine seems to be prefential. Threonine in position -2 is absolutely conserved among set of high risk HPV E6 proteins. The sequence alignment shows a high degree of conservation of a glutamate residue in the –3 position for high affinity peptides (Fig. 9B). Peptides such as HPV58e6ct and HPV52e6ct with glutamine and valine residue at this position, show a marked increase in the K_D values (Fig. 9A,B). The negative charge at the –3 position therefore seems to be important for determining binding affinity. HSQC experiments showed that the chemical shifts of cross-peaks of residues beyond position –6 are affected little upon binding to MAGI1-PDZ1 and could therefore be neglected (Fig. 5A/B). Furthermore, the sequence alignment reveals a positive patch of variable length, encompassing positions –4 to –6. The peptides HPV33e6ct, HPV45e6ct and HPV18e6ct, each containing 3 arginine residues at positions –4 to –6, displayed the highest affinities for MAGI1-PDZ1. Position –6 appears to be particularly important as HPV16e6ct, which lacks a positive charge at that position, shows an increased K_D value (Fig. 9B). Substitution of the arginine residues by lysine residues seems also to be disadvantageous: in the hmNETct peptide the arginine residues are substituted by lysine residues retaining the positive charge at position –6 yet the mean K_D value is comparable to that of HPV16e6ct. Further reduction in positive charge and the introduction of proline residues, likely to limit the conformational possibilities for the peptides, further increased the mean K_D values, as can be seen for the HPV31e6ct, HPV58e6ct and HPV52e6ct peptides. The substitution of the conserved glutamine residue at position -1 by a residue with a shorter sidechain, such as alanine in HPV33e6ct, may reduce steric hindrance in the binding pocket and contribute to a lower K_D value. In contrast, the leucine residue at this position in hmNETct may increase steric hindrance and so increase the K_D value. Interestingly, neither hpTENct nor hbAr2ct, which contain the X-S/T-X-V/L consensus motif, contain any of these additional key residues and completely fail to bind MAGI-PDZ1.

DISCUSSION

The ability of high risk mucosal HPV E6 proteins to target and drive PDZ domain containing proteins to proteasomal degradation has been identified as one essential activity to induce cell immortalization and progression into cancer (14,15). For years structural investigation of the PDZ domain mediated binding of E6 to its target proteins was impossible due to the problems encountered to produce folded and active recombinant E6 oncoprotein. Recently, a stable mutant of the C-terminal zinc binding domain E6-C, bearing the PDZ binding consensus motif, was purified into a monomeric and active form and the

NMR structure solved (42,43,52). In the present study we first designed and produced a well folded and active recombinant MAGI-1 PDZ1 domain and then investigated the binding properties of E6-C with MAGI-1 PDZ1 by NMR as well as the kinetic parameters by Biacore.

Preliminary results indicated that the minimal MAGI1 PDZ1 construct according to sequence alignments against PDZ domains of known structure led to samples which were unstable and unsuitable for NMR analysis. In order to produce high quality recombinant MAGI-1 PDZ1 samples, applied different techniques. Combination of the pETM-vector series (Stier G., EMBL Heidelberg, Germany), which allows the rapid subcloning in parallel of one given cDNA into a variety of different vectors providing different features, with the domain phasing approach, small scale expression condition screenings and aggregation rate measurements, allowed to rapidly identify the optimal expression vector and the optimal expression conditions. Holdup assays as well as small scale ^{15}N-labelling followed by basic HSQC experiments and a rapid method based on peptide based Biacore screening allowed to identify the optimal PDZ1-5 construct for further structural and kinetic analysis. The great power of this approach relies in simple and rapid methods which can be performed in parallel and which were optimised to allow rapid screening. This work-flow-scheme could therefore be adapted to any protein or protein domain purification *ab initio* for subsequent *in vitro* assays.

The structural analysis of the HPV 16 E6-C domain revealed a globular domain consisting of three β-strands and three α-helices, presenting a novel zinc binding fold and where the last twelve C-terminal residues, presenting the PDZ binding motif, were in random conformation (43). HSQC titrations using labelled E6-C and unlabelled PDZ1-5$_{(456-580)}$ allowed us to identify residues of E6-C affected upon PDZ binding and to assign them into four clusters according to their localization in the domain structure. The first cluster concerns the last eight C-terminal residues of E6-C, which are directly implicated into PDZ-binding via the consensus motif. The second cluster groups residues of the E6-C zinc binding domain which are close in space to the C-terminus. In general C-terminal PDZ binding consensus motifs binds to the PDZ binding groove between the second β-strand and the second α-helix (). They arrange in an extended and antiparallel fashion in respect to the second β-strand of the PDZ domain and the free carboxylate of the binding motif is hydrogen bonded by a GLGF motif situated prior to the the second β-strand of the PDZ domain (25). This very conserved binding mechanism among PDZ-ligand pairs lets suggest that the C-terminus of E6-C gets structured upon PDZ binding and therefore induces chemical shift perturbations in the region of the second cluster due to local structural rearrangements. This is supported by the fact that the buried residue F125, which is a key residue of the

hydrophobic core of the E6-C domain, displays also chemical shift modifications upon PDZ-binding, which could be induced by structural changes in the zinc binding domain of E6-C. Concerning residues found in the clusters three and four it is difficult to judge if the perturbation of the chemical shifts is due to a direct interaction with the PDZ domain or to long range conformational changes. L88 and I128 have been identified as conserved and exposed hydrophobic residues in different HPV E6-C domains. I128 is located near to the C-terminus and L88 has been proposed to play a role as an anchoring amino acid implicated in heterodimerization of the E6-C and E6-N domains of full length E6 (43). However, whether or not binding of a PDZ domain implicates additional residues outside the C-terminus or induce some relevant conformational changes on E6-C or full length E6, which might influence the activity, remains unclear.

More than 400 PDZ domains in more than 250 proteins have been identified in the human genome and also in other organisms (36,55). The simple model of four different classes of four amino acid consensus binding motifs is not sufficient to explain ligand-PDZ specificity. It has been shown that additional residues N-terminal to the consensus binding motif are important for the ligand-PDZ recognition and specificity (37-39,56). However, no general rule has yet been derived and apart the classes of consensus motives, each interaction remains individual in terms of specificity and must be investigated individually. Concerning the MAGI-1 PDZ1 / E6-C couple, HSQC experiments on labelled PDZ1-5$_{(456-580)}$ titrated with unlabelled E6-C as well as with C-terminal peptides of E6 allowed to confine the minimal MAGI1-PDZ1 binding region of E6 to the last twelve C-terminal residues. A kinetic analysis of the binding behaviour of C-terminal peptides of seven high risk, one low risk and three other putative MAGI1-PDZ1 binding proteins revealed fast association and dissociation rates and mean K_D-values in the low to medium μM range. The class I PDZ consensus binding motif at the positions 0 and –2 has been revealed to be necessary but not sufficient for MAGI1-PDZ1 binding. It had previously been shown that a valine residue at position 0 was more advantageous for binding the human Dlg protein than a leucine residue (29). By comparison of the position 0 of HPV 33, 45 in respect to their binding affinities this cannot be confirmed for MAGI-1 PDZ1 binding. This might be explained by a deeper hydrophobic binding pocket which does not differentiate between a shorter valine or a longer leucine residue, but has to be structurally confirmed.

E148 at position -3 is a highly conserved residue. Its substitution by Q increases the K_D-value by a factor 10. In addition a positively charged patch at position –4 to –6 seems to be necessary for binding specificity. A key residue might be at position –5 as the positive charge is strictly conserved among all binding peptides. The C-terminal peptide bearing the consensus binding motif and a negative charge at position –3 but lacking these positive charges failed binding. Reduction of the net positive charge and substitution with Pro residues reduces binding affinity. Therefore E6 behaves like the cellular Rho family nucleotide exchange factor mNET1 (56) and could possibly compete with it for binding nuclear MAGI-1c and degrading it., thereby possibly deregulating some signalling pathways.

We showed further that the five amino acid peptide RETQL still binds to MAGI-1 PDZ1 but in a different way than the eleven amino acid peptide bearing all necessary residues for specificity or the E6-C domain. This indicates that PDZ-protein interaction follows the biological evolutionary rules of a "box", a well defined sequence in a protein with conserved residues in precise positions which allow binding, where additional residues at key position outside of this box could be modulated according to the specific needs of the interaction. In our case, this box would correspond to the class I consensus binding motif with the strongly conserved amino acids at the positions 0 and –2, which define a pool of PDZ-domains to be bound. Selectivity is then added by increasing or decreasing the affinity to certain PDZ domains of this pool by modulating residues at the key positions –1 and –3 to –6 or even beyond, as has been shown in other publications (37,40). This principle allows a high variety of specific interactions on a common basic principle and makes PDZ domains one of the most abundant protein-protein interaction domains found in eucaryotic cells. The efficient biochemical methods to produce protein domains combined with the powerful peptide based Biacore approach for ligand screening and affinity determination, described in this publication, makes it possible to study any protein-PDZ. NMR experiments allow to explain the results in a structural manner. Therefore it will be possible to study more in detail the interaction of E6 with MAGI1-PDZ as well as to rapidly screen for inhibitors or competitors, opening therapeutic perspectives on the oncogenic activities of the HPV E6 oncoprotein.

ACKNOWLEDGMENTS

FIGURE LEGENDS :

FIG. 1. **Phasing of MAGI1-PDZ1 domain in order to obtain a monodisperse and well folded protein domain.** A/ Structural alignment of four class I PDZ domains with known structures (PSD95-PDZ3 (25), HDLG-PDZ3 (26), NHERF-PDZ1 (57) and PSD95-PDZ1 (58) against the residues 456-580 of human MAGI1-PDZ1. The amino acid sequences correspond to the constructs used for structure determination. Black arrows represent beta strands and black squares alpha helices. The colours correspond to: magenta-hydrophobic; red-negatively charged; green-positively charged; yellow-polr; grey-proline. The residues implicated into secondary structure elements are outlined with a black box according to the calculated structures. The minimal amino acid sequence of a PDZ domain is highlighted by the big black box. B/ Design of the 5 different PDZ1 constructs. For PDZ1-1$_{(468\text{-}566)}$ a proline-rich region was added at the C-terminus of the minimal PDZ1-0$_{(468\text{-}555)}$ domain. PDZ1-2$_{(460\text{-}566)}$ results in adding a N-terminal elongation until the next hydrophobic patch in the primary sequence to the PDZ1-1$_{468\text{-}566)}$. PDZ1-3$_{(468\text{-}580)}$ results in adding a C-terminal sequence to PDZ1-1$_{(468\text{-}566)}$ until the next proline. PDZ1-4$_{(460\text{-}580)}$ results in adding the same C-terminal elongation to PDZ1-2$_{(460\text{-}566)}$. For PDZ1-5$_{(456\text{-}580)}$ two hydrophobic residues (Phe), a proline and a lysine were added to the N-terminus of the PDZ1-4$_{(460\text{-}580)}$ construct.

FIG. 2. **Assaying protein solubility and monodispersity.** The minimal PDZ1-0 domain as well as the PDZ1-constructs, originating from the domain phasing, were produced as MBP-fusions and micropurified on affinity resin. One fraction of this protein preparation was subjected to proteolytic cleavage, the digested sample centrifuged and both pellet and supernatant resuspended in equal volumes and subsequently analyzed by SDS-PAGE. Lanes 1 and 2 : pellet (P) and supernatant (S) of PDZ1-0$_{(468\text{-}555)}$; lanes 3 and 4 : (P) and (SN) of PDZ1-1$_{(468\text{-}566)}$; lanes 5 and 6 : (P) and (SN) of PDZ1-2$_{(460\text{-}566)}$; lanes 7 and 8 : (P) and (SN) of PDZ1-3$_{(468\text{-}580)}$; lanes 9 and 10 : (P) and (SN) of PDZ1-4$_{(460\text{-}580)}$; lanes 11 and 12 : (P) and (SN) of PDZ1-5$_{(456\text{-}580)}$. B, a second fraction of the MBP-PDZ1 constructs was used to determine the aggregation rate (Ragg) by static light scattering (45). The values for each PDZ1 construct are labelled 0 to 5, according to the PDZ-contruct and are plotted below the corresponding pellet/supernatant gel lanes.

FIG. 3. **Pure-pure MBP-holdup assay to test the E6-binding ability of the soluble and monodisperse PDZ1 domains.** Two equal batches of maltoheptaose resin presaturated with either MBP-PDZ1-3, -4 or –5 were incubated with purified E6 (6C/6S) protein. In the (+M)-batch the MBP-fusion was eluted from the resin by addition of maltose, in the (-M)-batch no elution was done and the liquid phase of each batch was extracted. Protein-protein interaction was detected upon disappearance of the E6 band in the liquid phase of the (-M)-batch, where E6 was retained by the MBP-PDZ1 construct. MBP-E6 was used as a negative control, as E6 does no bind to itself. Lane 1: MBP-PDZ1-3 + E6 - maltose; lane 2: same reaction + maltose; lane 3: MBP-PDZ1-4 + E6 - maltose; lane 4: same reaction + maltose; lane 5: MBP-PDZ1-5 + E6 - maltose; lane 6: same reaction + maltose; lane 7: MBP-E6 + E6 – maltose; lane 8 same reaction + maltose.

FIG. 4. **HSQC spectra and binding curves of the soluble and active PDZ1-3$_{(468\text{-}580)}$, PDZ1-4$_{(460\text{-}580)}$, PDZ1-5$_{(456\text{-}580)}$ constructs.** A, ^{1}H-^{15}N HSQC spectra were recorded at pH 6.8 and 298 K. PDZ1 concentration was adjusted to 200 µM. B, 700 RU of GST-none and GST-HPV16e6ct peptide were alternatively immobilized on a CM5 sensor chip surface and 1 µM of PDZ1-3$_{(468\text{-}580)}$, 10 µM of PDZ1-4$_{(460\text{-}580)}$ and 20 µM of PDZ1-5$_{(456\text{-}580)}$ were injected. Sensorgrams were recorded at 25 °C. Black curves: experimental data. Grey curve: fit of the experimental data for PDZ1-3$_{(468\text{-}580)}$ to a simple bimolecular Langmuir model.

FIG. 5. **^{1}H-^{15}N-HSQC spectra of ^{15}N-labelled HPV16-E6-C in the absence or presence of PDZ1-5$_{(456\text{-}580)}$.** A, chemical shift analysis of the HSQC spectrum of ^{15}N-labelled E6-C (E6-C*) upon PDZ1-5$_{(456\text{-}580)}$ binding. Spectra were recorded at pH 6.8 and 288 K. Black spectrum: E6-C* (50 µM). Red spectrum: E6-C* (50 µM) in presence of 150 µM of PDZ1-5$_{(456\text{-}580)}$. Amino acids showing chemical shifts were assigned to four different clusters. Residues belonging to the first cluster are typed in bold

and underlined. Amino acids belonging to the other clusters are underlined. For peak assignment see (52). B, highlighting of amino acids which experience chemical shifts on the amino acid sequence of HPV16-E6-C (6C/6S). Secondary structure elements are indicated as black arrows for beta strands and black squares for alpha-helices. Amino acids of the first cluster are outlined with a black box and are underlayed gray, amino acids of the second cluster are outlined with a black box, amino acids of the third cluster are underlayed with grey, amino acids of the fourth cluster are underlined and two additional residues are typed in grey.

FIG. 6. **^{1}H-^{15}N-HSQC spectra of ^{15}N-labelled PDZ1-5*$_{(456-580)}$ in the absence or presence of HPV16-E6-C.** Spectra were recorded at pH 6.8 and 298 K. Black spectrum: PDZ1-5*$_{(456-580)}$ (150 µM). Red spectrum: PDZ1-5*$_{(456-580)}$ (150 µM) in presence of 250 µM of E6-C. For peak assignment see (46).

FIG. 7. **Comparative ^{1}H-^{15}N-HSQC spectra of ^{15}N-labelled PDZ1-5*$_{(456-580)}$ in presence of different peptidic ligands.** Spectra were recorded at pH 6.8 and 298 K. A, black spectrum: PDZ1-5*$_{(456-580)}$ (70 µM). Blue spectrum: PDZ1-5*$_{(456-580)}$ (70 µM) in presence of 200 µM of e6ct(11) peptide. B, black spectrum PDZ1-5*$_{(456-580)}$ (70 µM). Green spectrum: PDZ1-5$_{(456-580)}$ (70 µM) in presence of 200 µM of e6ct(5) peptide. C, red spectrum: PDZ1-5$_{(456-580)}$ (150 µM) in presence of 250 µM of E6-C. Blue spectrum: PDZ1-5*$_{(456-580)}$ (70 µM) in presence of 200 µM of e6ct(11) peptide. Green spectrum: PDZ1-5*$_{(456-580)}$ (70 µM) in presence of 200 µM of e6ct(5) peptide. D, magnification of a representative window of the superposed spectra in FIG. 7C.

FIG. 8. **Kinetic analysis of different C-terminal peptides.** C-terminal peptides bearing class I PDZ domain binding consensus motifs, were expressed as GST fusions and fixed as ligands on a CM5 sensor chip surface via a coupled anti GST antibody. Injection of PDZ1-5$_{(456-580)}$ as an analyte at different concentrations and subsequent steady state fitting on 1:1 Langmuir binding isotherms allowed to evaluate the dissociation constant at equilibrium (K_D). A, representative data set of GST-HPV16e6ct peptide. Injected PDZ1-5$_{(456-580)}$ concentrations were from downwards to upwards 0.5, 1, 2, 5, 10, !5, 20, 30 and 60 µM. B, representative data set of the GST-HPV11e6ct peptide as a negative control. PDZ1-5$_{(456-580)}$ concentrations were the same than in Fig. 8A. C, Analysis of the interaction of various C-terminal GST-peptides against PDZ1-5$_{(456-580)}$. For each peptide a data set was recorded as shown in Fig. 8A. Equilibrium responses (Req) were plotted as a function of total PDZ1-5$_{(456-580)}$ concentration and fit to simple 1:1 binding isotherms. Mean equilibrium dissociation constants are reported in Table I. Tested GST-peptides are from upwards to downwards: black triangles: HPV45e6ct; black squares: HPV18e6ct; black circles: HPV31e6ct; grey crosses: mNETct; open circles:HPV33e6ct; open squares: HPV58e6ct; black lozenge: HPV16e6ct; open lozenge: HPV52e6ct; open triangle: HPV11e6ct; grey circles: hpTENct; grey triangles: hb2Ar2t.

FIG. 9. **A/ Sequence alignment of the C-terminal peptides.** The C-terminal peptides were aligned against each other with decreasing affinity. The amino acids of the class I PDZ binding consensus motif as well as possible key residues confering binding specificity are outlined with black boxes. The amino acid positions, according to the common PDZ consensus binding motif nomenclature, are written on top of the alignment. **B/ Kinetic analysis of the interaction of various C-terminal peptides to PDZ1-5$_{(456-580)}$.** C-terminal peptides of the high risk HPV's 16, 18, 31, 33, 45, 52, 58, the low risk HPV 11 as well as the proteins hmNET, hpTen and hbAR2, were produced as GST-fusions and captured on a CM5 sensorchip pre-immobilized with anti GST antibody. Injection of different concentrations of PDZ1-5$_{(456-580)}$ allowed the evaluation of the individual equilibrium dissociation konstants (K_D). Each experiment was repeated at least two times. Individual K_D-values are listed for each GST peptide. The mean K_D-values and the standard deviation are listed in bold type. Nb: no binding; %: experiment not repeated.

1. zur Hausen, H. (1991) *Virology* **184**, 9-13
2. Dell, G., and Gaston, K. (2001) *Cell Mol Life Sci* **58**, 1923-1942
3. zur Hausen, H., and de Villiers, E. M. (1994) *Annu Rev Microbiol* **48**, 427-447
4. de Villiers, E. M. (1994) *Curr Top Microbiol Immunol* **186**, 1-12
5. Bosch, F. X., and de Sanjose, S. (2002) *Curr Oncol Rep* **4**, 175-183
6. Clifford, G. M., Smith, J. S., Plummer, M., Munoz, N., and Franceschi, S. (2003) *Br J Cancer* **88**, 63-73
7. Parkin, D. M., Bray, F. I., and Devesa, S. S. (2001) *Eur J Cancer* **37 Suppl 8**, S4-66
8. Pecorelli, S., Favalli, G., Zigliani, L., and Odicino, F. (2003) *Int J Gynaecol Obstet* **82**, 369-379
9. Schwarz, E., Freese, U. K., Gissmann, L., Mayer, W., Roggenbuck, B., Stremlau, A., and zur Hausen, H. (1985) *Nature* **314**, 111-114
10. Scheurer, M. E., Tortolero-Luna, G., and Adler-Storthz, K. (2005) *Int J Gynecol Cancer* **15**, 727-746
11. Song, S., Pitot, H. C., and Lambert, P. F. (1999) *J Virol* **73**, 5887-5893
12. Kiyono, T. (1998) *Uirusu* **48**, 125-135
13. Kiyono, T., Hiraiwa, A., Fujita, M., Hayashi, Y., Akiyama, T., and Ishibashi, M. (1997) *Proc Natl Acad Sci U S A* **94**, 11612-11616
14. Nguyen, M. L., Nguyen, M. M., Lee, D., Griep, A. E., and Lambert, P. F. (2003) *J Virol* **77**, 6957-6964
15. Simonson, S. J., Difilippantonio, M. J., and Lambert, P. F. (2005) *Cancer Res* **65**, 8266-8273
16. Lee, S. S., Weiss, R. S., and Javier, R. T. (1997) *Proc Natl Acad Sci U S A* **94**, 6670-6675
17. Pim, D., Thomas, M., Javier, R., Gardiol, D., and Banks, L. (2000) *Oncogene* **19**, 719-725
18. Gardiol, D., Kuhne, C., Glaunsinger, B., Lee, S. S., Javier, R., and Banks, L. (1999) *Oncogene* **18**, 5487-5496
19. Nakagawa, S., and Huibregtse, J. M. (2000) *Mol Cell Biol* **20**, 8244-8253
20. Glaunsinger, B. A., Lee, S. S., Thomas, M., Banks, L., and Javier, R. (2000) *Oncogene* **19**, 5270-5280
21. Lee, S., Maler, L., and Dunn, R. J. (2000) *J Comp Neurol* **426**, 429-440
22. Kornau, H. C., Schenker, L. T., Kennedy, M. B., and Seeburg, P. H. (1995) *Science* **269**, 1737-1740
23. Matsumine, A., Ogai, A., Senda, T., Okumura, N., Satoh, K., Baeg, G. H., Kawahara, T., Kobayashi, S., Okada, M., Toyoshima, K., and Akiyama, T. (1996) *Science* **272**, 1020-1023
24. Niethammer, M., Kim, E., and Sheng, M. (1996) *J Neurosci* **16**, 2157-2163
25. Doyle, D. A., Lee, A., Lewis, J., Kim, E., Sheng, M., and MacKinnon, R. (1996) *Cell* **85**, 1067-1076
26. Morais Cabral, J. H., Petosa, C., Sutcliffe, M. J., Raza, S., Byron, O., Poy, F., Marfatia, S. M., Chishti, A. H., and Liddington, R. C. (1996) *Nature* **382**, 649-652
27. Harris, B. Z., and Lim, W. A. (2001) *J Cell Sci* **114**, 3219-3231
28. Hung, A. Y., and Sheng, M. (2002) *J Biol Chem* **277**, 5699-5702
29. Thomas, M., Glaunsinger, B., Pim, D., Javier, R., and Banks, L. (2001) *Oncogene* **20**, 5431-5439
30. Thomas, M., Laura, R., Hepner, K., Guccione, E., Sawyers, C., Lasky, L., and Banks, L. (2002) *Oncogene* **21**, 5088-5096
31. Dobrosotskaya, I., Guy, R. K., and James, G. L. (1997) *J Biol Chem* **272**, 31589-31597
32. Ide, N., Hata, Y., Deguchi, M., Hirao, K., Yao, I., and Takai, Y. (1999) *Biochem Biophys Res Commun* **256**, 456-461
33. Patrie, K. M., Drescher, A. J., Welihinda, A., Mundel, P., and Margolis, B. (2002) *J Biol Chem* **277**, 30183-30190
34. Hirabayashi, S., Mori, H., Kansaku, A., Kurihara, H., Sakai, T., Shimizu, F., Kawachi, H., and Hata, Y. (2005) *Lab Invest* **85**, 1528-1543
35. Kotelevets, L., van Hengel, J., Bruyneel, E., Mareel, M., van Roy, F., and Chastre, E. (2005) *Faseb J* **19**, 115-117
36. Sheng, M., and Sala, C. (2001) *Annu Rev Neurosci* **24**, 1-29
37. Cai, C., Coleman, S. K., Niemi, K., and Keinanen, K. (2002) *J Biol Chem* **277**, 31484-31490
38. Skelton, N. J., Koehler, M. F., Zobel, K., Wong, W. L., Yeh, S., Pisabarro, M. T., Yin, J. P., Lasky, L. A., and Sidhu, S. S. (2003) *J Biol Chem* **278**, 7645-7654
39. Wiedemann, U., Boisguerin, P., Leben, R., Leitner, D., Krause, G., Moelling, K., Volkmer-Engert, R., and Oschkinat, H. (2004) *J Mol Biol* **343**, 703-718
40. Kang, B. S., Cooper, D. R., Devedjiev, Y., Derewenda, U., and Derewenda, Z. S. (2003) *Structure* **11**, 845-853
41. Mantovani, F., and Banks, L. (2001) *Oncogene* **20**, 7874-7887
42. Nomine, Y., Charbonnier, S., Ristriani, T., Stier, G., Masson, M., Cavusoglu, N., Van Dorsselaer, A., Weiss, E., Kieffer, B., and Trave, G. (2003) *Biochemistry* **42**, 4909-4917
43. Nomine, Y., Masson, M., Charbonnier, S., Zanier, K., Ristriani, T., Deryckere, F., Sibler, A. P., Desplancq, D., Atkinson, R. A., Weiss, E., Orfanoudakis, G., Kieffer, B., and Trave, G. (2006) *Mol Cell* **21**, 665-678

44.   Nomine, Y., Ristriani, T., Laurent, C., Lefevre, J. F., Weiss, E., and Trave, G. (2001) *Protein Expr Purif* **23**, 22-32

45.   Nomine, Y., Ristriani, T., Laurent, C., Lefevre, J. F., Weiss, E., and Trave, G. (2001) *Protein Eng* **14**, 297-305

46.   Charbonnier, S., Zanier, K., Masson, M., and Trave, G. (2006) *Protein Expr Purif*

47.   Zanier, K., Charbonnier, S., Baltzinger, M., Nomine, Y., Altschuh, D., and Trave, G. (2005) *J Mol Biol* **349**, 401-412

48.   Piotto, M., Saudek, V., and Sklenar, V. (1992) *J Biomol NMR* **2**, 661-665

49.   Delaglio, F., Grzesiek, S., Vuister, G. W., Zhu, G., Pfeifer, J., and Bax, A. (1995) *J Biomol NMR* **6**, 277-293

50.   Eccles, C., Guntert, P., Billeter, M., and Wuthrich, K. (1991) *J Biomol NMR* **1**, 111-130

51.   Pfuhl, M., Chen, H. A., Kristensen, S. M., and Driscoll, P. C. (1999) *J Biomol NMR* **14**, 307-320

52.   Nomine, Y., Charbonnier, S., Miguet, L., Potier, N., Van Dorsselaer, A., Atkinson, R. A., Trave, G., and Kieffer, B. (2005) *J Biomol NMR* **31**, 129-141

53.   van Ham, M., and Hendriks, W. (2003) *Mol Biol Rep* **30**, 69-82

54.   Meyer, G., Varoqueaux, F., Neeb, A., Oschlies, M., and Brose, N. (2004) *Neuropharmacology* **47**, 724-733

55.   Schultz, J., Hoffmuller, U., Krause, G., Ashurst, J., Macias, M. J., Schmieder, P., Schneider-Mergener, J., and Oschkinat, H. (1998) *Nat Struct Biol* **5**, 19-24

56.   Dobrosotskaya, I. Y. (2001) *Biochem Biophys Res Commun* **283**, 969-975

57.   Karthikeyan, S., Leung, T., Birrane, G., Webster, G., and Ladias, J. A. (2001) *J Mol Biol* **308**, 963-973

58.   Piserchio, A., Salinas, G. D., Li, T., Marshall, J., Spaller, M. R., and Mierke, D. F. (2004) *Chem Biol* **11**, 469-473

Figure 1

A

PSD95_PDZ3	GSPEFLGEEDIPREPRRIVIHRGST.GLGFNIIGGEDG.......EGIFISFILAGGPADLSGELR
HDLG_PDZ3	.......DDEITREPRKVVLHRGST.GLGFNIVGGEDG.......EGIFISFILAGGPADLSGELR
SHANK	..........RMLPRLCCLEKGPN.GYGFHLH.GEKGKL......QYIRLVEPGSPAEKAGLLA
PSD95_PDZ1	.............EYEEITLERGNS.GLGFSIAGGTDNPHIGDDPSIFITKII.PGGAAAQDGRLR
MAGI1_PDZ1	.KPFFTRNPSELKGKFIHTKLRKSSRGFGFTVVGGDEPDE.....FLQIKSLVLDGPAALD.GKME

PSD95_PDZ3	KGDQILSVNGVDLRNASHEQAAIALKNAGQTVTIIAQY.KPEEYSRFEANSRVNSSGRIVTN....
HDLG_PDZ3	KGDRIISVNSVDLRAASHEQAAAALKNAGQAVTIVAQY.RPEEYSRFEA................
SHANK	.GDRLVEVNGENVEKETHQQVVSRIRAALNAVRLLVVD..PETDEQL.................
PSD95_PDZ1	VNDSILFVNEVDVREVTHSAAVEALKEAGSIVRLYVMRRKPP.....................
MAGI1_PDZ1	TGDVIVSVNDTCVLGHTHAQVVKIFQSIPIGASVDLELCRGYPLPFDPDDPNTSLVTSVAILDKEP

B

KPFF TRNPSELK GKF... ...CRG YPLPFDPDDPN TSLVTSVAILDKEP

456   460   468   556   567   580

PDZ1-0(468-555)

PDZ1-1(468-566)

PDZ1-2(460-566)

PDZ1-3(468-580)

PDZ1-4(460-580)

PDZ1-5(456-580)

Figure 2

A

B

Figure 3

Figure 4

A

PDZ1-3   **PDZ1-4**   **PDZ1-5**

^{15}N (ppm)

^{1}H (ppm)

B

PDZ1-3   **PDZ1-4**   **PDZ1-5**

Time (s)
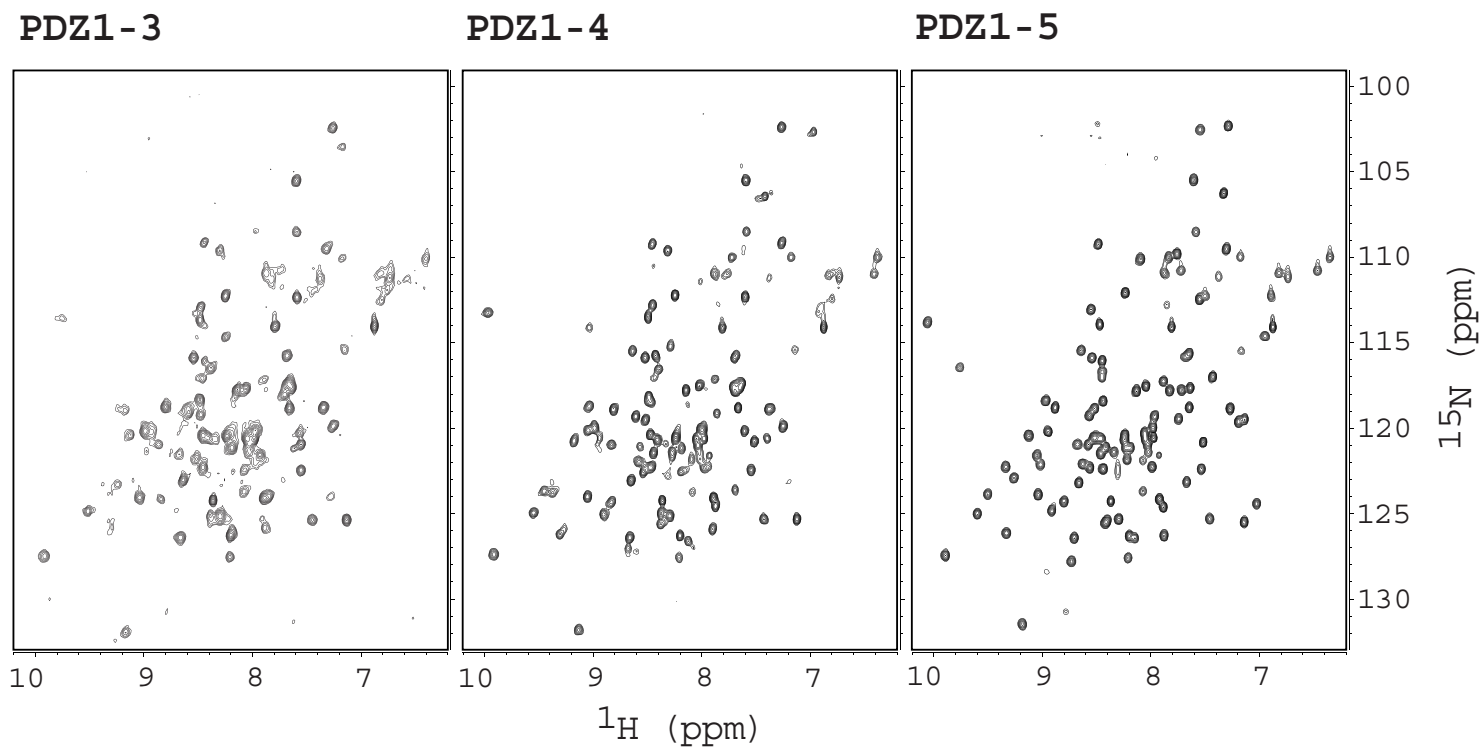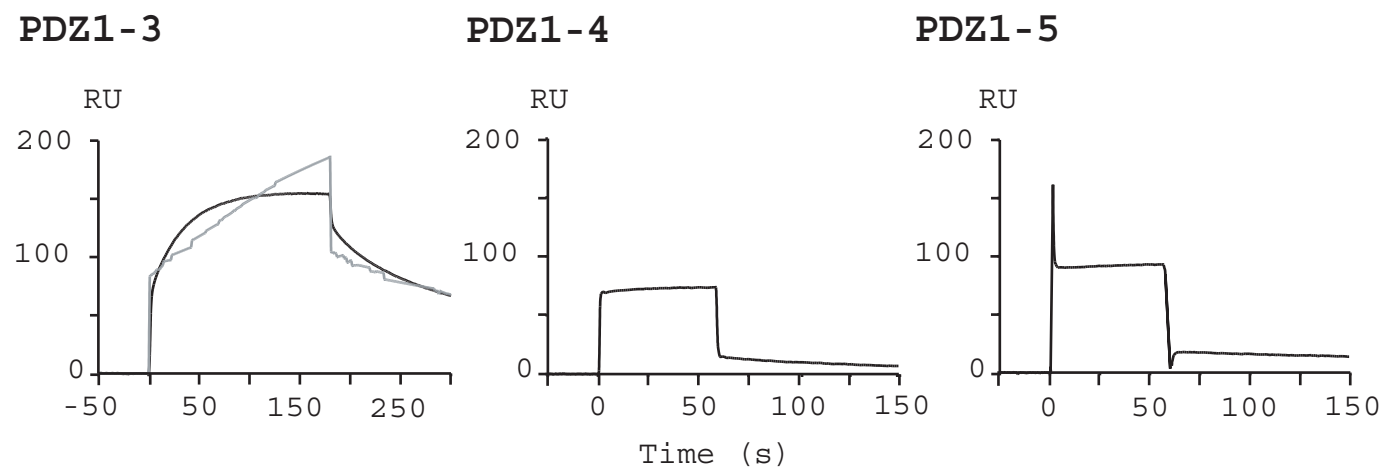
Figure 5

A
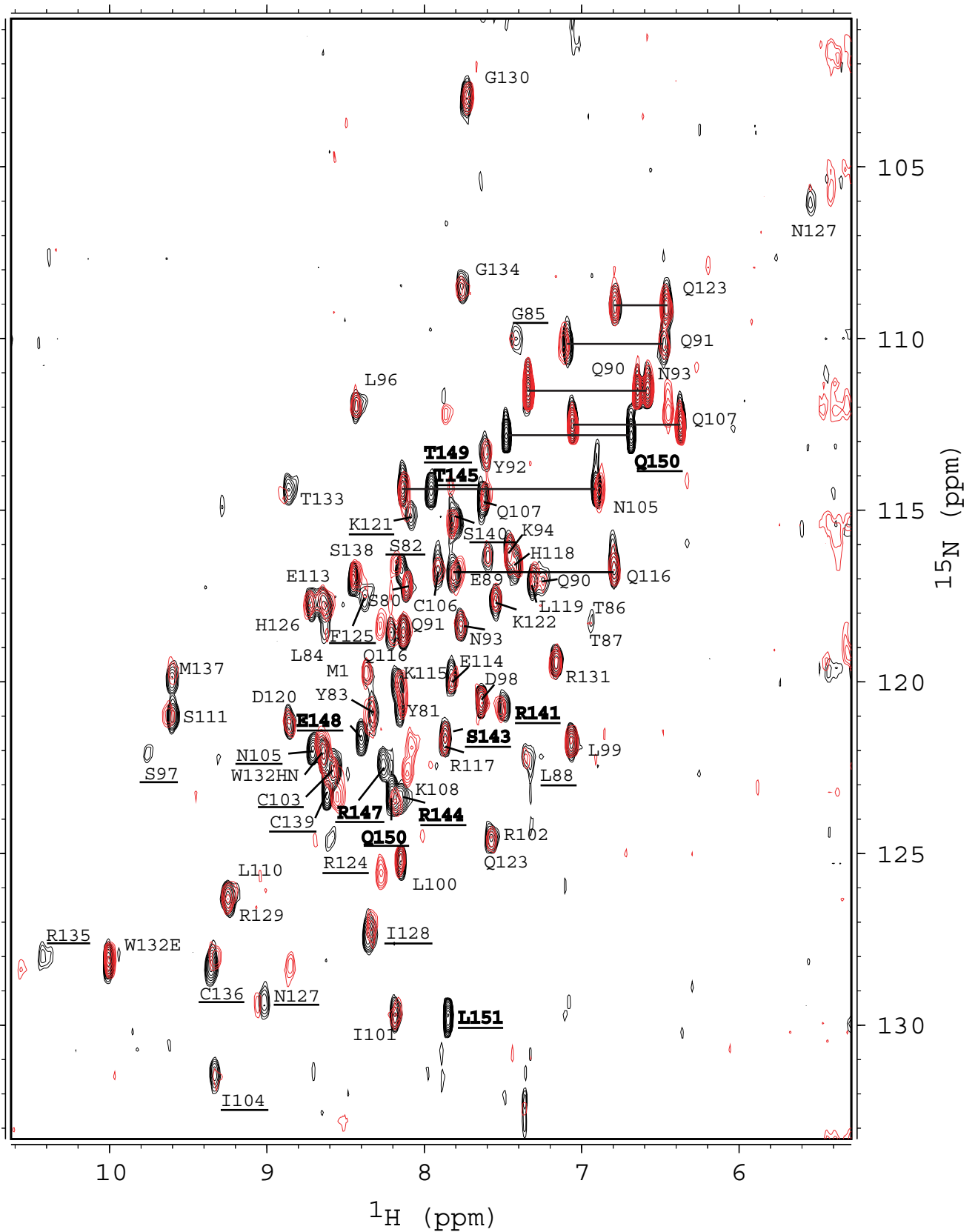
B

Figure 6

Figure 7

Figure 8

A

B

C
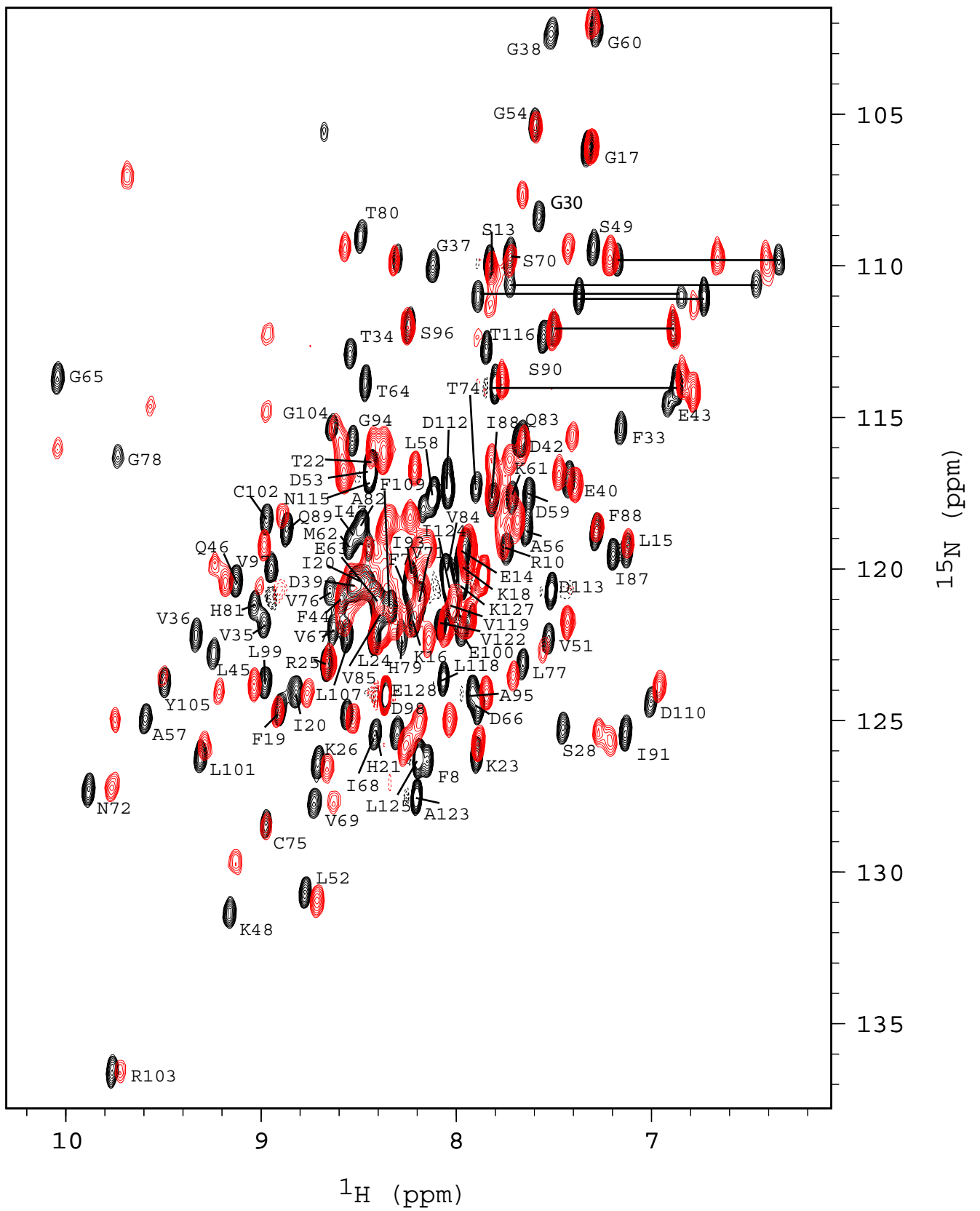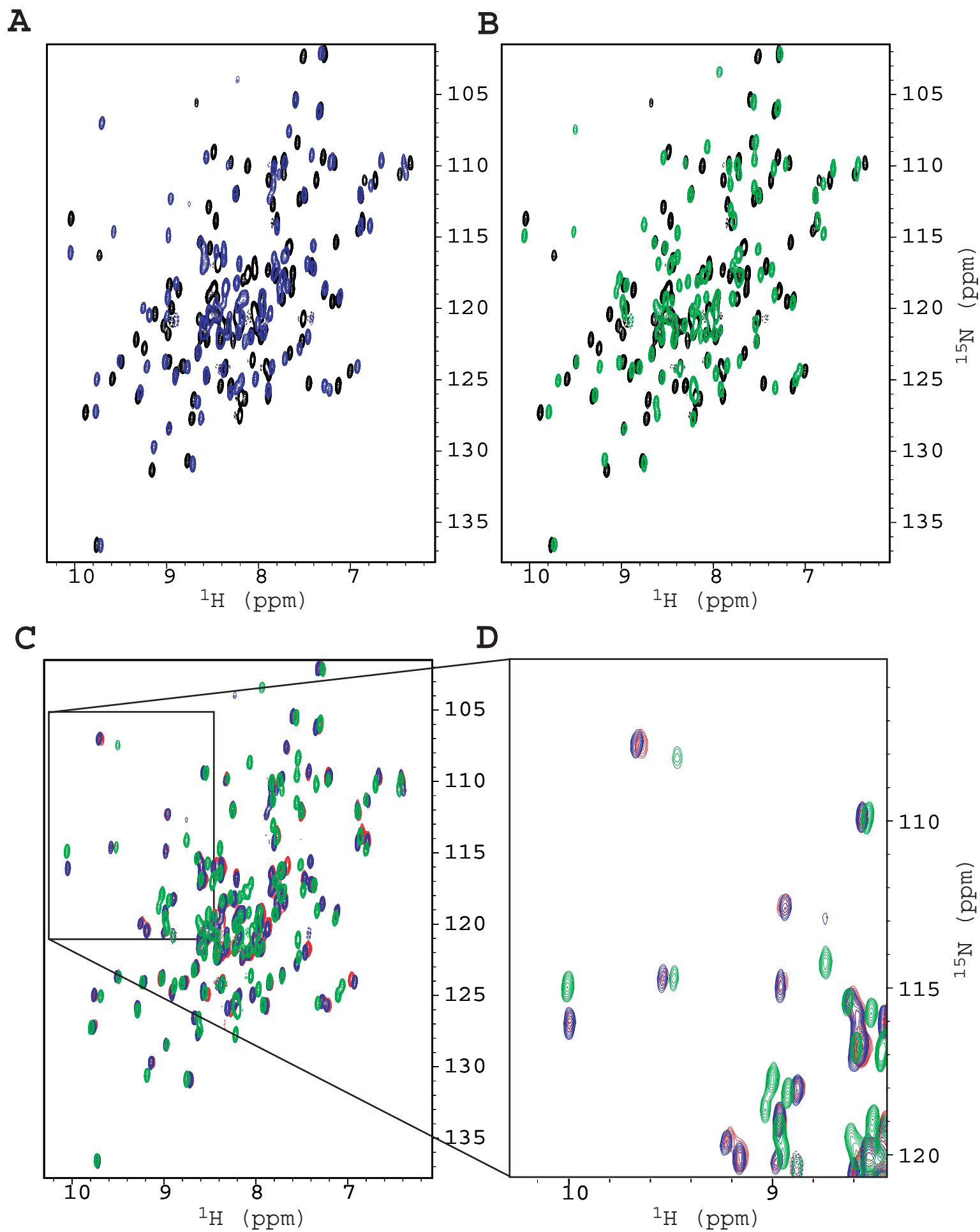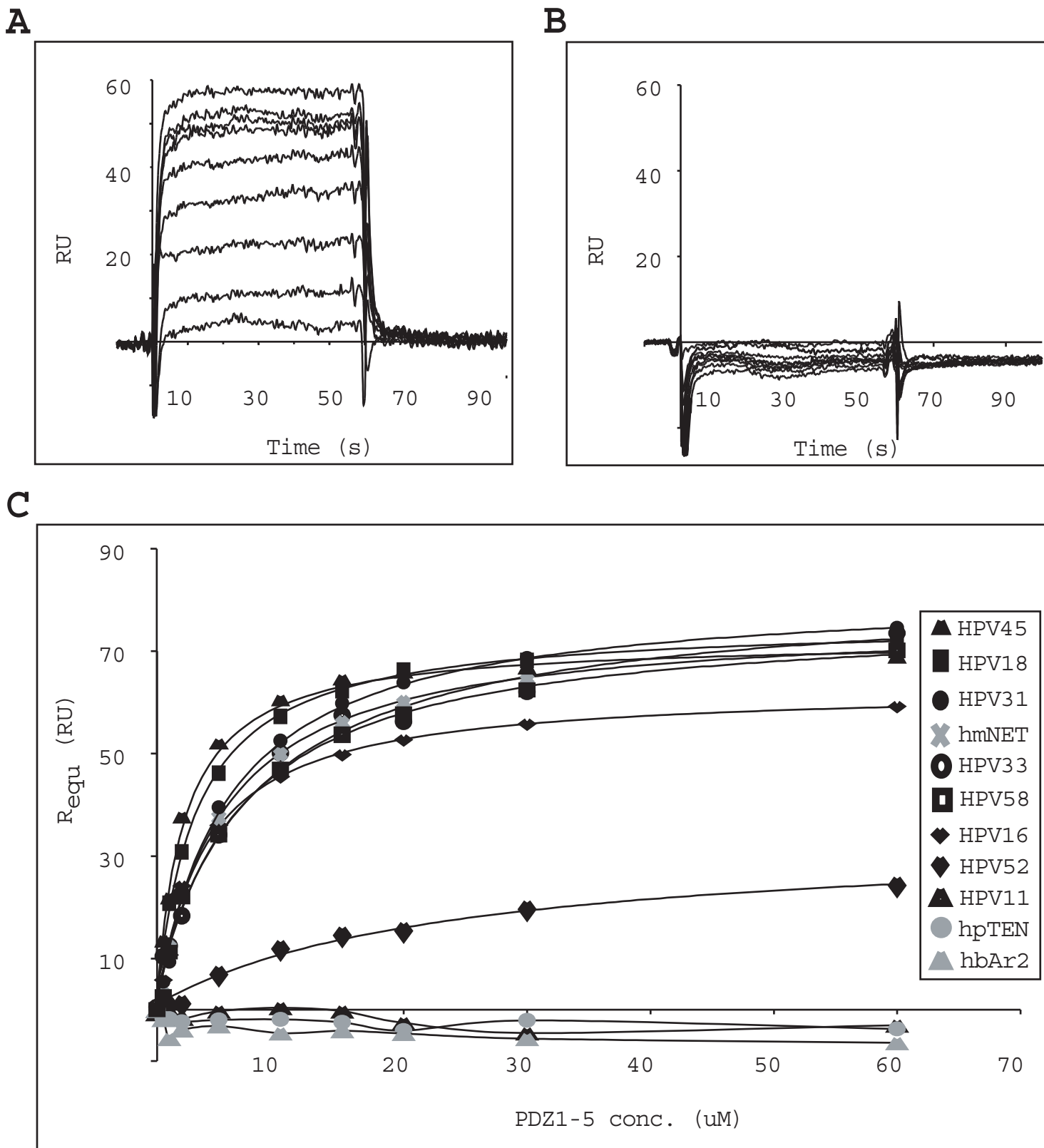
Figure 9

A
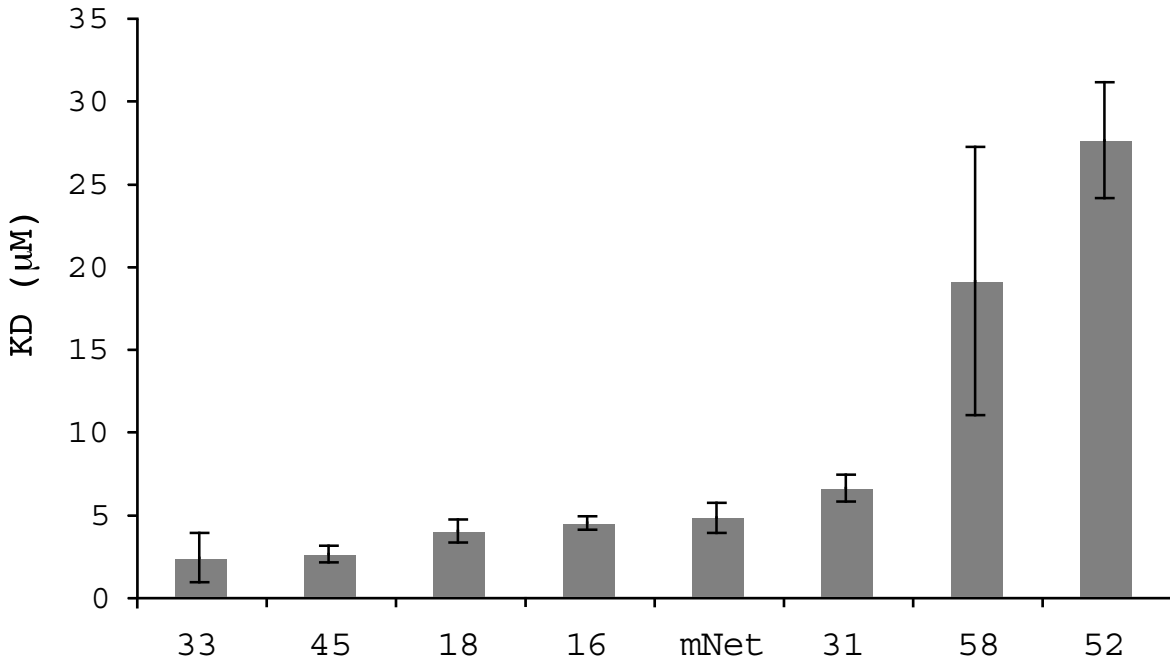
B

```
                                 -8 -6 -4 -2  0
HPV33e6ct              WRS R R R E T A L
HPV45e6ct     DQARQERLR R R R E T Q V
HPV18e6ct     NRARQERLQ R R R E T Q V
HPV16e6ct        SRSSRT R R E T Q L
hmNETct            QSGG K K K E T L V
HPV31e6ct          WRRP R T E T Q V
HPV58e6ct          WRP R R R Q T Q V
HPV52e6ct          WRP R P V T Q V
HPV11e6ct          WTTCMEDLLP
hpTENct        EPFDEDQHTQI T K V
hbAr2ct        DSQGRNCSTND S L L
```

Letter to the Editor

^{13}C, ^{15}N and ^{1}H resonance assignment of the PDZ1 domain of MAGI-1 using QUASI

Human Membrane Associated Guanylate kinase with Inverted domain structure-1 (MAGI-1) is localised predominantly at tight junctions of epithelial cells and is described as a scaffolding protein. The PDZ1 domain of human MAGI-1 binds to the oncoprotein E6 from human papillomavirus 16 (HPV-16) via a C-terminal consensus motif and mediates E6 degradation by the proteasome. This is thought to contribute to HPV16 E6-mediated cervical carcinogenesis (Thomas et al., 2001). The PDZ1 domain of human MAGI-1 (amino acids 456–580) was cloned into a pETM41 expression vector and uniformly ^{13}C,^{15}N-labelled by over-expression in *E. coli* (BL21-DE3) in M9 minimal medium supplemented with^{15}NH$_4$Cl and ^{13}C$_6$-glucose. From NMR experiments acquired at 22 °C on a Bruker DRX600 spectrometer, backbone ^{1}H, ^{15}N and ^{13}C assignments were obtained using QUASI (Coutouly et al., 2004) and side-chain resonances completed manually. Assignments are complete except for broad amide resonances of S24-G28, side-chains of K44, K57, E59, H75, R99 and some aromatic resonances. BMRB accession number: 6911.
References: Thomas et al. (2001) *Oncogene*, **20**, 5431–5439; Coutouly et al. (2004) *Comptes Rendus Chim.*, **7**, 335–341

Sebastian Charbonnier[a], Marie-Aude Coutouly[b], Bruno Kieffer[b], Gilles Travé[a] & R. Andrew Atkinson[b],*
[a]*Equipe Oncoprotéines, Institut Gilbert Laustriat (UMR 7175-LC1);* [b]*Biomolecular NMR Group, IGBMC (UMR 7104) ESBS, 67412, Illkirch, France*

*To whom correspondence should be addressed. E-mail: atkinson@titus.u-strasbg.fr

CONCLUSIONS & PERSPECTIVES

To fold or not to fold...

Production of proteins, whether for biochemical analysis, therapeutics or structural investigations, is dependent of three major individual factors: 1) expression, 2) solubility and monodispersity and 3) purification. Nowadays, protein expression and purification techniques have improved considerably and are in general no longer considered as limiting steps. The main bottleneck in the field remains the production of properly folded, monodisperse and soluble protein samples, especially when required in high concentration as for NMR studies. In fact as the number of high-throughput structural genomics projects increases, reported percentages of soluble heterologous proteins expressed in bacterial hosts (mainly *E. coli* strains) decrease continuously (Chambers et al., 2004). The reasons why it is difficult to express soluble proteins in bacteria, especially mammalian proteins, are multiple and not well known. Besides the obvious differences between bacteria and eucaryotic cells (S70 ribosomes, no compartimentation, no posttranslational modifications and protein processing, different chaperones...) one main difference which certainly contributes to the aggregation problem is the translation and protein folding rate which are much faster in *E. coli* as compared to eucaryotic cells (Widmann & Christen, 2000). Although eucaryotic expression systems are sometimes able to overcome aggregation problems, they show still major drawbacks in terms of ease of use, time, cost and experimental flexibility. Protein folding problems becoming the bottleneck of *in vitro* protein investigation have led to significant research in order to enhance the production of monodisperse well folded proteins with the current bacterial expression hosts. Although part of the efforts have been directed towards optimising expression conditions a majority of the work has focused on the discovery and the development of solubility increasing fusion tags.

To date many different solubilising tags have been described (for review see: Esposito & Chatterjee, 2006; Waugh, 2005). However a multitude of commercially available vectors, presenting different features make easy one step cloning and straightforward expression tests complicated. One way to circumvent this problem was the invention of the Gateway system (Invitrogen), which allows easy cDNA transfer from a shuttle vector into destination vectors by means of homologuous recombination. However, no ready to use expression vectors containing different solubilising tags are available. The alternative is the pETM vector series, which was

designed and created by Gunter Stier (EMBL, Heidelberg). This vector series has been especially developed for high-yield fusion protein expression. They combine a strong viral promoter with a multitude of solubilising tags in a unique vector backbone allowing parallel cloning of one given cDNA into unique restriction sites. Another striking feature is that besides the solubilising tag, which can sometimes be used for affinity purification, all vectors provide a 6His-tag allowing chelating affinity purification and, in case of an affinity tag, double affinity purifications. Finally the vectors contain a TEV protease restriction site which allows specific cleavage of the solubilising tag after purification. The fact that the pETM vector kit is provided with a expression vector for recombinant TEV protease makes this system even more attractive.

The structural and kinetic study of the interaction of the HPV 16 E6 oncoprotein with MAGI-1, presented in this manuscript, had to deal with major protein aggregation problems and needed extensive optimisation for generating well folded protein domains. We addressed these aggregation problems by different strategies according to the protein, among which domain phasing and primary sequence modifications at non-conserved positions. Besides that, screening to find the optimal solubilising tag and the optimal expression conditions was required. Introducing the pETM vector series with the afore mentioned features enabled us by applying a broad parallel screening approach to identify MBP as the optimal solubilising tag and for the first time to lower the expression temperature in order to produce well folded protein at NMR concentration, for both E6-C and MAGI-1 PDZ1. Protein quality, i.e. proper folding, was systematically monitored. However MBP fusions tend to form soluble inclusion bodies, rendering classical pellet-supernatant-based quality assays unreliable.

To address this problem and in order to get qualitative information about the proper folding, Nominé et al. invented a method based on static light scattering. The derived aggregation ratio gives information about the monodispersity of the analysed protein suspension (Nomine et al. 2001b). This assay is very sensitive and reliable, yet the present work has shown that three slightly different PDZ1 domain constructions showing identical aggregation rate values displayed significant differences at an atomic level concerning the stability of the protein fold when analysed by NMR.

NMR technology has made remarkable progress in the past years especially concerning sensitivity of the NMR probes and in the routine use of certain pulse sequences. In parallel isotopic labelling of proteins has become both affordable and routine laboratory practice. ^{1}H-^{15}N HSQC NMR experiments are suited to rapidly provide information at atomic scale about protein fold and molecular dynamics of the protein. In the present work I show that minor modifications of the domain boundaries (or even point mutations of residues which are expected to be exposed on the protein structure (data not shown)) alter significantly the stability of the fold. This becomes only visible, when analysed by 2D-HSQC and cannot be detected by classical solubility or monodispersity tests. Therefore small-scale production of isotopically labelled protein and systematic HSQC analysis should be included in routine protein quality monitoring, in case NMR equipment is available.

Meet and greet...

Protein-protein interactions are dynamic processes. Knowledge about the kinetic parameters which drive a given protein-ligand interaction is therefore an essential information. In the specific case study of the E6 oncoprotein binding to PDZ domains only little information was available. This was certainly related to the difficulties in producing E6 and handling it during *in vitro* assays. To overcome the latter problem we optimised a GST-peptide based Biacore method for measuring interaction kinetics of GST peptides against E6 or PDZ domains. The experimental setup combines low cost and medium throughput screening and provides real-time access to all kinetic parameters, including the kinetic association and dissociation rates.

Thereby Zanier et al. studied in detail the interaction kinetics of the E6AP minimal binding peptide of E6AP with E6. I measured the affinities of several C-terminal GST peptides derived from different high- and low-risk E6 proteins with MAGI-1 PDZ1 and identified important residues which contribute to binding affinity and specificity.

The Biacore technology, being a very sensitive tool, requires to work with homogeneous and monodisperse protein suspensions, especially when measuring low affinity interactions. Control of protein quality and monodispersity during sample preparations is therefore a crucial aspect in all interaction assays. Our knowledge in

this field and the optimisations we propose allow us therefore to guarantee a high degree of purity and monodispersity of the protein suspensions. This enables us to assay and measure even interactions in the µM range with accuracy.

The here described Biacore method presents a great advance in both the study of high-risk HPV E6 oncoproteins and for the query of therapeutical molecules directed against E6. Virtually any compounds susceptible to bind and maybe alter or block oncogenic activites of E6 can be screened and the affinity measured. It could also be used to optimise already existing binders via in vitro evolution or by more rational approaches.

However, although this approach has been designed to lower the cost by allowing to test several interactions on the same sensorchip, it remains still a cost intensive method requiring special equipment. Based on an idea of Gilles Travé, I set up a very simple, yet powerful method, named holdup assay. Comparative chromatographic retention allows easy qualitative and, with some optimisation, quantitative analysis of protein-protein interactions. This assay is powerful for both technical and experimental reasons: first it requires only basic laboratory equipment and can therefore be performed in every standard laboratory; second it allows to analyse all reaction species in the binding equilibrium with sufficient accuracy, enabling accurate estimations of mean K_D values. Besides the specific interactions presented in the publication we have already validated this assay in a screening approach where we identified E6-binding GST-peptides out of 35 potential binders. All peptides identified via holdup were confirmed via Biacore (data to be published). Another advantage of this method in such a screening approach is its principle of comparative analyte depletion in small volumes and in presence of high protein concentrations without washing steps. This allows to work at equilibrium as well as to detect medium and low affinity interactions.

A future application of the holdup could be to take advantage of this feature to detect medium and low affinity binders by implementing it into a proteomic project. Holdup experiments using recombinant GST-E6, or virtually any other protein, could be performed on concentrated cell lysates and the resulting extracts analysed by 2D gel electrophoresis and standard mass spectrometry.

However, one major drawback of this method to date is that it depends much on the manipulating person, especially with respect to the accuracy of resin handling.

The recent commercialisation of GST-capturing membranes, could be a way to substitute resin handling and to increase the accuracy and reproducibility of the assay. In the near future it would be interesting to adapt this method for automation using robotic platforms. This would make qualitative high-throughput-screening possible ensuring high accuracy and reproducibility for generating quantitative data. Automation of this assay would certainly make it interesting to a broader range of potential users, especially of the industrial sector. We have already started to get into contact with potentially interested laboratories to do so.

Where atoms interact...

In parallel to the kinetic study I analysed the molecular determinants of the HPV 16 E6 / MAGI-1 PDZ1 interaction by NMR. Combinatorial titration experiments using labelled and unlabelled E6, E6-C and different C-terminal peptides of E6 showed that the interaction did not affect the folded E6-C domain and involved essentially the last 11 C-terminal residues of E6. In solution this region appeared to be unstructured in the E6-C domain in absence of a ligand (Nomine et al., 2005). These data combined with the kinetic data allowed to identify key residues important for ligand binding affinity and specificity.

For the near future the aim will be to generate the 3 dimensional structure of the MAGI-1 PDZ1 complexed to this minimal HPV 16 E6 C-terminal peptide. We have already assigned most of the ^{1}H, ^{13}C and ^{15}N frequencies of the PDZ1 domain in the free (BMRB code: 6911) and complexed form (unpublished data). We have started first structure calculations for the complexed PDZ1 domain and performed additional experiments to position some atoms of the peptide on the PDZ structure and to calculate a reliable model of the peptide-PDZ1 interaction. The structural analysis will allow to understand the precise interaction mechanism at atomic scale.

One great advantage of the NMR is that in addition to the structure it provides also information about the molecular dynamics of the protein complex.

From first dynamic experiments we have evidence, that the C-terminal elongation of the PDZ1 domain, which is essential for the solubility and the monodispersity of the domain, is unstructured in the free PDZ1 domain in solution. However, in the complex it seems to adopt a more constrained conformation and

secondary structure analysis show possible organisation in a β-strand like conformation.

In the era of structural genomics, protein domains requiring some domain phasing optimisation to be soluble and monodisperse, might be lost during the high throughput solubility screenings. Every piece of additional knowledge about the function of the elongations described in my work, could be implemented into high-throughput processes and might increase the amount of soluble protein domains to work with.

However, the most interesting part will be to investigate the dynamic behaviour of the peptide-PDZ interaction. This will generate precious information to better understand this peptide-PDZ interaction according to modulation of affinity and specificity and perhaps to allow to extract some general rules.

Once the structural data available we plan also to analyse the interaction by a rational mutagenesis study, in order to validate important residues experimentally. Complementary to this, random mutagenesis combined to an interaction screening by yeast two hybrid, holdup or Biacore could serve to identify high affinity peptidic ligands which could compete with E6 for MAGI-1 PDZ1 binding.

Concluding remarks

The ability of high-risk HPV E6 oncoprotein has been described as one essential activity for generating hyperplasias *in vivo* as well as for malignant progression into cancer. It has been shown, that E6 is more implicated into malignant progression than in promotion of cancer (Simonson et al., 2005). This seems to correlate with the hypothesis that degradation of PDZ-domain containing proteins is important in the late stages of cancer development and contributes to loss of cell polarity and inhibition contact thus driving the cells towards invasive phenotypes (Mantovani & Banks, 2001). Detailed information of the interaction of high-risk HPV E6 oncoproteins with PDZ domains is therefore of great interest in order to find new potential targets for potential therapeutic applications.

In a more general context of therapeutic applications and drug delivery, recent progress in the analysis of protein-protein interactions has shown that PDZ domain containing proteins regulate expression and/or function of drug transporters (Anzai et al., 2004; Gisler et al., 2003; Kato et al., 2004; Kato et al., 2005). Various

transporters are localised at apical membranes and thought to be incorporated into protein networks. Considering the multiple interactions between PDZ domains and cytosolic free C-termini of transmembrane transporters, PDZ proteins influence potentially transmembrane transport of a multitude of molecules. In general, the diversity of PDZ proteins, their roles in cell function and the diversity of interactions make them promising candidates for modulation by low molecular mass inhibitors (Arkin & Wells, 2004; Berg, 2003; Fisher & Lane, 2004; Fuji et al., 2003). In addition it has been shown that PDZ protein expression can be tissue specific. For these reasons PDZ domain proteins have been proposed as novel targets for drug discovery (Dev, 2004). Detailed understanding of peptide-PDZ interaction could contribute to improve or to control drug uptake in specific cell types.

Different ways to address the targeting of PDZ mediated interactions have been described. Blocking peptides have been proposed, which mimic the PDZ binding motif and occupy the targeted PDZ domain. Such blocking peptides have already been tested successfully for functional analysis of a precise neuronal PDZ protein (Daw et al., 2000; Hirbec et al., 2003; Nishimune et al., 1998). However such peptides show limited cell permeability and require delivery into the cell. Synthesis of small chemical compounds which either bind to C-terminal PDZ binding motifs or into the PDZ binding groove, could therefore be an alternative. Recently interaction between PTEN and MAGI3 has been selectively inhibited by such a chemical compound (Fujii et al., 2003). Another interesting feature of PDZ domains is their implication into different molecular complexes. Disruption of one specific PDZ interaction might hit more than one disease related protein simultaneously and could potentially improve efficacy.

On the other hand efforts have been made to isolate high affinity PDZ domains. Computer-based design of PDZ domains has been performed and yielded PDZ variants with affinity changes ranging from 1-100 µM (Reina et al., 2002). Another approach consisted in applying in vitro protein evolution techniques in order to modify and select a high affinity PDZ domain against a given peptide, yielding a PDZ domain with an affinity as high as 620 nM to the desired peptide (Ferrer et al., 2005).

Finally, this specific interaction study shows again that PDZ-protein interactions are based on a very simple, yet very modulable and highly specific principle. It can be speculated that they have evolved from a basic mechanism involving the binding of a consensus peptide into a binding groove. This simple mechanism might have evolved towards creation of specific binding pockets in the binding groove of the PDZ domain combined with differentiation of distinct peptidic consensus sequences. Finally addition or subtraction of additional amino acid determinants to this consensus increased the flexibility of this mechanism.

PDZ domains are the most abundant protein interaction domains and implicated into a multitude of cellular pathways, rendering these family of protein domains an interesting target for possible therapeutic applications.

This work clearly demonstrates that the HPV 16 E6 – MAGI-1 PDZ1 interaction can not simply be classified as a class I interaction. It is more complex, implicating additional residues and showing a particular binding dynamic. The detailed data generated for this specific interaction might contribute to the general understanding of PDZ binding mechanism and hopefully be of interest in the quest for treating HPV mediated cervical cancer.

MATERIALS & METHODS

Culture media

Luria Broth (LB) medium

This medium is classified as a complex medium for bacterial expression of unlabelled recombinant proteins. 1 L of LB contains 10.0 g of tryptone, 5.0 g of yeast extract and 10.0 g of NaCl at pH 7 ("*Molecular Cloning; a laboratory manual*" (1989) Sambrook J, Fritsch EF and Maniatis T.). For E6 and E6-C expression, the medium was adjusted to 100 µM $ZnSO_4$ 10 min before induction.

M9 minimal medium

This medium is classified as a minimal medium for bacterial expression and labelling of recombinant proteins. 1 L of basic medium contains: 8 g Na_2HPO_4 (2 H_2O), 3 g KH_2PO_4, 0.5 g NaCl, 0.5 g NH_4Cl, 0.85 mg $ZnSO_4$, 8.4 g $FeCl_3$ (6 H_2O), 0.10 mg $CuCl_2$, 0.1 mg $CoCl_2$ (6 H_2O), 11 mg $MnCl_2$ (2 H_2O), 18 mM EDTA, 1 mM $MgSO_4$, 0.33 mM $CaCl_2$, 4% D-glucose, 1 mg Biotin, 1 mg Thiamin (pH 7.5). For ^{15}N-labelling, $^{14}NH_4Cl$ was substituted for $^{15}NH_4Cl$ (M9^{15}N), for ^{13}C-labelling, ^{12}C-glucose was substituted for ^{13}C-glucose (M9^{13}C or M9^{13}C ^{15}N, respectively). For E6 or E6-C expressions, EDTA was not added to the minimal medium, to avoid chelating the zinc ions.

According to the vector systems and the bacterial strains used, we added different antibiotics. The standard concentrations used were: 15 µg/ml kanamycin, 35 µg/ml chloramphenicol.

Bacterial protein expression

Expression vectors

M-series expression vectors are pET-derived vectors, which have been developed by Gunter Stier (EMBL, Heidelberg, Germany). All these vectors contain a kanamycin resitance gene, a lacI gene as well as a 6 histidine tagged (6His-tag) carrier protein expressed under control of the T7-promoter. The T7 promoter is under control of the lactose or IPTG inducible lac operator. M-series vectors require a bacterial host containing a chromosomal copy of the T7 RNA polymerase. The cDNA of a protein of interest can be cloned at the C-terminus of a carrier protein via a standard MCS. A peptidic linker sequence between the carrier protein and the protein of interest provides a TEV protease cleavage site, which allows the separation of the carrier from the protein by proteolytic cleavage with recombinant TEV protease. The vectors pETM-41, pETM-30, pETM-11, code for the following carrier proteins: *E. coli* Maltose Binding Protein, *S. japonicum* Glutathione S Transferase and a 6His-tag alone, respectively
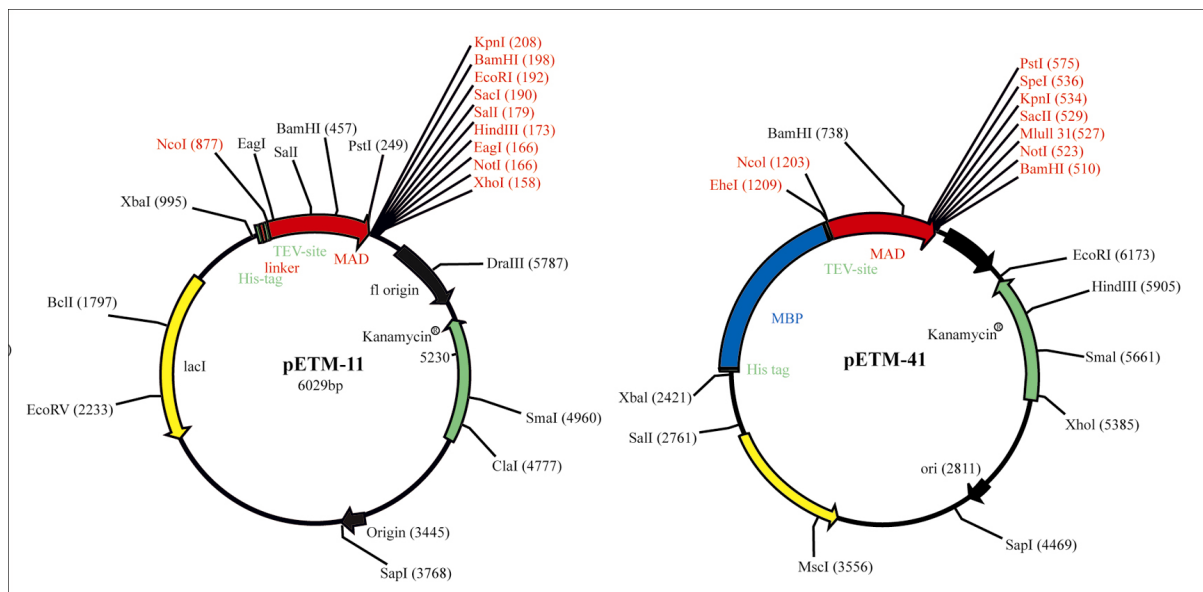
E. coli strains

I worked with three *E.coli* strains: DH5α for plasmid amplification and preparation, as well as BL21 (DE3) and BL21 (DE3) pLysS for protein expression. DH5α allow efficient transformation of unmethylated DNA from PCR reactions by a knockout of the hsdR subunit of typeI endonucleases (hsdR), have a reduced occurrence of unwanted recombination due to a knockout of the recA-protein (recA1), allow clean preparations of DNA and are deficient in Endonuclease I (endA1).

BL21 (DE3) *E. coli* are ideal when using T7 promoter-based expression vectors. They carry the lambda DE3 lysogen encoding for IPTG inducible T7 RNA polymerase, which is under the control of the lac operator. Proper recombinant protein expression is possible due to the lack of the *omp* and *lon* protease. These three factors, inducibility, high expression yields due to the strong viral promoter and the lack of endogeneous proteases make this system very powerful for protein production at amounts required for structural investigations. BL21 (DE3) pLysS contains the episomal pLysS plasmid which constitutively expresses low levels of T7 lysozyme, reducing basal expression of recombinant genes by inhibiting basal levels of T7 RNA polymerase expression. This system is suitable for expressing toxic proteins or proteins which tend to aggregate.

Electroporation

50 µl of electro-competent bacteria were transformed with 100-150 ng of plasmid DNA by electroporation with a Gene Pulser™ (Biorad) (200 Ω, 2.5 V, 25 µFa). Cells were incubated in 1 ml of LB for 30 min at 37 °C.

Preculture conditions

Usually after electroporation, BL21 were directly added at a 1:100 dilution in the preculture medium. For M9 minimal medium precultures, bacteria were harvested by centrifugation and resuspended in 1 ml of M9 before performing the 1:100 dilution in the preculture medium. In case of proteins that reveal difficult to be expressed, bacteria were plated out on LB-Agar plates and

incubated at 37 °C over night. One or several of the biggest colonies were picked and added to the preculture medium. Antibiotic concentrations were 15 µg/ml kanamycin for the pETM-vectors and 35 µg/ml chloramphenicol for the pLysS episome. Precultures were incubated over night at 37 °C and 250 rpm.

GST-peptide expression protocol

100 ng of pETM-30, coding for the required GST-peptide fusion were electroporated into BL21 (DE3) *E. coli* cells. The volume of the over night preculture was chosen according to the volume of the expression culture. For Biacore or holdup assay purposes, 25 ml of fresh LB and for peptide purification purposes, 500-1000 ml of fresh LB medium, adjusted to 15 µg/ml of kanamycin, were inoculated with 1:40 (v/v) of preculture. Expression cultures were grown at 37 °C until an OD_{600} of 0.6 was reached, induced with a final concentration of 0.5 mM IPTG and further grown at 37 °C for 2 h. Expression cultures were harvested by centrifugation, the pellets shockfrozen in liquid nitrogen and stored at −20 °C.

Recombinant TEV protease characteristics and expression protocol

The TEV protease originates from the tobacco etch virus (TEV) and is characterised by its high specificity, its high degree of purity, its broad temperature range and its efficient cleavage at low temperatures (Parks et al., 1994). It is a site specific protease that has a seven amino acid recognition site, E-N-L-Y-F-Q*G. The residues E, Y, Q, G are required for efficient cleavage. Cleavage occurs between the residues Q and G. A pET vector expressing a 6his-tagged mutated version of TEV protease (rTEV) with improved expression and solubility was provided by Gunter Stier (EMBL, Heidelberg, Germany). The 6his-tag allows easy one step removal of both rTEV and his-tagged carrier protein after proteolytic cleavage by chelating chromatography. The rTEV and the carrier protein would thus be retained on the NiNTA resin. Pure passenger protein could be easily and rapidly obtained in the flowthrough. In addition experiments showed that the presence of complete antiprotease cocktail (Roche) at 1/4th of the manufacturers recommended concentration did not affect rTEV activity. Above this limit rTEV activity can be slightly diminished but never completely suppressed. To produce the recombinant TEV protease, the rTEV-encoding plasmid was transformed in BL21 DE3 pLysS *E. coli* cells. Preculture was diluted at a 1/40 ratio into LB medium, 15 µg/ml kanamycin, 35 µg/ml chloramphenicol. Expression was performed during 9 h at 25 °C to improve folding of rTEV. 1.5 L aliquots of expression culture were harvested by centrifugation, shockfrozen in liquid nitrogen and stored at −20 °C.

E6-C expression protocol

100 ng of pETM-41, coding for the MBP-E6-C fusion were electroporated into BL21 (DE3) *E. coli* cells and 200 ml precultures grown. For expression of unlabelled protein we used LB medium for ^{15}N-labelling, M9^{15}N medium, both adjusted to 15 µg/ml of kanamycine. The preculture was diluted 1:40 in fresh medium and the culture incubated at 37 °C and 250 rpm until an OD_{600} of 0.6. LB

cultures were adjusted to 100 µM ZnSO$_4$. Cultures were transferred to 22 °C, induced by adding IPTG at a final concentration of 500 µM and incubated over night (12-14h). Cells were harvested by centrifugation, the pellets shockfrozen in liquid nitrogen and stored -20 °C.

E6 expression protocol

100 ng of pETM-41, coding for the MBP-E6 fusion were electroporated into BL21 (DE3) *E. coli* cells. E6-151 expression was performed as described for E6-C. LB expression cultures were adjusted to 100 µM of ZnSO$_4$ 10 min before induction. For ^{15}N-labelled cultures a mixture of M9^{15}N containing 5% ^{15}N-labelled *E. coli* OD2 N synthetic medium (Silantes) was used. Expression temperature was set to 27 °C. The expression time was 3-4 h for LB cultures and 4-8 h for M9 cultures.

MAGI-1 PDZ1 expression protocol

100 ng of pETM-41, coding for the MBP-MAGI-1 PDZ1 or GST-MAGI-1 PDZ1 fusion were electroporated into BL21 (DE3) *E. coli* cells and precultures were grown according to the volume of the expression culture. For expression of unlabelled protein we used LB medium for ^{15}N-labelling, M9^{15}N medium and for ^{13}C-^{15}N- labelling M9^{13}C^{15}N medium, all adjusted to 15 µg/ml of kanamycine. For protein quality monitoring 25 ml of fresh LB and for large scale protein purification 200-500 ml of fresh LB or M9 (^{15}N or ^{13}C^{15}N) medium were inoculated with 1:40 (v/v) of preculture. Expression cultures were grown at 37 °C until an OD$_{600}$ of 0.6 was reached, induced with a final concentration of 0.5 mM IPTG and further grown at 27 °C for 4 h in case of LB or 6-8 h in case of M9 cultures. Expression cultures were harvested by centrifugation, the pellets shockfrozen in liquid nitrogen and stored at –20°C.

Protein purification

Note: I have decided to list each protocol independently. This may sometimes be redundant and I could certainly have grouped the different working steps, but I prefer writing these standard operating protocols to leave ready-to-use protocols for any person who will continue working on that project.

All following protocols were performed at low temperature, i.e. on ice, at 4 °C and for gel filtration at 11-13 °C. For E6-C and E6 preparations all buffers were degassed and bubbled extensively with argon to avoid oxidation of the cysteine residues.

rTEV purification protocol

A pellet, corresponding to 1.5 L of expression culture, was resuspended in 100 ml of sonication buffer (50 mM Tris-HCl (pH 6.8), 200 mM NaCl, 2 mM DTT, 2.5 µg/ml DNaseI, 2.5 µg/ml RNase A, Complete EDTA-free antiprotease cocktail (Roche), 5% glycerol) containing 10 µM imidazole and sonicated in aliquots of 25 ml with the 20 mm probe of a S-450D Digital Sonifier (Branson). 4 runs of 1.5 min with a pulse frequency of 1 second pulse for 0.5 second pause and an amplitude of

25% were performed. The lysed bacteria were ultracentrifuged at 10500 g (12000 rpm in a Ti70 rotor) and filtered with 0.22 μm Millex®-GS filters (Millipore). A 2 ml NiNTA Superflow column (Qiagen) was equilibrated in buffer N1 (50 mM Tris-HCl (pH 8.0), 200 mM NaCl, 1 mM DTT). Protein solution was applied and the column washed with 20 vol buffer N2 (N1 + 10 mM Imidazole), 20 vol buffer N3 (N2 adjusted to 1M NaCl) and 20 vol buffer N4 (N1 + 30 mM Imidazole). Protein was eluted with buffer N5 (N1 + 300 mM Imidazole). Eluted fractions were pooled and the buffer exchanged into buffer N1 via a PD 10 Desalting Column (Amersham Biosciences). The rTEV preparation was adjusted to 50% glycerol, aliquoted into 100 μl aliquots, shockfrozen in liquid nitrogen and stored at -80 ˚C.

GST-peptide purification

Bacterial pellets corresponding to a 500 ml expression culture of MBP-E6-C were resuspended in 50 ml of sonication buffer. Aliquots of 25 ml were sonicated with the 20 mm probe of a S-450D Digital Sonifier (Branson). 6 cycles of 1 min were performed with a pulse frequency of 1 second pulse for 0.5 second pause and an amplitude of 25%. The lysed bacteria were ultracentrifuged at 10500 g (12000 rpm when using a Ti70 rotor) and filtered with 0.22 μm Millex ®-GS filters (Millipore). The filtered supernatant was loaded 3 times on a 1.5 ml NiNTA agarose resin (QIAGEN) pre-equilibrated with buffer N1. The column was washed with 10-15 vol of buffer C1 and 5 vol buffer N4. The GST-peptide was eluted with 10 ml buffer N5. The eluted protein was concentrated to 2.5 ml and buffer exchanged with buffer C1 (50 mM Tris-HCl (pH 6.8), 200 mM NaCl, 1 mM DTT) via a PD 10 Desalting Column (Amersham Bioscience). The GST-peptide concentration was measured by light absorption spectroscopy at 280 nm. The GST-peptide was further concentrated to a final concentration of about 1 mM (500-1000 μl in general). The GST-tag was separated from the peptide by adding 50 μl of rTEV as well as 1:10 of the recommended concentration of Complete antiprotease cocktail and performing an over night digestion at 16 °C. GST was finally removed by NiNTA amylose resin. 400 μl of NiNTA agarose were equilibrated in buffer N1. 100 μl aliquots of NiNTA resin beads were aliquoted into Eppendorf tubes, the superfluous buffer discarded and the 500-1000 μl of digestion mixture incubated for 5-10 min on a rotating wheel. The liquid phase was extracted by centrifugating the resin liquid mixture in a 0.5 ml Eppendorf tube possessing a pin-hole, filled with glass wool and placed on top of 2 ml tubes. This procedure was repeated 3 more times in order to remove all GST protein without increasing remarkably the reaction volume. Depletion of GST peptide was verified by SDS PAGE. Peptide was aliquoted into 100 μl aliquots and stored at −20 °C. (NOTE: As the peptide cannot further be concentrated once it is separated from the GST-tag and as it is often very difficult to measure its concentration due to the lack of tyrosine or tryptophane residues, it is important to stay in the same volume. This allows making the assumption that the peptide concentration should be in the range of what was measured by light absorption in presence of the GST-tag.)

E6-C purification protocol

Bacterial pellets corresponding to a 1 L expression culture of MBP-E6-C were resuspended in 100-150 ml sonication buffer and sonicated in aliquots of 25 ml with the 20 mm probe of a S-450D Digital Sonifier (Branson). 5 cycles of 1 minute were performed with a pulse frequency of 1 second pulse for 0.5 second pause and an amplitude of 25%. The lysed bacteria were ultracentrifuged at 10500 g (12000rpm when using a Ti70 rotor) and filtered with 0.22 µm Millex®-GS filters (Millipore). The filtered supernatant was loaded onto an 80 ml amylose column (New England Biolabs) pre-equilibrated with buffer C1. The column was washed with 10-15 vol buffer C2 (C1 + 1:10 of the recommended concentration of protease inhibitor cocktail). The flowthrough of this washing step was collected in 50 ml fractions. 50-80% of the recombinant MBP-E6-C leaks from the amylose column during late washing steps, after most of all contaminating proteins, and can thus be collected in a highly pure form in the washing fractions. Remaining protein was eluted with buffer C1 supplemented with 10 mM maltose. Purity level and amounts of MBP-E6-C were estimated by Tris-Tricine PAGE. The fractions containing MBP-E6-C were pooled and ultracentrifuged at 160000 g (36000 rpm when using a SW-41 rotor) and 4 °C during 14-16 h. The supernatant was concentrated to 10 ml using Vivaspin 15 filter devices with a MWCO of 5 kDa (VIVASCIENCE) (VS 5/15). (Note: these filters possess a polyethersulfone membrane which has been tested as the most appropriate for E6 concentration, minimising losses by unspecific protein binding to the membrane (Never use membranes made of regenerated cellulose for E6 and its domains!)). The MBP moiety was removed by adding 500 µl of rTEV and performing an over night digestion at 16 °C. The digestion mix was concentrated to 2 ml and separation of the MBP fragment, the rTEV and E6-C was achieved by gel filtration on a HiLoad 16/60 Superdex 75 pg ($3x10^3$ x $7x10^4$ Da) gel filtration column (Amersham Biosciences) previously equilibrated with 240 ml buffer C1. Eluted E6-C was tested by Tris-Tricine PAGE (Schagger and von Jagow, 1987). Fractions were pooled, the concentration estimated by absorption spectroscopy at 280 nm and stored at 4 °C.

E6 purification protocol

All buffers were adjusted to 400 mM NaCl. Bacterial pellets corresponding to 1 L of MBP-E6 expression culture in LB were resuspended in 100-150 ml sonication buffer and sonicated in aliquots of 25 ml with the 20 mm probe of a S-450D Digital Sonifier (Branson). 5-10 cycles of 1 minute were performed with a pulse frequency of 1 second pulse for 0.5 second pause and an amplitude of 25%. Lysed bacteria were loaded on an amylose column as described for E6-C, with the difference that MBP-E6-151 does not leak off the column during the washing steps. MBP-E6-151 was eluted with buffer C2 containing 10 mM maltose. The eluate was tested via Tris-Tricine PAGE and further processed like described for E6-C. The added rTEV volumes for digestion were adjusted according to the volumes described in the E6-C protocol. Gel filtration was performed on a HiLoad 16/60 Superdex 75 pg ($3x10^3$ x $7x10^4$ Da) gel filtration column (Amersham Biosciences). Eluted E6 was tested by Tris-Tricine PAGE. Fractions were pooled, the concentration estimated by absorption spectroscopy at 280 nm and stored at 4 °C.

MAGI1-PDZ1 purification protocol

Bacterial pellets corresponding to a 200-500 ml of expression culture of MBP-MAGI-1 PDZ1 were resuspended in 50-100 ml sonication buffer and sonicated in aliquots of 25 ml with the 20 mm probe of a S-450D Digital Sonifier (Branson). 3-5 cycles of 1 minute were performed with a pulse frequency of 1 second pulse for 0.5 second pause and an amplitude of 25%. The lysed bacteria were ultracentrifuged at 10500 g (12000 rpm when using a Ti70 rotor) and filtered with 0.22 µm Millex®-GS filters (Millipore). Cleared bacterial extracts were loaded on an amylose column (New England Biolabs), pre-equilibrated with buffer C1. The protein was eluted with buffer C1 supplemented with 10 mM maltose. The MBP tag was removed by proteolytic cleavage with rTEV at 8-10 °C over night. The His6-MBP tag and the 6His-rTEV were removed by loading the digestion mix 3 times on a 2 ml NiNTA resin (QIAGEN), pre-equilibrated with buffer C1. The flowthrough, containing essentially the PDZ1-domain was collected, concentrated to 2 ml and subjected to a final gel-filtration on a Hiload 16/60 Superdex 75 pg ($3x10^3$ x $7x10^4$ Da) gel filtration column (Amersham Biosciences) equilibrated in buffer C1, resulting in pure and monodisperse PDZ1.

Protein Quality Monitoring

Expression tests

Expression tests were systematically performed after each expression culture. 100 µl of expression culture were centrifuged at 1500 g (4000 rpm in a benchtop centrifuge) for 5 minutes. The supernatant was discarded and the bacterial pellet resuspended in 50 µl SDS PAGE sample buffer (120 mM Tris-HCl (pH 6.8), 4% SDS, 20% glycerol, 2.5 mg/ml bromophenol blue, 10% ß-mercaptoethanol) and 50 µl H_2O. 20 µl samples were analysed by SDS PAGE.

Small-scale purification of fusion-proteins

A standard protocol was developed for routine small-scale production of fusion proteins. It was modified at different steps depending on the parameters tested. Samples of 10-25 ml expression culture were pelleted at 1500 g at 4 °C, resuspended in 1 ml sonication buffer in a 2 ml Eppendorf tube and sonicated with the 3 mm probe of a Vibracell 72412 sonicator (Bioblock Scientific) performing 2 runs of 15 seconds with a pulse frequency of 1 second pulse for 0.5 second pause and an amplitude of 6%. (These sonicated extracts can then be used for direct pellet supernatant test.)

Lysed bacteria were centrifuged at 16000 g (12000 rpm in a benchtop centrifuge) for 15 min at 4 °C. Supernatants were incubated for 30 min with either 200 µl of amylose resin (New England Biolabs), 150 µl of glutathione sepharose resin (Amersham Biosciences) or 60 µl of NiNTA Superflow resin (Qiagen) pre-equilibrated with buffer C1. The resin was washed 5 times for 2 min with 1.5 ml of buffer C1 containing 1:10 of the manufacturers recommended concentration of complete EDTA free antiprotease cocktail (Roche). The washed resin beads were finally incubated for 5 min with 1 volume (bed volume of resin) of buffer C1 containing either 20 mM maltose (for amylose resin), 20 mM reduced glutathione (for glutathione resin) or 600 mM imidazole (for NiNTA resin) to elute the fusions

from the resin. Extraction of total liquid was achieved by transferring the resin-liquid mixture to a 96 well MultiScreen-HV filterplate possessing a 0.45 µm Durapore membrane (Millipore) and centrifugation. Alternatively, extraction can be achieved by centrifugating the liquid/resin mixtures through 0.5 ml Eppendorf tubes possessing a pin-hole, filled with glass wool and placed on top of 2 ml tubes. (The purified fusions could be subjected to static light scattering measurements or to pellet/supernatant tests after proteolytic removal of the solubilising tag)

Pellet/supernatant test on lysed bacteria

Samples of 50 µl of a lysed bacterial expression culture were centrifuged at 16000 g (12000 rpm in a benchtop centrifuge) for 15 minutes. The supernatant was transferred to a new tube and mixed with 50 µl SDS PAGE sample buffer. The pellet was resuspended in 50 µl H_2O and 50 µl SDS PAGE sample buffer. 20 µl samples of pellet and supernatant were tested by SDS-PAGE. The ratio of soluble protein compared to aggregated protein can be determined with Quantity One™ quantification software (BIORAD) or, for rough estimation, simply by eye.

NOTE: This test is adapted to reveal the proportion of soluble protein and the amount of solid aggregates (insoluble inclusion bodies) out of a protein expression. It is assumed that soluble protein is well folded and monodisperse. This method is not adapted, when the protein is fused to a solubilising tag. Especially MBP has been reported to form soluble aggregates. Therefore folded and misfolded MBP-fusion protein stays always in the supernatant and no conclusion about the amount of folded protein is possible. In this case a pellet/supernatant test after proteolytic cleavage of the solubilising tag or an aggregation ratio measurement by static light scattering is required.

Pellet/supernatant test on purified protein samples

50 µl of purified fusion protein were incubated with 5 µl of rTEV for 3 h at room temperature in an Eppendorf tube. The digestion mix was centrifuged at 16000 g (12000 rpm in a benchtop centrifuge) and 4 °C for 15 minutes. The supernatant was transferred to a new tube and one volume of SDS PAGE sample buffer added. 50 µl of the working buffer and 50 µl of SDS PAGE sample buffer were added to the original tube containing the pellet and the tube vortexed vigorously in order to resuspend the pellet. 20 µl samples of pellet and supernatant were tested by SDS-PAGE. The ratio of soluble protein compared to aggregated protein can be determined with Quantity One™ quantification software (BIORAD) or simply by eye.
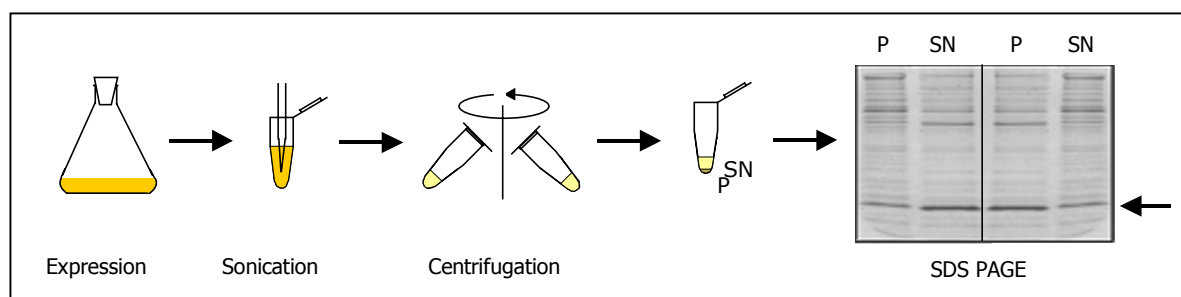
Fluorescence, static light scattering and aggregation ratio determination

Experiments were performed with a SPEX Fluorolog-2 spectrofluorimeter (SPEX Industries, Inc., Edison, NJ) equipped with a 450-W Xe lamp, a double-grating excitation monochromator and a single-grating emission monochromator. Data were acquired with a photon counting photomultiplier (linear up to 10^7 counts/s) where high voltages were fixed at 800 V. Slitwidths were adjusted in order to avoid saturation of the detectors. 2 ml of a diluted sample of purified fusion protein was placed in a quartz cuvette maintained at 20 °C and excited by monochromatic light at 350 nm. The scattered light was recorded at 350 nm. Measurement of the buffer alone served as substraction baseline. Fluorescence was measured by exciting the sample at 280 nm and recording the emission spectrum in the range between 300 and 400 nm. The maximum fluorescence intensity was recorded at about 340 nm.
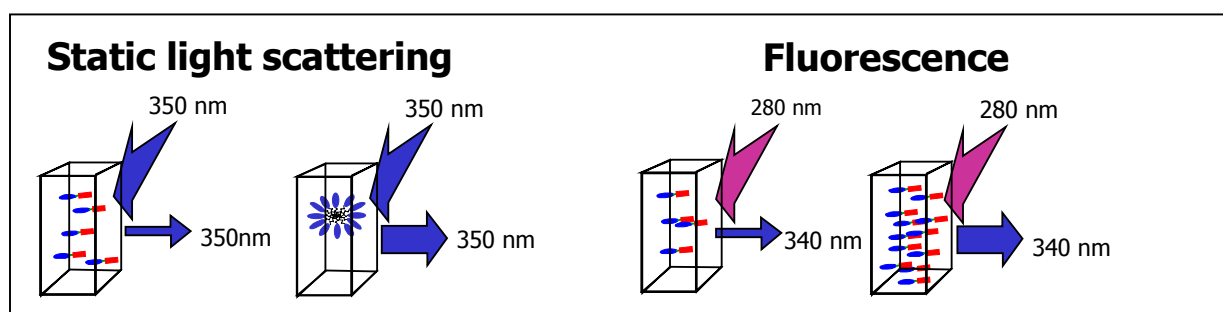
From these two values the aggregation ratio (R_{agg}) was calculated as follows:

$$R_{agg} = I350 / I340 \, ,$$

where $I350$ and $I340$ represent the intensities of maximal scattered light and maximal fluorescence, respectively. $I350$ is proportional to both protein concentration and the averaged mass of the scattering particles whereas $I340$ is proportional to protein concentration only. Therefore R_{agg} depends on the averaged molecular weight of the scattering particles only. It is important to notice that R_{agg} values will vary depending on the instrument used and on the instrumental parameters chosen for the measurement. These parameters have to be initially calibrated using one sample of aggregated fusion protein and one sample of monodisperse fusion protein at the same concentrations.

Quick assay of protein folding by ^{1}H-^{15}N HSQC (NMR)

A soluble and monodisperse protein suspension does not necessarily implicate a stable and unique fold. The protein could be monomeric but still be trapped in a folding intermediate stage or explore a given conformational space. To assay this it is possible to perform rapid and small scale production of ^{15}N-labelled protein followed by ^{1}H-^{15}N HSQC experiments. ^{1}H-^{15}N HSQC is a standard NMR experiment which generates in some way a fingerprint of the protein, giving information about the fold stability of the protein. In general a 200-300 ml expression culture in M9^{15}N minimal medium yields sufficient recombinant protein for this kind of NMR experiment. Since all pETM vectors bear a 6His-tag as well as the rTEV protease a double NiNTA purification approach can be adopted to purify

the protein. Briefly, the bacteria are lysed by sonication and the fusion protein is rapidly purified on a first NiNTA column (in general not more than 1 ml of resin bed volume). The protein is buffer exchanged with working buffer by a PD10 column, to remove the imidazole, and the fusion protein is digested with rTEV. The 6His-tagged carrier protein and the rTEV are removed by incubating the digestion mix several times on a second NiNTA column, yielding pure protein in the flow trough. This protein can be concentrated to 50-100 µM and analysed by ^{1}H-^{15}N HSQC, which gives directly information about the fold stability of the protein and if it is suitable for further NMR investigation or structure calculation. (for more detailed information about NMR, please refer to section "NMR analysis" in the materials and methods).

Protein interaction assays

For all protein-protein interaction assays I will define the protein which is fixed on a solid phase as the "ligand" protein and the interaction partner in the liquid phase as the "analyte" protein, referring to the nomenclature commonly used when working with the Biacore technology.

Protein-protein interaction assay using comparative chromatographic retention

NOTE: In this material and methods section I will only present the protocol for the basic holdup or catchup assay. This assay can also be adapted for assaying protein-protein interaction using eucaryotic cell extracts as well as for the estimation of the KD value for a given interaction. All these different modifications concerning this basic protocol are listed in publication 3 in the results section.

Here again I chose deliberately to present the two assays separately in order to provide a ready to use standard operation protocol.

All recombinant proteins used to perform these assays resulted of 25 ml expression cultures, which were harvested by centrifugation and lysed by sonication and cleared by centrifugation as described in "small scale purification of proteins" in the protein quality monitoring section.

The holdup assay

Affinity resins were pre-equilibrated in buffer C1 (50 mM Tris-HCl (pH 6.8), 200 mM NaCl, 1 mM DTT). 1 ml cleared bacterial ligand protein expression lysates were batch-incubated at 4 °C for 30 min in the presence of 100-300 µl Glutathione Sepharose (Amersham Pharmacia Biotech) for GST-ligands, or 100-300 µl Maltoheptaose Agarose (Sigma) for MBP-ligands. Beads were then washed once for 10 min in 15 ml of buffer C1 supplemented with Complete anti-protease cocktail (Roche), resulting in resin beads saturated with a concentration of 50-100 µM of pure ligand. For each interaction experiment, we prepared two 500 µl Eppendorf tubes containing 10 µl (bed volume) of ligand-saturated resin each. For pure-pure holdup assays, both tubes were supplemented with 30 µl of buffer C1 containing 5-10 µM of pure analyte protein. For pure-crude holdup assays, both tubes were

supplemented with 2-4 µl of cleared bacterial analyte expression lysates adjusted with buffer C1 to 30 µl. Ligand-analyte interaction was achieved by a 15 min incubation on ice interspaced every minute with a quick step of resin resuspension by gentle vortexing, avoiding to spread resin on the tube walls. For each pair of samples, one tube (labelled +G, or +M) was adjusted to 10 mM of competitor molecule (reduced glutathione for glutathione resin, or maltose for maltoheptaose resin, respectively). This was achieved by adding 4 µl of a 110 mM stock solution of competitor (note that the pH of glutathione stock solutions must be adjusted to 7). The other tube (labelled -G, or -M) was adjusted with 4 µl of water. After another 5 min incubation at 4 °C interspaced with gentle vortexing, the total liquid present outside and inside the resin was extracted by transferring the liquid/resin mixtures to a 96 well MultiScreen-HV filterplate possessing a 0.45 µm Durapore membrane (Millipore) and centrifugation. Alternatively, extraction can be achieved by centrifugating the liquid/resin mixtures through 0.5 ml Eppendorf tubes possessing a pin-hole, filled with glass wool and placed on top of 2 ml tubes.

Extracted liquid phases of experiments involving purified proteins or extracts of bacterially overexpressed proteins were adjusted with one volume of 2x concentrated SDS-PAGE sample buffer and run on 15% SDS gels which were stained with Coomassie-blue.

The catchup assay

Affinity resins were pre-equilibrated in buffer C1 (50 mM Tris-HCl (pH 6.8), 200 mM NaCl, 1 mM DTT). Ligand-analyte interaction was performed in solution before incubation with resin. For pure-pure catchup assays we diluted pure ligand and pure analyte into 70 µl of buffer C1, at a final concentrations of 10-30 µM for the ligand and 2-5 µM for the analyte. For crude-crude catchup assays, using two cleared bacterial overexpression lysates, we mixed 17 µl of ligand extract with 8 µl of analyte extract into a total volume of 70 µl (buffer C1). Note that these proportions were chosen after a preliminary SDS PAGE analysis to ensure that the ligand would be in slight excess as compared to the analyte in the reaction. Ligand-analyte interaction was performed by incubating the mixtures on ice for at least 15 min. For each interaction experiment, we then prepared two 500 µl Eppendorf tubes each containing 10 µl of appropriate affinity resin beads. Both tubes were supplemented with 30 µl of ligand-analyte mixture. Capture of ligand fusion protein, free or bound to analyte, was finally achieved via 15 min incubation on ice interspaced with a gentle vortexing step every minute. Addition of competitor molecule, liquid phase extraction and SDS-PAGE analysis were performed as described for the holdup assay.

Protein-peptide interaction assay using Biacore technology

Biacore technology allows quantifying molecular interaction in real time by a physical process which is called surface plasmon resonance (SPR). A monochromatic light beam is focused on a gold plated glass surface. The other side of the glass surface, also called sensorchip, is coated with a thin dextran layer which is immerged in a system of microchannels with a constant flow of working buffer.

On this dextran layer a first molecule or protein (the ligand) can be fixed and later put into contact with a pontentially binding molecule (the analyte) via injection into the buffer stream. The angle of the light beam is chosen that it is totally reflected. This condition allows the phenomenon of SPR to take place. Incident photons get into resonance with free electrons of the gold layer and induce a fine shadow in the reflected light beam. The angle of the shadow varies as a function of the density on the dextran chip surface. Measurement of the angle variations of this shadow allows thus following in real time association and dissociation of molecules. The resonance signal is quantified in relative units (RU). A variation of 1000 RU corresponds to an angular deviation of 0.1° of the shadow and binding
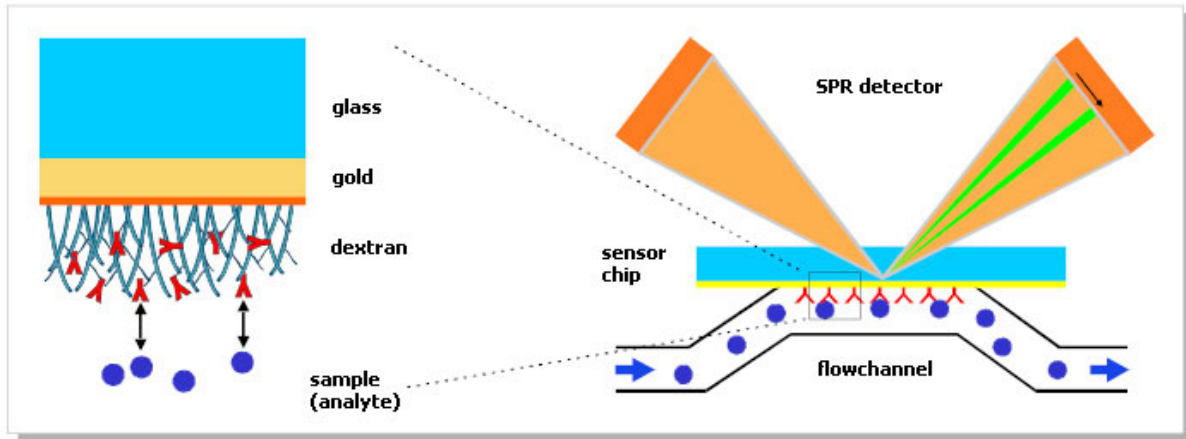
of 1 ng of protein per mm^2 on the dextran surface.

In a classical Biacore experiment, first, the ligand molecule is fixed either covalently or by means of antibodies on the dextran surface. The dextran surface in facts allows increasing accessible surface and "mimics a soluble interface". Second, the analyte molecule is injected in the buffer flow in the flowchannel which is supposed to interact with the ligand. If the interaction takes place, modification of the signal occurs. A classical sensorgram allows monitoring in real time the association phase, the steady state level, the dissociation phase and the regeneration of the surface. If regeneration does not harm the surface it can be re-used several times.
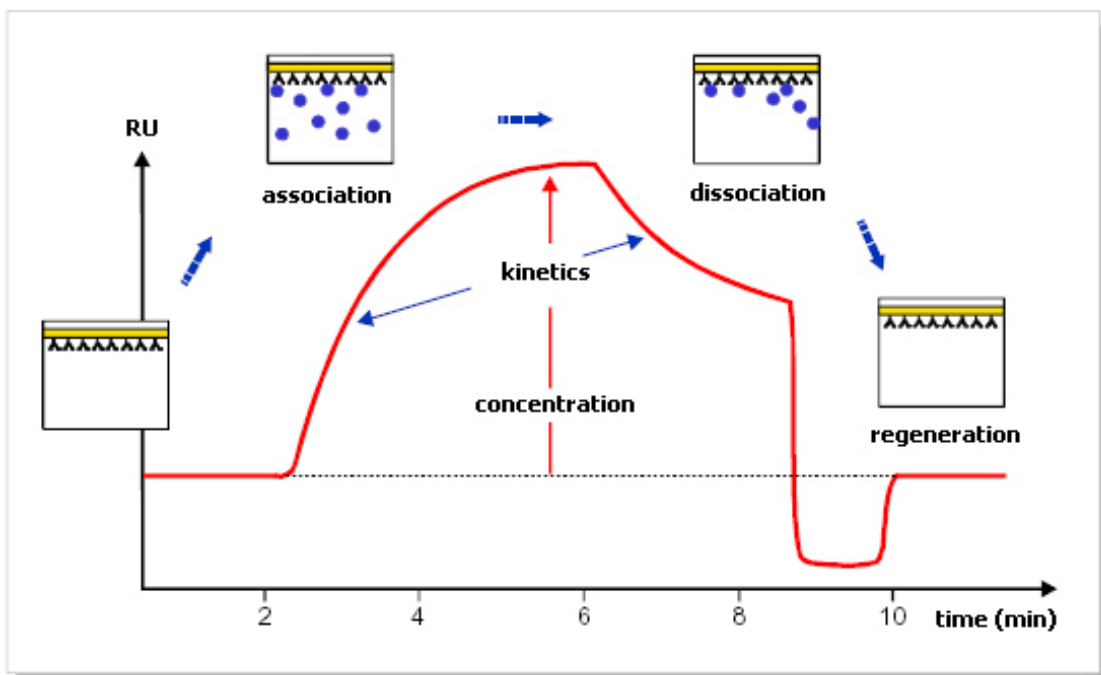
All experiments were performed on a Biacore 2000 apparatus at 25 °C with the autosampler rack base cooled to 10 °C.

Buffer preparations

-Running buffer 1 for anti-GST antibody immobilisation:
10 mM HEPES, 150 mM NaCl, 0.005% P-20

-GST-peptide lysis buffer:
50 mM Tris-HCl (pH 6.8), 400 mM NaCl, 2 mM DTT, 1 µg/ml DNase I, 1 µg/ml RNase A, 1% NP40, 5% glycerol, 1:1 of the manufacturers recommended concentration of Complete antiprotease cocktail (Roche)

-GST-peptide dilution buffer:
50 mM Tris-HCl (pH6.8), 200 mM NaCl, 1 mM DTT, 1% NP40, 1:2 of the manufacturers recommended concentration of Complete antiprotease cocktail (Roche)

-Running buffer 2 for interaction assays:
50 mM Tris-HCL (pH6.8), 200 mM NaCl, 0.5 mM DTT, 0.005% P-20

-Regeneration buffer:
50 mM glycine (pH 2.2)

GST-peptide lysis protocol
Bacterial pellets corresponding to 20 ml of GST-peptide expression culture were resuspended with 1 ml of peptide lysis buffer in a 2 ml Ependorf tube and sonicated with the 3 mm probe of a Vibracell 72412 sonicator (Bioblock Scientific). 6 cycles of 1 min were performed with a pulse frequency of 1 second pulse for 0.5 second pause and an amplitude of 6%. The lysed bacteria were ultracentrifuged at 10500 g (12000 rpm when using a Ti70 rotor) and 4 °C and filtered with 0.22 µm Millex ®-GS filters (Millipore). 200 µl aliquots of filtered supernatant were shockfrozen in liquid nitrogen and stored at –80 °C.

Anti-GST antibody immobilisation
A new CM5 sensor surface was washed with two 30 s pulses of 50 mM NaOH, 500 mM NaCl and one 30 s pulse of regeneration solution and then left for 1 h at a flow-rate of 20 µl/min in running buffer 1. Each flow-cell of the sensor chip surface was activated independently for 10 min with a solution containing N-ethyl-N′-[3′-(diethylamino)propyl]carbodiimide (EDC) and N-hydroxysuccinimide (NHS) at a flow-rate of 10 µl/min. Anti-GST-antibody was fixed by manual injection of a 8 µg/ml dilution of anti-GST polyclonal antibody in 10 mM sodium acetate

coupling solution (pH 5.5) at a flow-rate of 40 µl/min until 1500 RU was reached. The surface was deactivated by injecting ethanolamine for 10 min followed by four times 1 min pulse of regeneration solution at a flow-rate of 10 µl/min.

Ligand immobilisation

The running buffer for peptide-protein interaction assays was changed to running buffer 2. Cleared bacterial GST-peptide expression lysates were diluted 1:10 (v/v) in GST peptide dilution buffer. 60 µl of this diluted lysate were injected at flow 10 µl/min on an individual channel. Each flow-cell could thus be saturated with GST-peptide, where flow-cell 1 usually contained the GST-none peptide (negative control).

Analyte injection

For kinetic measurements of a given GST-peptide / PDZ1 interaction, the flow-rate was set to 20 µl/min. Purified PDZ1 diluted in running buffer 2 was injected at 8 different concentrations (0, 1, 2, 5, 15, 20, 30 and 60 µM) for 30 s followed by a 1 min pulse of the synthetic HPV16 E6 C-terminal peptide concentrated at 25 µM using the command "coinject". This prevented rebinding of the PDZ1 domains to the surface and allowed a rapid return to a stable baseline. At least, two independent experiments were performed for each GST-peptide / PDZ1 interaction. Sensorgrams were recorded with the Biacore software. (Note: In case of a very slow dissociation and of a baseline which does not come back to zero rapidly, one has to change the protocol by working on only one flowcell of the Biacor chip (see publication 4 Zanier et al.).

Regeneration

The flowcells were regenerated by three successive injections of 1 min of regeneration solution. This removes the GST-peptide and liberates the chip surface for another round of GST-peptide immobilisation.

Data analysis and KD calculation

Data were analysed using the BiaEvaluation 3.2 software (Biacore). Steady state analysis of the data was performed using a simple Langmuir interaction model. The kinetic parameters were obtained from fits with X^2 values smaller than 5% of the globally fitted R_{max} values. Equilibrium responses (Req) as a function of total analyte concentrations were fitted to simple 1:1 interaction binding isotherms.

NMR analysis

Nuclear magnetic resonance spectroscopy (NMR spectroscopy)

NMR spectroscopy is a biophysical technique which exploits the magnetic properties of the nuclei of atoms and takes an important place in modern natural science. This technique can be used to perform a broad range of investigations including: 1) identification of signals from individual atoms in a pure molecule; 2) identification of functional groups or how many atoms of one type exist within a sample; 3) studying mixtures of different molecules; 4) analysing dynamic effects, for example after a change in temperature or following reaction mechanisms 5) determining protein and nucleic acid structures with a view to understanding their function. The following chapters will give only a brief introduction into the basic protein NMR spectroscopy methods that I used during my Ph.D. studies.

Some NMR basics

When placed in a magnetic field, the energy levels of certain atomic nuclei become non-degenerate, such that the nuclei resonate at distinct frequencies. This resonant frequency is sensitive to the magnetic (chemical) environment of the nucleus: the nuclei of the same isotope in an identical magnetic environment will resonate at the identical frequencies. However, nuclei of the same isotype will resonate at different frequencies if their magnetic environments are not identical. Since resonant frequencies depend linearly on the strength of the external magnetic field, they may be expressed as a field-independent value, the chemical shift in parts per million (ppm), by simple division of the observed frequency (in Hertz) by that of the magnetic field. Chemical shifts are reported relative to a defined reference frequency. In protein NMR, which deals mainly with ^{1}H, ^{13}C and ^{15}N nuclei, the reference commonly used is tetramethylsilane. In protein NMR the spread of chemical shifts caused by the diverse magnetic environments in a folded protein permits the identification of signals from individual nuclei in a macromolecule.

Protein nuclear magnetic resonance spectroscopy

Protein NMR is a field of structural biology in which NMR spectroscopy is applied to the study of the structural and dynamic behaviour of a polypeptide or protein. The field was pioneered by, among others, the laboratory of Kurt Wüthrich, for which he was awarded the Nobel prize in 2002. Structure determination by NMR is organised in several following steps with each requesting separate specialised techniques. Briefly the sample is prepared, the frequencies of the nuclei measured and assigned, distances between nuclei measured, a list of spatial restraints generated and finally the structure calculated. Here, I will not discuss the fundamental aspects of NMR as applied to proteins, as I am not a structural biologist and lack the necessary theoretical background to introduce this topic decently. I would refer the reader to (Cavanagh et al., 1995). Nevertheless I would like to introduce briefly some basic experiments which can provide a lot of information about the behaviour of a protein at the atomic level to a biochemist or a molecular biologist.

1D proton NMR

The simplest NMR experiment is the recording of a 1 dimensional (1D) ^{1}H spectrum [There is no strict size limit for this experiment, although resonances become broader and overlap becomes more severe with increasing molecular weight]. It allows to measure proteins up to 100 kDa and the actual cryoprobes allow to measure at protein concentrations as low as 10 μM. Spectra can be acquired in less than 1 hour and analysed directly. The 1D ^{1}H spectrum measures the frequencies of each proton (^{1}H nucleus) in the protein and allows qualitative differentiation between a folded and an unfolded protein. Unfolded proteins show poor dispersion with most of the backbone amide proton frequencies overlapping around 7.5 to 8.5 ppm. Folded proteins provide a variable chemical environment for the different amino acids and therefore they show a greater dispersion of the proton frequencies which can exceed 10 ppm and lie below 0.5 ppm. In general, amide protons resonate around 7 to 10 ppm, side chain amides around 5 to 7 ppm, protons from α-carbons around 4-5 ppm and protons from aliphatic residues lower

than 4 ppm with folded methyl protons typically at the lowest ppm values. The figure shows the spectra of an unfolded and a folded protein, illustrating the greater peak dispersion for the folded protein compared to the unfolded protein, with especially the peaks below 0 ppm corresponding to the proton of folded methyl groups, characteristic of a folded protein.
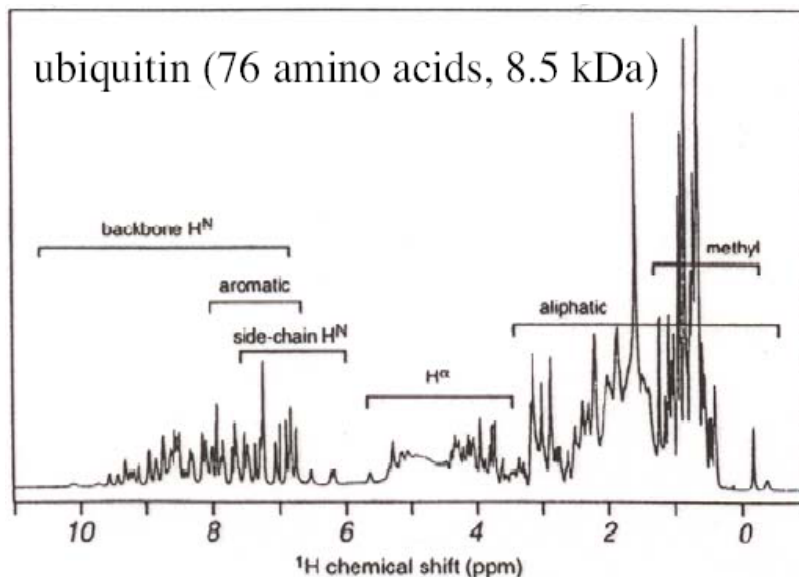
However, it is usually impossible to assign each peak to specific atoms in the polypeptide chain, due to the complexity and the high degree of peak overlap in this kind of spectra. To do so and therefore to identify each proton frequency, 2D or even 3D spectra have to be recorded. These spectra allow to visualise specifically [1]H-^{1}H scalar (through-bond) and dipolar (through-space) couplings and in the case of isotopically labelled proteins exploit couplings to ^{15}N and ^{13}C nuclei. This allows the problem of resonance overlap to be overcome and gives access to individual frequencies.
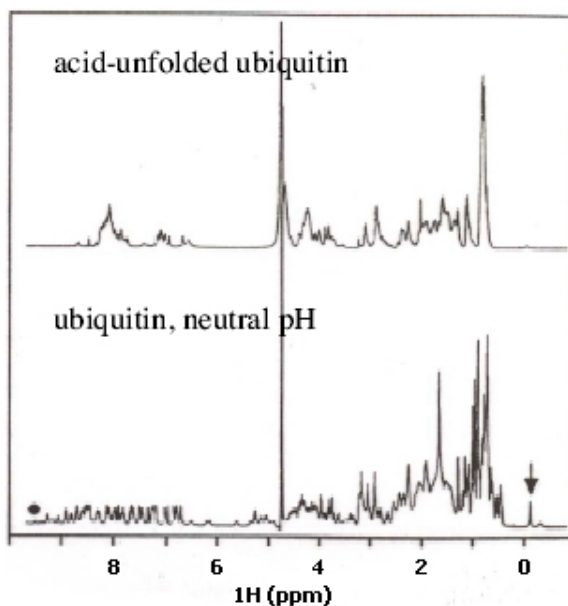
^{1}H-^{15}N HSQC NMR

The ^{1}H-^{15}N HSQC experiment is a 2D experiment which requires imperatively a ^{15}N-labelled protein. It provides correlation peaks between the ^{1}H frequency and the corresponding ^{15}N frequency of amide groups of the peptide bonds in the polypeptide main chain, as well as of side chain amides. Folded proteins show a broad dispersion of frequencies giving a characteristic peak pattern, like a fingerprint of the protein, where each correlation corresponds to one amino acid (except for the side chain amides which give two peaks and the "invisible" proline residues).
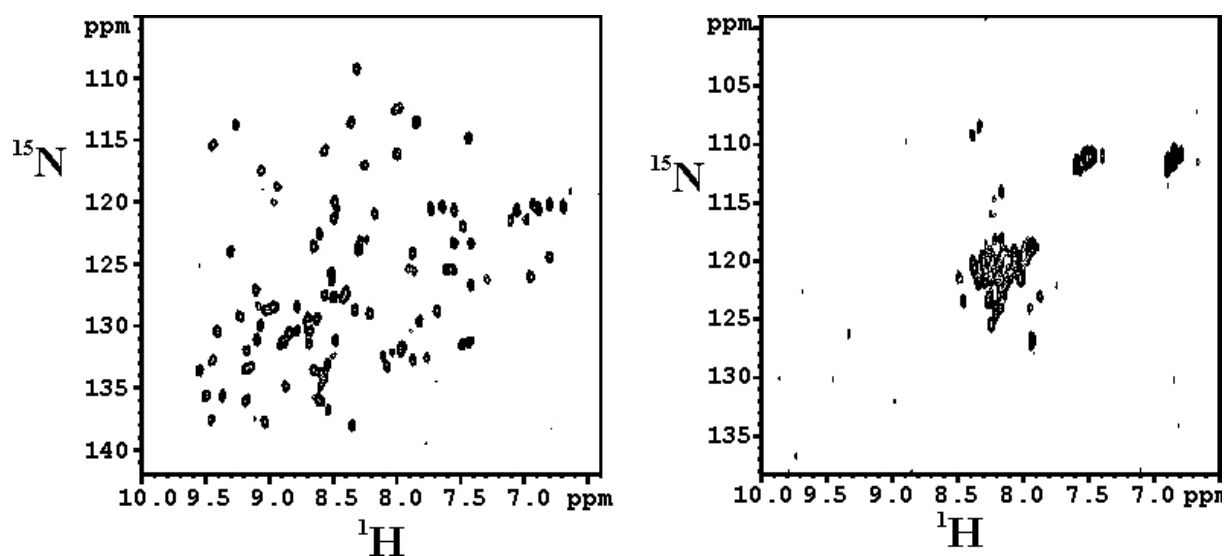
The figure shows in the left panel a ^{1}H-^{15}N HSQC spectrum of a folded protein and in the right panel a ^{1}H-^{15}N HSQC spectrum of a unfolded protein, with most of the backbone amide correlation peaks around 8 ppm and the NH_2 sidechain correlation peaks overlapping around 7.5 and 6.8 ppm. Besides providing information on protein quality, i.e. protein folding, the HSQC also gives information about the dynamic behaviour of the amino acids. The linewidth of the peaks of a homogeneous protein sample gives information about the motions of the residues. Line broadening is representative of slow time-scale motions, such as for residues implicated in a region of the protein which can adopt several distinct conformations. On the contrary, sharp peaks indicate fast time-scale motions like for example residues in a flexible loop or in an unstructured, random coil conformation. Additional 2D or 3D experiments allow each correlation peak of the HSQC spectrum to be assigned to one specific amide proton. Once this information is acquired, the HSQC experiment provides a fast and reliable tool for studying the protein at an atomic level, like for example the structural effect of mutations or for mapping protein-ligand interaction, by analysing the chemical shift differences upon modification of the protein.

NMR data acquisition

Samples for NMR experiments were prepared in 25 mM Tris-HCl (pH 6.8), 50 mM NaCl, 1 mM DTT, 10% D_2O. In order to investigate the binding between MAGI-1 PDZ1 and E6-C, ^{1}H-^{15}N heteronuclear single quantum correlation (HSQC) spectra were acquired at 600 MHz and 295K on a Bruker DRX600 spectrometer aquipped with a z-gradient triple resonance cryoprobe. The water signal

was suppressed using the WATERGATE sequence (Piotto et al., 1992). Data were processed using NMRPipe (Delaglio et al., 1995) and analysed with NEASY. was suppressed using the WATERGATE sequence (Piotto et al., 1992). Data were processed using NMRPipe (Delaglio et al., 1995) and analysed with NEASY.

REFERENCES

References

Abarzua P, LoSardo JE, Gubler ML, Spathis R, Lu YA, Felix A, Neri A. Restoration of the transcription activation function to mutant p53 in human cancer cells. *Oncogene.* (1996) Dec 5;**13**(11):2477-82.

Agorastos T, Miliaras D, Lambropoulos AF, Chrisafi S, Kotsis A, Manthos A, Bontis J. Detection and typing of human papillomavirus DNA in uterine cervices with coexistent grade I and grade III intraepithelial neoplasia: biologic progression or independent lesions? *Eur J Obstet Gynecol Reprod Biol.* (2005) Jul 1;**121**(1):99-103.

Alewine C, Olsen O, Wade JB, Welling PA. TIP-1 Has PDZ Scaffold Antagonist Activity. *Mol Biol Cell.* (2006) Oct;**17**(10):42000-11.

Alonso LG, Garcia-Alai MM, Nadra AD, Lapena AN, Almeida FL, Gualfetti P, Prat-Gay GD. High-risk (HPV16) human papillomavirus E7 oncoprotein is highly stable and extended, with conformational transitions that could explain its multiple cellular binding partners. *Biochemistry.* (2002) Aug 20;**41**(33):10510-8.

Alvarez-Salas LM, Cullinan AE, Siwkowski A, Hampel A, DiPaolo JA. Inhibition of HPV-16 E6/E7 immortalization of normal keratinocytes by hairpin ribozymes. *Proc Natl Acad Sci U S A.* (1998) Feb 3;**95**(3):1189-94.

Anderson JM. Cell signalling: MAGUK magic. *Curr Biol.* (1996) Apr 1;**6**(4):382-4.

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

Anzai N, Miyazaki H, Noshiro R, Khamdang S, Chairoungdua A, Shin HJ, Enomoto A, Sakamoto S, Hirata T, Tomita K, Kanai Y, Endou H. The multivalent PDZ domain-containing protein PDZK1 regulates transport activity of renal urate-anion exchanger URAT1 via its C terminus. *J Biol Chem.* (2004) Oct 29;**279**(44):45942-50.

Arkin MR, Wells JA. Small-molecule inhibitors of protein-protein interactions: progressing towards the dream. *Nat Rev Drug Discov.* (2004) Apr;**3**(4):301-17.

Auster AS, Joshua-Tor L. The DNA-binding domain of human papillomavirus type 18 E1. Crystal structure, dimerization, and DNA binding. *J Biol Chem.* (2004) Jan 30;**279**(5):3733-42.

Bahner M, Sander P, Paulsen R, Huber A. The visual G protein of fly photoreceptors interacts with the PDZ domain assembled INAD signaling complex via direct binding of activated Galpha(q) to phospholipase cbeta. *J Biol Chem.* (2000) Jan 28;**275**(4):2901-4.

Bai L, Wei L, Wang J, Li X, He P. Extended effects of human papillomavirus 16 E6-specific short hairpin RNA on cervical carcinoma cells. *Int J Gynecol Cancer.* (2006) Mar-Apr;**16**(2):718-29.

Baker CC, Phelps WC, Lindgren V, Braun MJ, Gonda MA, Howley PM. Structural and transcriptional analysis of human papillomavirus type 16 sequences in cervical carcinoma cell lines. *J Virol.* (1987) Apr;**61**(4):962-71.

Baldwin A, Pirisi L, Creek KE. NFI-Ski interactions mediate transforming growth factor beta modulation of human papillomavirus type 16 early gene expression. *J Virol.* (2004) Apr;**78**(8):3953-64.

Band V, De Caprio JA, Delmolino L, Kulesa V, Sager R. Loss of p53 protein in human papillomavirus type 16 E6 immortalized human mammary epithelial cells. *J Virol.* (1991) Dec;**65**(12):6671-6.

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

Banks L, Matlashewski G, Pim D, Churcher M, Roberts C, Crawford L. Expression of human papillomavirus type 6 and type 16 capsid proteins in bacteria and their antigenic characterization. *J Gen Virol.* (1987) Dec;**68** (Pt 12):3081-9.

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

Barbosa MS, Lowy DR, Schiller JT. Papillomavirus polypeptides E6 and E7 are zinc-binding proteins. *J Virol.* (1989) Mar;**63**(3):1404-7.

Bargonetti J, Manfredi JJ, Chen X, Marshak DR, Prives C. A proteolytic fragment from the central region of p53 has marked sequence-specific DNA-binding activity when generated from wild-type but not from oncogenic mutant p53 protein. *Genes Dev.* (1993) Dec;**7**(12B):2565-74.

Barth AI, Nathke IS, Nelson WJ. Cadherins, catenins and APC protein: interplay between cytoskeletal complexes and signaling pathways. *Curr Opin Cell Biol.* (1997) Oct;**9**(5):683-90.

References

Barton SE, Maddox PH, Jenkins D, Edwards R, Cuzick J, Singer A. Effect of cigarette smoking on cervical epithelial immunity: a mechanism for neoplastic change? *Lancet.* (1988) Sep 17;**2**(8612):652-4.

Baseman JG, Koutsky LA. The epidemiology of human papillomavirus infections. *J Clin Virol.* (2005) Mar;**32** Suppl 1:S16-24.

Be X, Hong Y, Wei J, Androphy EJ, Chen JJ, Baleja JD. Solution structure determination and mutational analysis of the papillomavirus E6 interacting peptide of E6AP. *Biochemistry.* (2001) Feb 6;**40**(5):1293-9.

Becamel C, Figge A, Poliak S, Dumuis A, Peles E, Bockaert J, Lubbert H, Ullmer C. Interaction of serotonin 5 hydroxytryptamine type 2C receptors with PDZ10 of the multi-PDZ domain protein MUPP1. *J Biol Chem.* (2001) Apr 20;**276**(16):12974-82.

Bellanger S, Blachon S, Mechali F, Bonne-Andrea C, Thierry F. High-risk but not low-risk HPV E2 proteins bind to the APC activators Cdh1 and Cdc20 and cause genomic instability. *Cell Cycle.* (2005) Nov;**4**(11):1608-15.

Berezutskaya E, Yu B, Morozov A, Raychaudhuri P, Bagchi S. Differential regulation of the pocket domains of the retinoblastoma family proteins by the HPV16 E7 oncoprotein. *Cell Growth Differ.* (1997) Dec;**8**(12):1277-86.

Berg T. Modulation of protein-protein interactions with small organic molecules. *Angew Chem Int Ed Engl.* (2003) Jun 6;**42**(22):2462-81.

Bezprozvanny I, Maximov A. PDZ domains: More than just a glue. Proc Natl Acad Sci U S A. 2001 Jan 30;98(3):787-9.

Bilder D, Li M, Perrimon N. Cooperative regulation of cell polarity and growth by Drosophila tumor suppressors. *Science.* (2000) Jul 7;**289**(5476):113-6.

Bilder D, Perrimon N. Localization of apical epithelial determinants by the basolateral PDZ protein Scribble. *Nature.* (2000) Feb 10;**403**(6770):676-80.

Birrane G, Chung J, Ladias JA. Novel mode of ligand recognition by the Erbin PDZ domain. *J Biol Chem.* (2003) Jan 17;**278**(3):1399-402.

Blachon S, Bellanger S, Demeret C, Thierry F. Nucleo-cytoplasmic shuttling of high risk human Papillomavirus E2 proteins induces apoptosis. *J Biol Chem.* (2005) Oct 28;**280**(43):36088-98.

Blachon S, Demeret C. The regulatory E2 proteins of human genital papillomaviruses are pro-apoptotic. *Biochimie.* (2003) Aug;**85**(8):813-9.

Bohl J, Das K, Dasgupta B, Vande Pol SB. Competitive binding to a charged leucine motif represses transformation by a papillomavirus E6 oncoprotein. *Virology.* (2000) May 25;**271**(1):163-70.

Bonne-Andrea C, Santucci S, Clertant P. Bovine papillomavirus E1 protein can, by itself, efficiently drive multiple rounds of DNA synthesis in vitro. *J Virol.* (1995) May;**69**(5):3201-5.

Boshart M, Gissmann L, Ikenberg H, Klein heinz A, Scheurlen W, zur Hausen H. A new type of papillomavirus DNA, its presence in genital cancer biopsies and in cell lines derived from cervical cancer. *EMBO J.* (1984) May;**3**(5):1151-7.

Bouvard V, Storey A, Pim D, Banks L. Characterization of the human papillomavirus E2 protein: evidence of trans-activation and trans-repression in cervical keratinocytes. *EMBO J.* (1994) Nov 15;**13**(22):5451-9.

Boyer SN, Wazer DE, Band V. E7 protein of human papilloma virus-16 induces degradation of retinoblastoma protein through the ubiquitin-proteasome pathway. *Cancer Res.* (1996) Oct 15;**56**(20):4620-4.

Bredt DS, Hwang PM, Glatt CE, Lowenstein C, Reed RR, Snyder SH. Cloned and expressed nitric oxide synthase structurally resembles cytochrome P-450 reductase. *Nature.* (1991) Jun 27;**351**(6329):714-8.

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

Brenman JE, Chao DS, Gee SH, McGee AW, Craven SE, Santillano DR, Wu Z, Huang F, Xia H, Peters MF, Froehner SC, Bredt DS. Interaction of nitric oxide synthase with the postsynaptic density protein PSD-95 and alpha1-syntrophin mediated by PDZ domains. *Cell.* (1996) Mar 8;**84**(5):757-67.

Brisson J, Morin C, Fortier M, Roy M, Bouchard C, Leclerc J, Christen A, Guimont C, Penault F, Meisels A. Risk factors for cervical intraepithelial neoplasia: differences between low- and high-grade lesions. *Am J Epidemiol.* (1994) Oct 15;**140**(8):700-10.

Bryant PJ, Watson KL, Justice RW, Woods DF. Tumor suppressor genes encoding proteins required for cell interactions and signal transduction in Drosophila. *Dev Suppl.* (1993);:239-49.

References

Burd EM. Human papillomavirus and cervical cancer. *Clin Microbiol Rev.* (2003) Jan;**16**(1):1-17.

Burkhardt A, Willingham M, Gay C, Jeang KT, Schlegel R. The E5 oncoprotein of bovine papillomavirus is oriented asymmetrically in Golgi and plasma membranes. *Virology.* (1989) May;**170**(1):334-9.

Burnett AF, Barnes WA, Johnson JC, Grendys E, Willett GD, Barter JF, Doniger J. Prognostic significance of polymerase chain reaction detected human papillomavirus of tumors and lymph nodes in surgically treated stage IB cervical cancer. *Gynecol Oncol.* (1992) Dec;**47**(3):343-7.

Butz K, Ristriani T, Hengstermann A, Denk C, Scheffner M, Hoppe-Seyler F. siRNA targeting of the viral E6 oncogene efficiently kills human papillomavirus-positive cancer cells. *Oncogene.* (2003) Sep 4;**22**(38):5938-45.

Cai C, Coleman SK, Niemi K, Keinanen K. Selective binding of synapse-associated protein 97 to GluR-A alpha-amino-5 hydroxy-3-methyl-4-isoxazole propionate receptor subunit is determined by a novel sequence motif. *J Biol Chem.* (2002) Aug 30;**277**(35):31484-90.

Cai C, Li H, Rivera C, Keinanen K. Interaction between SAP97 and PSD-95, two Maguk proteins involved in synaptic trafficking of AMPA receptors. *J Biol Chem.* (2006) Feb 17;**281**(7):4267-73.

Cao TT, Deacon HW, Reczek D, Bretscher A, von Zastrow M. A kinase-regulated PDZ-domain interaction controls endocytic sorting of the beta2-adrenergic receptor. *Nature.* (1999) Sep 16;**401**(6750):286-90.

Caruana G, Bernstein A. Craniofacial dysmorphogenesis including cleft palate in mice with an insertional mutation in the discs large gene. *Mol Cell Biol.* (2001) Mar;**21**(5):1475-83.

Cavanagh J, Fairbrother WJ, Palmer AG, Skelton NJ. Protein NMR Spectroscopy Principles and Practice. *Academic Press* (1995) ISBN: 978-0-12-164490-1

Chambers SP, Austen DA, Fulghum JR, Kim WM. High-throughput screening recombinant kinases in Escherichia coli and insect cells. *Protein Expr Purif.* (2004) Jul ;**36**(1):40-7

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

Charbonnier S, Coutouly MA, Kieffer B, Trave G, Atkinson RA. (13)C, (15)N and (1)H Resonance Assignment of the PDZ1 domain of MAGI-1 using QUASI. *J Biomol NMR.* (2006) Apr **24**; [Epub ahead of print] PMID: 16636753

Charbonnier S, Zanier K, Masson M, Trave G. Capturing protein-protein complexes at equilibrium: The holdup comparative chromatographic retention assay. *Protein Expr Purif.* (2006) Nov;**50**(1):89-101

Charest A, Lane K, McMahon K, Housman DE. Association of a novel PDZ domain-containing peripheral Golgi protein with the Q-SNARE (Q-soluble N-ethylmaleimide-sensitive fusion protein (NSF) attachment protein receptor) protein syntaxin 6. *J Biol Chem.* (2001) Aug 3;**276**(31):29456-65.

Charych EI, Akum BF, Goldberg JS, Jornsten RJ, Rongo C, Zheng JQ, Firestein BL. Activity-independent regulation of dendrite patterning by postsynaptic density protein PSD-95. *J Neurosci.* (2006) Oct 4;**26**(40):10164-76.

Chellappan S, Kraus VB, Kroger B, Munger K, Howley PM, Phelps WC, Nevins JR. Adenovirus E1A, simian virus 40 tumor antigen, and human papillomavirus E7 protein share the capacity to disrupt the interaction between transcription factor E2F and the retinoblastoma gene product. *Proc Natl Acad Sci U S A.* (1992) May 15;**89**(10):4549-53.

Chen G, Stenlund A. Characterization of the DNA-binding domain of the bovine papillomavirus replication initiator E1. *J Virol.* (1998) Apr;**72**(4):2567-76.

Chen JJ, Hong Y, Rustamzadeh E, Baleja JD, Androphy EJ. Identification of an alpha helical motif sufficient for association with papillomavirus E6. *J Biol Chem.* (1998) May 29;**273**(22):13537-44.

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

Chen YH, Huang LH, Chen TM. Differential effects of progestins and estrogens on long control regions of human papillomavirus types 16 and 18. *Biochem Biophys Res Commun.* (1996) Jul 25;**224**(3):651-9.

Chen Z, Kamath P, Zhang S, Weil MM, Shillitoe EJ. Effectiveness of three ribozymes for cleavage of an RNA transcript from human papillomavirus type 18. *Cancer Gene Ther.* (1995) Dec;**2**(4):263-71.

References

Cheng J, Moyer BD, Milewski M, Loffing J, Ikeda M, Mickle JE, Cutting GR, Li M, Stanton BA, Guggino WB. A Golgi-associated PDZ domain protein modulates cystic fibrosis transmembrane regulator plasma membrane expression. *J Biol Chem.* (2002) Feb 1;**277**(5):3520-9.

Cheng J, Wang H, Guggino WB. Regulation of cystic fibrosis transmembrane regulator trafficking and protein expression by a Rho family small GTPase TC10. *J Biol Chem.* (2005) Feb 4;**280**(5):3731-9.

Chevesich J, Kreuz AJ, Montell C. Requirement for the PDZ domain protein, INAD, for localization of the TRP store-operated channel to a signaling complex. *Neuron.* (1997) Jan;**18**(1):95-105.

Chiang SH, Baumann CA, Kanzaki M, Thurmond DC, Watson RT, Neudauer CL, Macara IG, Pessin JE, Saltiel AR. Insulin-stimulated GLUT4 translocation requires the CAP-dependent activation of TC10. *Nature.* (2001) Apr 19;**410**(6831):944-8.

Cho Y, Gorina S, Jeffrey PD, Pavletich NP. Crystal structure of a p53 tumor suppressor-DNA complex: understanding tumorigenic mutations. *Science.* (1994) Jul 15;**265**(5170):346-55.

Cho KO, Hunt CA, Kennedy MB. The rat brain postsynaptic density fraction contains a homolog of the Drosophila discs-large tumor suppressor protein. *Neuron.* (1992) Nov;**9**(5):929-42.

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

Clements A, Johnston K, Mazzarelli JM, Ricciardi RP, Marmorstein R. Oligomerization properties of the viral oncoproteins adenovirus E1A and human papillomavirus E7 and their complexes with the retinoblastoma protein. *Biochemistry.* (2000) Dec 26;**39**(51):16033-45.

Clore GM, Omichinski JG, Sakaguchi K, Zambrano N, Sakamoto H, Appella E, Gronenborn AM. High-resolution structure of the oligomerization domain of p53 by multidimensional NMR. *Science.* (1994) Jul 15;**265**(5170):386-91. **Erratum in:** *Science.* (1995) Mar 10;**267**(5203):1515-6.

Cohen NA, Brenman JE, Snyder SH, Bredt DS. Binding of the inward rectifier K+ channel Kir 2.3 to PSD-95 is regulated by protein kinase A phosphorylation. *Neuron.* (1996) Oct;**17**(4):759-67.

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

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

Conrad M, Goldstein D, Andresson T, Schlegel R. The E5 protein of HPV-6, but not HPV-16, associates efficiently with cellular growth factor receptors. *Virology.* (1994) May 1;**200**(2):796-800.

Cooper B, Schneider S, Bohl J, Jiang Y, Beaudet A, Vande Pol S. Requirement of E6AP and the features of human papillomavirus E6 necessary to support degradation of p53. *Virology.* (2003) Feb 1;**306**(1):87-99.

Corden SA, Sant-Cassia LJ, Easton AJ, Morris AG. The integration of HPV-18 DNA in cervical carcinoma. *Mol Pathol.* (1999) Oct;**52**(5):275-82.

Cowburn D. Peptide recognition by PTB and PDZ domains. *Curr Opin Struct Biol.* (1997) Dec;**7**(6):835-8.

Cox LS, Lane DP. Tumour suppressors, kinases and clamps: how p53 regulates the cell cycle in response to DNA damage. *Bioessays.* (1995) Jun;**17**(6):501-8.

Craven SE, Bredt DS. PDZ proteins organize synaptic signaling pathways. *Cell.* (1998) May 15;**93**(4):495-8.

Crook T, Wrede D, Tidy J, Scholefield J, Crawford L, Vousden KH. Status of c-myc, p53 and retinoblastoma genes in human papillomavirus positive and negative squamous cell carcinomas of the anus. *Oncogene.* (1991) Jul;**6**(7):1251-7.

Crook T, Marston NJ, Sara EA, Vousden KH. Transcriptional activation by p53 correlates with suppression of growth but not transformation. *Cell.* (1994) Dec 2;**79**(5):817-27.

Dang CV, Lee WM. Nuclear and nucleolar targeting sequences of c-erb-A, c-myb, N-myc, p53, HSP70, and HIV tat proteins. *J Biol Chem.* (1989) Oct 25;**264**(30):18019-23.

Daniels PR, Sanders CM, Coulson P, Maitland NJ. Molecular analysis of the interaction between HPV type 16 E6 and human E6-associated protein. *FEBS Lett.* (1997) Oct 13;**416**(1):6-10.

References

Daniels PR, Sanders CM, Maitland NJ. Characterization of the interactions of human papillomavirus type 16 E6 with p53 and E6-associated protein in insect and human cells. *J Gen Virol.* (1998) Mar;**79** (Pt 3):489-99.

Davy CE, Jackson DJ, Wang Q, Raj K, Masterson PJ, Fenner NF, Southern S, Cuthill S, Millar JB, Doorbar J. Identification of a G(2) arrest domain in the E1 wedge E4 protein of human papillomavirus type 16. *J Virol.* (2002) Oct;**76**(19):9806-18.

Daw MI, Chittajallu R, Bortolotto ZA, Dev KK, Duprat F, Henley JM, Collingridge GL, Isaac JT. PDZ proteins interacting with C-terminal GluR2/3 are involved in a PKC-dependent regulation of AMPA receptors at hippocampal synapses. *Neuron.* (2000) Dec;**28**(3):873-86.

De Villiers EM, Fauquet C, Broker TR, Bernard HU, zur Hausen H. Classification of papillomaviruses. *Virology.* (2004) Jun 20;**324**(1):17-27.

De Vries L, Lou X, Zhao G, Zheng B, Farquhar MG. GIPC, a PDZ domain containing protein, interacts specifically with the C terminus of RGS-GAIP. *Proc Natl Acad Sci U S A.* (1998) Oct 13;**95**(21):12340-5.

Delaglio F, Torchia DA, Bax A. Measurement of 15N-13C J couplings in staphylococcal nuclease. *J Biomol NMR.* (1991) Nov;**1**(4):439-46.

Delaglio F, Grzesiek S, Vuister GW, Zhu G, Pfeifer J, Bax A. NMRPipe: a multidimensional spectral processing system based on UNIX pipes. *J Biomol NMR.* (1995) Nov;**6**(3):277-93.

Dell G, Gaston K. Human papillomaviruses and their role in cervical cancer. *Cell Mol Life Sci.* (2001) Nov;**58**(12-13):1923-42.

Demers GW, Foster SA, Halbert CL, Galloway DA. Growth arrest by induction of p53 in DNA damaged keratinocytes is bypassed by human papillomavirus 16 E7. *Proc Natl Acad Sci U S A.* (1994) May 10;**91**(10):4382-6.

Denny LA, Wright TC Jr. Human papillomavirus testing and screening. *Best Pract Res Clin Obstet Gynaecol.* (2005) Aug;**19**(4):501-15.

Desaintes C, Demeret C, Goyat S, Yaniv M, Thierry F. Expression of the papillomavirus E2 protein in HeLa cells leads to apoptosis. *EMBO J.* (1997) Feb 3;**16**(3):504-14.

Dev KK. Making protein interactions druggable: targeting PDZ domains. *Nat Rev Drug Discov.* (2004) Dec;**3**(12):1047-56.

Dey A, Atcha IA, Bagchi S. HPV16 E6 oncoprotein stimulates the transforming growth factor-beta 1 promoter in fibroblasts through a specific GC-rich sequence. *Virology.* (1997) Feb 17;**228**(2):190-9.

Dimitratos SD, Woods DF, Stathakis DG, Bryant PJ. Signaling pathways are focused at specialized regions of the plasma membrane by scaffolding proteins of the MAGUK family. *Bioessays.* (1999) Nov;**21**(11):912-21.

Dinitto JP, Lambright DG. Membrane and juxtamembrane targeting by PH and PTB domains. *Biochim Biophys Acta.* (2006) Aug;**1761**(8):850-67.

Dobrosotskaya IY. Identification of mNET1 as a candidate ligand for the first PDZ domain of MAGI-1. *Biochem Biophys Res Commun.* (2001) May 18;**283**(4):969-75.

Dobrosotskaya I, Guy RK, James GL. MAGI-1, a membrane-associated guanylate kinase with a unique arrangement of protein-protein interaction domains. *J Biol Chem.* (1997) Dec 12;**272**(50):31589-97.

Dobrosotskaya IY, James GL. MAGI-1 interacts with beta-catenin and is associated with cell-cell adhesion structures. *Biochem Biophys Res Commun.* (2000) Apr 21;**270**(3):903-9.

Dong H, O'Brien RJ, Fung ET, Lanahan AA, Worley PF, Huganir RL. GRIP: a synaptic PDZ domain-containing protein that interacts with AMPA receptors. *Nature.* (1997) Mar 20;**386**(6622):279-84.

Doorbar J. The papillomavirus life cycle. *J Clin Virol.* (2005) Mar;**32** Suppl 1:S7-15.

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

Doorbar J, Foo C, Coleman N, Medcalf L, Hartley O, Prospero T, Napthine S, Sterling J, Winter G, Griffin H. Characterization of events during the late stages of HPV16 infection in vivo using high-affinity synthetic Fabs to E4. *Virology.* (1997) Nov 10;**238**(1):40-52.

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

References

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

Dostatni N, Lambert PF, Sousa R, Ham J, Howley PM, Yaniv M. The functional BPV-1 E2 trans-activating protein can act as a repressor by preventing formation of the initiation complex. *Genes Dev.* (1991) Sep;**5**(9):1657-71.

Dow LE, Brumby AM, Muratore R, Coombe ML, Sedelies KA, Trapani JA, Russell SM, Richardson HE, Humbert PO. hScrib is a functional homologue of the Drosophila tumour suppressor Scribble. *Oncogene.* (2003) Dec 18;**22**(58):9225-30.

Dowhanick JJ, McBride AA, Howley PM. Suppression of cellular proliferation by the papillomavirus E2 protein. *J Virol.* (1995) Dec;**69**(12):7791-9.

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

Duensing S, Lee LY, Duensing A, Basile J, Piboonniyom S, Gonzalez S, Crum CP, Munger K. The human papillomavirus type 16 E6 and E7 oncoproteins cooperate to induce mitotic defects and genomic instability by uncoupling centrosome duplication from the cell division cycle. *Proc Natl Acad Sci U S A.* (2000) Aug 29;**97**(18):10002-7.

Duensing S, Munger K. Mechanisms of genomic instability in human cancer: insights from studies with human papillomavirus oncoproteins. *Int J Cancer.* (2004) Mar 20;**109**(2):157-62.

Durst M, Gissmann L, Ikenberg H, zur Hausen H. A papillomavirus DNA from a cervical carcinoma and its prevalence in cancer biopsy samples from different geographic regions. *Proc Natl Acad Sci USA.* (1983) Jun;**80**(12):3812-5.

Durst M, Kleinheinz A, Hotz M, Gissmann L. The physical state of human papillomavirus type 16 DNA in benign and malignant genital tumours. *J Gen Virol.* (1985) Jul;**66** (Pt 7):1515-22.

Dyson N, Guida P, Munger K, Harlow E. Homologous sequences in adenovirus E1A and human papillomavirus E7 proteins mediate interaction with the same set of cellular proteins. *J Virol.* (1992) Dec;**66**(12):6893-902.

Eck MJ, Dhe-Paganon S, Trub T, Nolte RT, Shoelson SE. Structure of the IRS-1 PTB domain bound to the juxtamembrane region of the insulin receptor. *Cell.* (1996) May 31;**85**(5):695-705.

el-Deiry WS, Kern SE, Pietenpol JA, Kinzler KW, Vogelstein B. Definition of a consensus binding site for p53. *Nat Genet.* (1992) Apr;**1**(1):45-9.

El-Husseini AE, Schnell E, Chetkovich DM, Nicoll RA, Bredt DS. PSD-95 involvement in maturation of excitatory synapses. *Science.* (2000) Nov 17;**290**(5495):1364-8.

Elson DA, Riley RR, Lacey A, Thordarson G, Talamantes FJ, Arbeit JM. Sensitivity of the cervical transformation zone to estrogen-induced squamous carcinogenesis. *Cancer Res.* (2000) Mar 1;**60**(5):1267-75.

Elston RC, Napthine S, Doorbar J. The identification of a conserved binding motif within human papillomavirus type 16 E6 binding peptides, E6AP and E6BP. *J Gen Virol.* (1998) Feb;**79** (Pt 2):371-4.

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

Esposito D, Chatterjee DK. Enhancement of soluble protein expression through the use of fusion tags. *Curr Opin Biotechnol.* (2006) Aug;**17**(4):353-8.

Fahey MT, Irwig L, Macaskill P. Meta-analysis of Pap test accuracy. *Am J Epidemiol.* (1995) Apr 1;**141**(7):680-9.

Fanning AS, Anderson JM. PDZ domains: fundamental building blocks in the organization of protein complexes at the plasma membrane. *J Clin Invest.* (1999) Mar;**103**(6):767-72.

Faulkner G, Pallavicini A, Formentin E, Comelli A, Ievolella C, Trevisan S, Bortoletto G, Scannapieco P, Salamon M, Mouly V, Valle G, Lanfranchi G. ZASP: a new Z-band alternatively spliced PDZ-motif protein. *J Cell Biol.* (1999) Jul 26;**146**(2):465-75.

Favre-Bonvin A, Reynaud C, Kretz-Remy C, Jalinot P. 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.* (2005) Apr;**79**(7):4229-37.

Ferrer M, Maiolo J, Kratz P, Jackowski JL, Murphy DJ, Delagrave S, Inglese J. Directed evolution of PDZ variants to generate high-affinity detection reagents. *Protein Eng Des Sel.* (2005) Apr;**18**(4):165-73.

References

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

Finbow ME, Pitts JD. Is the gap junction channel--the connexon--made of connexin or ductin? *J Cell Sci.* (1993) Oct;**106** (Pt 2):463-71.

Fischer PM, Lane DP. Small-molecule inhibitors of the p53 suppressor HDM2: have protein-protein interactions come of age as drug targets? *Trends Pharmacol Sci.* (2004) Jul;**25**(7):343-6.

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

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

Francis DA, Schmid SI, Howley PM. Repression of the integrated papillomavirus E6/E7 promoter is required for growth suppression of cervical cancer cells. *J Virol.* (2000) Mar;**74**(6):2679-86.

Frattini MG, Hurst SD, Lim HB, Swaminathan S, Laimins LA. Abrogation of a mitotic checkpoint by E2 proteins from oncogenic human papillomaviruses correlates with increased turnover of the p53 tumor suppressor protein. *EMBO J.* (1997) Jan 15;**16**(2):318-31.

Frattini MG, Laimins LA. Binding of the human papillomavirus E1 origin-recognition protein is regulated through complex formation with the E2 enhancer-binding protein. *Proc Natl Acad Sci U S A.* (1994) Dec 20;**91**(26):12398-402.

Fujii N, Haresco JJ, Novak KA, Stokoe D, Kuntz ID, Guy RK. A selective irreversible inhibitor targeting a PDZ protein interaction domain. *J Am Chem Soc.* (2003) Oct 8;**125**(40):12074-5.

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

Funke L, Dakoji S, Bredt DS. Membrane-associated guanylate kinases regulate adhesion and plasticity at cell junctions. *Annu Rev Biochem.* (2005);**74**:219-45.

Gao Q, Kumar A, Srinivasan S, Singh L, Mukai H, Ono Y, Wazer DE, Band V. PKN binds and phosphorylates human papillomavirus E6 oncoprotein. *J Biol Chem.* (2000) May 19;**275**(20):14824-30.

Gardiol D, Kuhne C, Glaunsinger B, Lee SS, Javier R, Banks L. Oncogenic human papillomavirus E6 proteins target the discs large tumour suppressor for proteasome-mediated degradation. *Oncogene.* (1999) Sep 30;**18**(40):5487-96.

Gates J, Peifer M. Can 1000 reviews be wrong? Actin, alpha-Catenin, and adherens junctions. *Cell.* (2005) Dec 2;**123**(5):769-72.

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

Gentzsch M, Cui L, Mengos A, Chang XB, Chen JH, Riordan JR. The PDZ-binding chloride channel ClC-3B localizes to the Golgi and associates with cystic fibrosis transmembrane conductance regulator-interacting PDZ proteins. *J Biol Chem.* (2003) Feb 21;**278**(8):6440-9.

Gewin L, Myers H, Kiyono T, Galloway DA. Identification of a novel telomerase repressor that interacts with the human papillomavirus type-16 E6/E6-AP complex. *Genes Dev.* (2004) Sep 15;**18**(18):2269-82.

Gisler SM, Pribanic S, Bacic D, Forrer P, Gantenbein A, Sabourin LA, Tsuji A, Zhao ZS, Manser E, Biber J, Murer H. PDZK1: I. a major scaffolder in brush borders of proximal tubular cells. *Kidney Int.* (2003) Nov;**64**(5):1733-45.

Glahder JA, Hansen CN, Vinther J, Madsen BS, Norrild B. A promoter within the E6 ORF of human papillomavirus type 16 contributes to the expression of the E7 oncoprotein from a monocistronic mRNA. *J Gen Virol.* (2003) Dec;**84**(Pt 12):3429-41.

Glaunsinger BA, Lee SS, Thomas M, Banks L, Javier R. Interactions of the PDZ-protein MAGI-1 with adenovirus E4-ORF1 and high-risk papillomavirus E6 oncoproteins. *Oncogene.* (2000) Nov 2;**19**(46):5270-80.

Gonzalez-Mariscal L, Betanzos A, Avila-Flores A. MAGUK proteins: structure and role in the tight junction. *Semin Cell Dev Biol.* (2000) Aug;**11**(4):315-24.

Gonzalez-Mariscal L, Betanzos A, Nava P, Jaramillo BE. Tight junction proteins. *Prog Biophys Mol Biol.* (2003) Jan;**81**(1):1-44.

References

Goodwin EC, DiMaio D. Repression of human papillomavirus oncogenes in HeLa cervical carcinoma cells causes the orderly reactivation of dormant tumor suppressor pathways. *Proc Natl Acad Sci U S A.* (2000) Nov 7;**97**(23):12513-8.

Goodwin EC, Naeger LK, Breiding DE, Androphy EJ, DiMaio D. Transactivation-competent bovine papillomavirus E2 protein is specifically required for efficient repression of human papillomavirus oncogene expression and for acute growth inhibition of cervical carcinoma cell lines. *J Virol.* (1998) May;**72**(5):3925-34.

Gorina S, Pavletich NP. Structure of the p53 tumor suppressor bound to the ankyrin and SH3 domains of 53BP2. *Science.* (1996) Nov 8;**274**(5289):1001-5.

Grana X, Garriga J, Mayol X. Role of the retinoblastoma protein family, pRB, p107 and p130 in the negative control of cell growth. *Oncogene.* (1998) Dec 24;**17**(25):3365-83.

Greaves S. Growth and polarity: the case for scribble. *Nat Cell Biol.* (2000) Aug;**2**(8):E140.

Grm HS, Massimi P, Gammoh N, Banks L. Crosstalk between the human papillomavirus E2 transcriptional activator and the E6 oncoprotein. *Oncogene.* (2005) Aug 4;**24**(33):5149-64.

Grootjans JJ, Reekmans G, Ceulemans H, David G. Syntenin-syndecan binding requires syndecan-synteny and the co operation of both PDZ domains of syntenin. *J Biol Chem.* (2000) Jun 30;**275**(26):19933-41.

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

Guccione E, Massimi P, Bernat A, Banks L. Comparative analysis of the intracellular location of the high- and low-risk human papillomavirus oncoproteins. *Virology.* (2002) Feb 1;**293**(1):20-5.

Hakama M, Miller AB, Day NE. Screening for Cancer of the Uterine Cervix IARC Scientific Publications no. 76. Lyon: *International Agency for Research on Cancer*, (1986), pp. 47–60.

Hakama M, Louhivuori K. A screening programme for cervical cancer that worked. *Cancer Surveys* (1988); **17**: 403–416.

Hall RA, Ostedgaard LS, Premont RT, Blitzer JT, Rahman N, Welsh MJ, Lefkowitz RJ. A C-terminal motif found in the beta2-adrenergic receptor, P2Y1 receptor and cystic fibrosis transmembrane conductance regulator determines binding to the Na+/H+ exchanger regulatory factor family of PDZ proteins. *Proc Natl Acad Sci U S A.* (1998a) Jul 21;**95**(15):8496-501.

Hall RA, Premont RT, Chow CW, Blitzer JT, Pitcher JA, Claing A, Stoffel RH, Barak LS, Shenolikar S, Weinman EJ, Grinstein S, Lefkowitz RJ. The beta2-adrenergic receptor interacts with the Na+/H+-exchanger regulatory factor to control Na+/H+ exchange. *Nature.* (1998b) Apr 9;**392**(6676):626-30.

Halpert R, Fruchter RG, Sedlis A, Butt K, Boyce JG, Sillman FH. Human papillomavirus and lower genital neoplasia in renal transplant patients. *Obstet Gynecol.* (1986) Aug;**68**(2):251-8.

Ham J, Steger G, Yaniv M. Cooperativity in vivo between the E2 transactivator and the TATA box binding protein depends on core promoter structure. *EMBO J.* (1994) Jan 1;**13**(1):147-57.

Hampson L, Li C, Oliver AW, Kitchener HC, Hampson IN. The PDZ protein Tip-1 is a gain of function target of the HPV16 E6 oncoprotein. *Int J Oncol.* (2004) Nov;**25**(5):1249-56.

Harris BZ, Lim WA. Mechanism and role of PDZ domains in signaling complex assembly. *J Cell Sci.* (2001) Sep;**114**(Pt 18):3219-31.

Harrison SC. Peptide-surface association: the case of PDZ and PTB domains. *Cell.* (1996) Aug 9;**86**(3):341-3.

Harteneck C. Proteins modulating TRP channel function. *Cell Calcium.* (2003) May-Jun;**33**(5-6):303-10.

Hassel B, Schreff M, Stube EM, Blaich U, Schumacher S. CALEB/NGC interacts with the Golgi-associated protein PIST. *J Biol Chem.* (2003) Oct 10;**278**(41):40136-43.

Haupt Y, Maya R, Kazaz A, Oren M. Mdm2 promotes the rapid degradation of p53. *Nature.* (1997) May 15;**387**(6630):296-9.

Hawley-Nelson P, Vousden KH, Hubbert NL, Lowy DR, Schiller JT. HPV16 E6 and E7 proteins cooperate to immortalize human foreskin keratinocytes. *EMBO J.* (1989) Dec 1;**8**(12):3905-10.

Hillier BJ, Christopherson KS, Prehoda KE, Bredt DS, Lim WA. Unexpected modes of PDZ domain scaffolding revealed by structure of nNOS-syntrophin complex. *Science.* (1999) Apr 30;**284**(5415):812-5.

References

Hirabayashi S, Tajima M, Yao I, Nishimura W, Mori H, Hata Y. JAM4, a junctional cell adhesion molecule interacting with a tight junction protein, MAGI-1. *Mol Cell Biol.* (2003) Jun;**23**(12):4267-82.

Hirao K, Hata Y, Ide N, Takeuchi M, Irie M, Yao I, Deguchi M, Toyoda A, Sudhof TC, Takai Y. A novel multiple PDZ domain-containing molecule interacting with N-methyl-D-aspartate receptors and neuronal cell adhesion proteins. *J Biol Chem.* (1998) Aug 14;**273**(33):21105-10.

Hirao K, Hata Y, Yao I, Deguchi M, Kawabe H, Mizoguchi A, Takai Y. Three isoforms of synaptic scaffolding molecule and their characterization. Multimerization between the isoforms and their interaction with N-methyl-D-aspartate receptors and SAP90/PSD-95-associated protein. *J Biol Chem.* (2000) Jan 28;**275**(4):2966-72.

Hirbec H, Francis JC, Lauri SE, Braithwaite SP, Coussen F, Mulle C, Dev KK, Coutinho V, Meyer G, Isaac JT, Collingridge GL, Henley JM. Rapid and differential regulation of AMPA and kainate receptors at hippocampal mossy fibre synapses by PICK1 and GRIP. *Neuron.* (2003) Feb 20;**37**(4):625-38. **Erratum in:** *Neuron.* (2003) May 22;**38**(4):673.

Hollstein M, Shomer B, Greenblatt M, Soussi T, Hovig E, Montesano R, Harris CC. Somatic point mutations in the p53 gene of human tumors and cell lines: updated compilation. *Nucleic Acids Res.* (1996) Jan 1;**24**(1):141-6.

Honore B, Vorum H. The CREC family, a novel family of multiple EF-hand, low-affinity Ca(2+)-binding proteins localised to the secretory pathway of mammalian cells. *FEBS Lett.* (2000) Jan 21;**466**(1):11-8.

Huang L, Kinnucan E, Wang G, Beaudenon S, Howley PM, Huibregtse JM, Pavletich NP. Structure of an E6AP-UbcH7 complex: insights into ubiquitination by the E2-E3 enzyme cascade. *Science.* (1999) Nov 12;**286**(5443):1321-6.

Hudson JB, Bedell MA, McCance DJ, Laiminis LA. Immortalization and altered differentiation of human keratinocytes in vitro by the E6 and E7 open reading frames of human papillomavirus type 18. *J Virol.* (1990) Feb;**64**(2):519-26.

Huibregtse JM, Beaudenon SL. Mechanism of HPV E6 proteins in cellular transformation. *Semin Cancer Biol.* (1996) Dec;**7**(6):317-26.

Huibregtse JM, Scheffner M, Beaudenon S, Howley PM. A family of proteins structurally and functionally related to the E6-AP ubiquitin-protein ligase. *Proc Natl Acad Sci U S A.* (1995) Mar 28;**92**(7):2563-7. **Erratum in:** *Proc Natl Acad Sci U S A.* (1995) May 23;**92**(11):5249.

Huibregtse JM, Scheffner M, Howley PM. A cellular protein mediates association of p53 with the E6 oncoprotein of human papillomavirus types 16 or 18. *EMBO J.* (1991) Dec;**10**(13):4129-35.

Huibregtse JM, Scheffner M, Howley PM. Localization of the E6-AP regions that direct human papillomavirus E6 binding, association with p53, and ubiquitination of associated proteins. *Mol Cell Biol.* (1993) Aug;**13**(8):4918-27.

Humbert P, Russell S, Richardson H. Dlg, Scribble and Lgl in cell polarity, cell proliferation and cancer. *Bioessays.* (2003) Jun;**25**(6):542-53.

Hung AY, Sheng M. PDZ domains: structural modules for protein complex assembly. *J Biol Chem.* (2002) Feb 22;**277**(8):5699-702.

Hupp TR, Lane DP. Two distinct signaling pathways activate the latent DNA binding function of p53 in a casein kinase II-independent manner. *J Biol Chem.* (1995) Jul 28;**270**(30):18165-74.

Hwang ES, Riese DJ 2nd, Settleman J, Nilson LA, Honig J, Flynn S, DiMaio D. Inhibition of cervical carcinoma cell line proliferation by the introduction of a bovine papillomavirus regulatory gene. *J Virol.* (1993) Jul;**67**(7):3720-9.

Hwang ES, Naeger LK, DiMaio D. Activation of the endogenous p53 growth inhibitory pathway in HeLa cervical carcinoma cells by expression of the bovine papillomavirus E2 gene. *Oncogene.* (1996) Feb 15;**12**(4):795-803.

Ide N, Hata Y, Nishioka H, Hirao K, Yao I, Deguchi M, Mizoguchi A, Nishimori H, Tokino T, Nakamura Y, Takai Y. Localization of membrane-associated guanylate kinase (MAGI)-1/BAI-associated protein (BAP) 1 at tight junctions of epithelial cells. *Oncogene.* (1999) Dec 16;**18**(54):7810-5.

Ishidate T, Matsumine A, Toyoshima K, Akiyama T. The APC-hDLG complex negatively regulates cell cycle progression from the G0/G1 to S phase. *Oncogene.* (2000) Jan 20;**19**(3):365-72.

Iwabuchi K, Bartel PL, Li B, Marraccino R, Fields S. Two cellular proteins that bind to wild-type but not mutant p53. *Proc Natl Acad Sci U S A.* (1994) Jun 21;**91**(13):6098-102.

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

References

Jeffrey PD, Gorina S, Pavletich NP. Crystal structure of the tetramerization domain of the p53 tumor suppressor at 1.7 angstroms. *Science.* (1995) Mar 10;**267**(5203):1498-502.

Jelen F, Oleksy A, Smietana K, Otlewski J. PDZ domains - common players in the cell signaling. *Acta Biochim Pol.* (2003);**50**(4):985-1017.

Jenkins JR, Chumakov P, Addison C, Sturzbecher HW, Wade-Evans A. Two distinct regions of the murine p53 primary amino acid sequence are implicated in stable complex formation with simian virus 40 T antigen. *J Virol.* (1988) Oct;**62**(10):3903-6.

Jeong KW, Kim HZ, Kim S, Kim YS, Choe J. 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.* (2006) Jul 31; [Epub ahead of print] PMID: 16878151

Jesaitis LA, Goodenough DA. Molecular characterization and tissue distribution of ZO-2, a tight junction protein homologous to ZO-1 an the Drosophila discs-large tumor suppressor protein. *J Cell Biol.* (1994) Mar;**124**(6):949-61.

Jian Y, Schmidt-Grimminger DC, Chien WM, Wu X, Broker TR, Chow LT. Post-transcriptional induction of p21cip1 protein by human papillomavirus E7 inhibits unscheduled DNA synthesis reactivated in differentiated keratinocytes. *Oncogene.* (1998) Oct 22;**17**(16):2027-38.

Jones DL, Alani RM, Munger K. The human papillomavirus E7 oncoprotein can uncouple cellular differentiation and proliferation in human keratinocytes by abrogating p21Cip1-mediated inhibition of cdk2. *Genes Dev.* (1997) Aug 15;**11**(16):2101-11.

Jordan LB, Monaghan H. Pathology of the cervix: recent developments. *Clin Oncol (R Coll Radiol).* (2004) Jun;**16**(4):248-54.

Kalantari M, Blennow E, Hagmar B, Johansson B. Physical state of HPV16 and chromosomal mapping of the integrated form in cervical carcinomas. *Diagn Mol Pathol.* (2001) Mar;**10**(1):46-54.

Kanamori M, Sandy P, Marzinotto S, Benetti R, Kai C, Hayashizaki Y, Schneider C, Suzuki H. The PDZ protein tax-interacting protein-1 inhibits beta-catenin transcriptional activity and growth of colorectal cancer cells. *J Biol Chem.* (2003) Oct 3;**278**(40):38758-64.

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

Kang BS, Cooper DR, Devedjiev Y, Derewenda U, Derewenda ZS. Molecular roots of degenerate specificity in syntenin's PDZ2 domain: reassessment of the PDZ recognition paradigm. *Structure.* (2003) Jul;**11**(7):845-53.

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

Kato Y, Sai Y, Yoshida K, Watanabe C, Hirata T, Tsuji A. PDZK1 directly regulates the function of organic cation/carnitine transporter OCTN2. *Mol Pharmacol.* (2005) Mar;**67**(3):734-43.

Kato Y, Yoshida K, Watanabe C, Sai Y, Tsuji A. Screening of the interaction between xenobiotic transporters and PDZ proteins. *Pharm Res.* (2004) Oct;**21**(10):1886-94.

Kavanaugh WM, Williams LT. An alternative to SH2 domains for binding tyrosine-phosphorylated proteins. *Science.* (1994) Dec 16;**266**(5192):1862-5.

Kell B, Jewers RJ, Cason J, Pakarian F, Kaye JN, Best JM. Detection of E5 oncoprotein in human papillomavirus type 16-positive cervical scrapes using antibodies raised to synthetic peptides. *J Gen Virol.* (1994) Sep;**75** (Pt 9):2451-6.

Kelley ML, Keiger KE, Lee CJ, Huibregtse JM. The global transcriptional effects of the human papillomavirus E6 protein in cervical carcinoma cell lines are mediated by the E6AP ubiquitin ligase. *J Virol.* (2005) Mar;**79**(6):3737-47.

Kennedy MB. Origin of PDZ (DHR, GLGF) domains. *Trends Biochem Sci.* (1995) Sep;**20**(9):350.

Kennedy MB. Signal-processing machines at the postsynaptic density. *Science.* (2000) Oct 27;**290**(5492):750-4.

Kim E, Cho KO, Rothschild A, Sheng M. Heteromultimerization and NMDA receptor-clustering activity of Chapsyn-110, a member of the PSD-95 family of proteins. *Neuron.* (1996) Jul;**17**(1):103-13.

Kim E, Niethammer M, Rothschild A, Jan YN, Sheng M. Clustering of Shaker-type K+ channels by interaction with a family of membrane-associated guanylate kinases. *Nature.* (1995) Nov 2;**378**(6552):85-8.

Kim CJ, Um SJ, Kim TY, Kim EJ, Park TC, Kim SJ, Namkoong SE, Park JS. Regulation of cell growth and HPV genes by exogenous estrogen in cervical cancer cells. *Int J Gynecol Cancer.* (2000) Mar;**10**(2):157-164.

206

References

Kinzler KW, Vogelstein B. Lessons from hereditary colorectal cancer. *Cell.* (1996) Oct 18;**87**(2):159-70.

Kiselyov K, Kim JY, Zeng W, Muallem S. Protein-protein interaction and functionTRPC channels. *Pflugers Arch.* (2005) Oct;**451**(1):116-24.

Kistner U, Wenzel BM, Veh RW, Cases-Langhoff C, Garner AM, Appeltauer U, Voss B, Gundelfinger ED, Garner CC. SAP90, a rat presynaptic protein related to the product of the Drosophila tumor suppressor gene dlg-A. *J Biol Chem.* (1993) Mar 5;**268**(7):4580-3.

Kiyono T, Foster SA, Koop JI, McDougall JK, Galloway DA, Klingelhutz AJ. Both Rb/p16INK4a inactivation and telomerase activity are required to immortalize human epithelial cells. *Nature.* (1998) Nov 5;**396**(6706):84-8.

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

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

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

Ko LJ, Prives C. p53: puzzle and paradigm. *Genes Dev.* (1996) May 1;**10**(9):1054-72.

Kornau HC, Schenker LT, Kennedy MB, Seeburg PH. Domain interaction between NMDA receptor subunits and the postsynaptic density protein PSD-95. *Science.* (1995) Sep 22;**269**(5231):1737-40.

Koroll M, Rathjen FG, Volkmer H. The neural cell recognition molecule neurofascin interacts with syntenin-1 but not with syntenin-2, both of which reveal self-associating activity. *J Biol Chem.* (2001) Apr 6;**276**(14):10646-54.

Kotelevets L, van Hengel J, Bruyneel E, Mareel M, van Roy F, Chastre E. Implication of the MAGI-1b/PTEN signalosome in stabilization of adherens junctions and suppression of invasiveness. *FASEB J.* (2005) Jan;**19**(1):115-7.

Kozlov G, Banville D, Gehring K, Ekiel I. Solution structure of the PDZ2 domain from cytosolic human phosphatase hPTP1E complexed with a peptide reveals contribution of the beta2-beta3 loop to PDZ domain-ligand interactions. *J Mol Biol.* (2002) Jul 19;**320**(4):813-20.

Kraiss S, Quaiser A, Oren M, Montenarh M. Oligomerization of oncoprotein p53. *J Virol.* (1988) Dec;**62**(12):4737-44.

Kubbutat MH, Ludwig RL, Ashcroft M, Vousden KH. Regulation of Mdm2-directed degradation by the C terminus of p53. *Mol Cell Biol.* (1998) Oct;**18**(10):5690-8.

Kuhne C, Banks L. E3-ubiquitin ligase/E6-AP links multicopy maintenance protein 7 to the ubiquitination pathway by a novel motif, the L2G box. *J Biol Chem.* (1998) Dec 18;**273**(51):34302-9.

Kuhne C, Gardiol D, Guarnaccia C, Amenitsch H, Banks L. Differential regulation of human papillomavirus E6 by protein kinase A: conditional degradation of human discs large protein by oncogenic E6. *Oncogene.* (2000) Nov 30;**19**(51):5884-91.

Kukimoto I, Aihara S, Yoshiike K, Kanda T. Human papillomavirus oncoprotein E6 binds to the C-terminal region of human minichromosome maintenance 7 protein. *Biochem Biophys Res Commun* .(1998) Aug 10;**249**(1):258-62.

Kurman RJ, Schiffman MH, Lancaster WD, Reid R, Jenson AB, Temple GF, Lorincz AT. Analysis of individual human papillomavirus types in cervical neoplasia: a possible role for type 18 in rapid progression. *Am J Obstet Gynecol.* (1988) Aug;**159**(2):293-6.

Lagrange M, Charbonnier S, Orfanoudakis G, Robinson P, Zanier K, Masson M, Lutz Y, Trave G, Weiss E, Deryckere F. Binding of human papillomavirus 16 E6 to p53 and E6AP is impaired by monoclonal antibodies directed against the second zinc-binding domain of E6. *J Gen Virol.* (2005) Apr;**86**(Pt 4):1001-7.

Lahey T, Gorczyca M, Jia XX, Budnik V. The Drosophila tumor suppressor gene dlg is required for normal synaptic bouton structure. *Neuron.* (1994) Oct;**13**(4):823-35.

Lamprecht G, Weinman EJ, Yun CH. The role of NHERF and E3KARP in the cAMP-mediated inhibition of NHE3. *J Biol Chem.* (1998) Nov 6;**273**(45):29972-8.

Latorre IJ, Roh MH, Frese KK, Weiss RS, Margolis B, Javier RT. Viral oncoprotein-induced mislocalization of select PDZ proteins disrupts tight junctions and causes polarity defects in epithelial cells. *J Cell Sci.* (2005) Sep 15;**118**(Pt 18):4283-93.

References

Laura RP, Ross S, Koeppen H, Lasky LA. MAGI-1: a widely expressed, alternatively spliced tight junction protein. *Exp Cell Res.* (2002) May 1;**275**(2):155-70.

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

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

Lee JO, Russo AA, Pavletich NP. Structure of the retinoblastoma tumour-suppressor pocket domain bound to a peptide from HPV E7. *Nature.* (1998) Feb 26;**391**(6670):859-65.

Lee SS, Weiss RS, Javier RT. Binding of human virus oncoproteins to hDlg/SAP97, a mammalian homolog of the Drosophila discs large tumor suppressor protein. *Proc Natl Acad Sci U S A.* (1997) Jun 24;**94**(13):6670-5.

Lemmers C, Medina E, Delgrossi MH, Michel D, Arsanto JP, Le Bivic A. hINADl/PATJ, a homolog of discs lost, interacts with crumbs and localizes to tight junctions in human epithelial cells. *J Biol Chem.* (2002) Jul 12;**277**(28):25408-15.

Levine AJ. p53, the cellular gatekeeper for growth and division. *Cell.* (1997) Feb 7;**88**(3):323-31.

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

Li S, Labrecque S, Gauzzi MC, Cuddihy AR, Wong AH, Pellegrini S, Matlashewski GJ, Koromilas AE. The human papilloma virus (HPV)-18 E6 oncoprotein physically associates with Tyk2 and impairs Jak-STAT activation by interferon-alpha. *Oncogene.* (1999) Oct 14;**18**(42):5727-37.

Li HS, Montell C. TRP and the PDZ protein, INAD, form the core complex required for retention of the signalplex in Drosophila photoreceptor cells. *J Cell Biol.* (2000) Sep 18;**150**(6):1411-22.

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

Liu Y, Chen JJ, Gao Q, Dalal S, Hong Y, Mansur CP, Band V, Androphy EJ. Multiple functions of human papillomavirus type 16 E6 contribute to the immortalization of mammary epithelial cells. *J Virol.* (1999) Sep;**73**(9):7297-307.

Liu X, Clements A, Zhao K, Marmorstein R. Structure of the human Papillomavirus E7 oncoprotein and its mechanism for inactivation of the retinoblastoma tumor suppressor. *J Biol Chem.* (2006) Jan 6;**281**(1):578-86.

Liu Y, Liu Z, Androphy E, Chen J, Baleja JD. Design and characterization of helical peptides that inhibit the E6 protein of papillomavirus. *Biochemistry.* (2004) Jun 15;**43**(23):7421-31.

Longworth MS, Laimins LA. 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.* (2004a) Apr;**78**(7):3533-41.

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

Lorenzo HK, Farber D, Germain V, Acuto O, Alzari PM. The MBP fusion protein restores the activity of the first phosphatase domain of CD45. *FEBS Lett.* (1997) Jul 14;**411**(2-3):231-5.

Lu H, Levine AJ. Human TAFII31 protein is a transcriptional coactivator of the p53 protein. *Proc Natl Acad Sci U S A.* (1995) May 23;**92**(11):5154-8.

Lue RA, Marfatia SM, Branton D, Chishti AH. Cloning and characterization of hdlg: the human homologue of the Drosophila discs large tumor suppressor binds to protein 4.1. *Proc Natl Acad Sci U S A.* (1994) Oct 11;**91**(21):9818-22.

Lusky M, Hurwitz J, Seo YS. The bovine papillomavirus E2 protein modulates the assembly of but is not stably maintained in a replication-competent multimeric E1-replication origin complex. *Proc Natl Acad Sci U S A.* (1994) Sep 13;**91**(19):8895-9.

Macias MJ, Hyvonen M, Baraldi E, Schultz J, Sudol M, Saraste M, Oschkinat H. Structure of the WW domain of a kinase associated protein complexed with a proline-rich peptide. *Nature.* (1996) Aug 15;**382**(6592):646-9.

Mancini A, Koch A, Stefan M, Niemann H, Tamura T. The direct association of the multiple PDZ domain containing proteins (MUPP-1) with the human c-Kit C-terminus is regulated by tyrosine kinase activity. *FEBS Lett.* (2000) Sep 29;**482**(1-2):54-8.

References

Maniatis T, Fritsch EF, Sambrook J. Molecular Cloning - A laboratory manual. *Cold Spring Harbor Laboratory* Lit. **72** Maniatis et al. 1989

Mantovani F, Banks L. The human papillomavirus E6 protein and its contribution to malignant progression. *Oncogene.* (2001) Nov 26;**20**(54):7874-87.

Marfatia SM, Morais Cabral JH, Lin L, Hough C, Bryant PJ, Stolz L, Chishti AH. Modular organization of the PDZ domains in the human discs-large protein suggests a mechanism for coupling PDZ domain-binding proteins to ATP and the membrane cytoskeleton. *J Cell Biol.* (1996) Nov;**135**(3):753-66.

Marte BM, Downward J. PKB/Akt: connecting phosphoinositide 3-kinase to cell survival and beyond. *Trends Biochem Sci.* (1997) Sep;**22**(9):355-8.

Martin LG, Demers GW, Galloway DA. Disruption of the G1/S transition in human papillomavirus type 16 E7-expressing human cells is associated with altered regulation of cyclin E. *J Virol.* (1998) Feb;**72**(2):975-85.

Massimi P, Gammoh N, Thomas M, Banks L. HPV E6 specifically targets different cellular pools of its PDZ domain-containing tumour suppressor substrates for proteasome-mediated degradation. *Oncogene.* (2004) Oct 21;**23**(49):8033-9.

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

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

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

Matsuda S, Mikawa S, Hirai H. Phosphorylation of serine-880 in GluR2 by protein kinase C prevents its C terminus from binding with glutamate receptor-interacting protein. *J Neurochem.* (1999) Oct;**73**(4):1765-8.

Matsumine A, Ogai A, Senda T, Okumura N, Satoh K, Baeg GH, Kawahara T, Kobayashi S, Okada M, Toyoshima K, Akiyama T. Binding of APC to the human homolog of the Drosophila discs large tumor suppressor protein. *Science.* (1996) May 17;**272**(5264):1020-3.

Matsumoto Y, Nakagawa S, Yano T, Takizawa S, Nagasaka K, Nakagawa K, Minaguchi T, Wada O, Ooishi H, Matsumoto K, Yasugi T, Kanda T, Huibregtse JM, Taketani Y. Involvement of a cellular ubiquitin-protein ligase E6AP in the ubiquitin-mediated degradation of extensive substrates of high-risk human papillomavirus E6. *J Med Virol.* (2006) Apr;**78**(4):501-7.

May P, May E. Twenty years of p53 research: structural and functional aspects of the p53 protein. *Oncogene.* (1999) Dec 13;**18**(53):7621-36. **Erratum in:** *Oncogene* (2000) Mar 23;**19**(13):1734.

McIntyre MC, Frattini MG, Grossman SR, Laimins LA. Human papillomavirus type 18 E7 protein requires intact Cys-X-X Cys motifs for zinc binding, dimerization, and transformation but not for Rb binding. *J Virol.* (1993) Jun;**67**(6):3142-50.

McIntyre MC, Ruesch MN, Laimins LA. Human papillomavirus E7 oncoproteins bind a single form of cyclin E in a complex with cdk2 and p107. *Virology.* (1996) Jan 1;**215**(1):73-82.

Middleton K, Peh W, Southern S, Griffin H, Sotlar K, Nakahara T, El-Sherif A, Morris L, Seth R, Hibma M, Jenkins D, Lambert P, Coleman N, Doorbar J. Organization of human papillomavirus productive cycle during neoplastic progression provides a basis for selection of diagnostic markers. *J Virol.* (2003) Oct;**77**(19):10186-201.

Migaud M, Charlesworth P, Dempster M, Webster LC, Watabe AM, Makhinson M, He Y, Ramsay MF, Morris RG, Morrison JH, O'Dell TJ, Grant SG. Enhanced long-term potentiation and impaired learning in mice with mutant postsynaptic density-95 protein. *Nature.* (1998) Dec 3;**396**(6710):433-9.

Mino A, Ohtsuka T, Inoue E, Takai Y. Membrane-associated guanylate kinase with inverted orientation (MAGI)-1/brain angiogenesis inhibitor 1-associated protein (BAP1) as a scaffolding molecule for Rap small G protein GDP/GTP exchange protein at tight junctions. *Genes Cells.* (2000) Dec;**5**(12):1009-16.

Mitrani-Rosenbaum S, Tsvieli R, Tur-Kaspa R. Oestrogen stimulates differential transcription of human papillomavirus type 16 in SiHa cervical carcinoma cells. *J Gen Virol.* (1989) Aug;**70** (Pt 8):2227-32.

Mittal R, Tsutsumi K, Pater A, Pater MM. Human papillomavirus type 16 expression in cervical keratinocytes: role of progesterone and glucocorticoid hormones. *Obstet Gynecol.* (1993) Jan;**81**(1):5-12.

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

References

Momand J, Zambetti GP, Olson DC, George D, Levine AJ. The mdm-2 oncogene product forms a complex with the p53 protein and inhibits p53-mediated transactivation. *Cell.* (1992) Jun 26;**69**(7):1237-45.

Montcouquiol M, Rachel RA, Lanford PJ, Copeland NG, Jenkins NA, Kelley MW. Identification of Vangl2 and Scrb1 as planar polarity genes in mammals. *Nature.* (2003) May 8;**423**(6936):173-7.

Moreno-Lopez J, Ahola H, Stenlund A, Osterhaus A, Pettersson U. Genome of an avian papillomavirus. *J Virol.* (1984) Sep;**51**(3):872-5.

Morosov A, Phelps WC, Raychaudhuri P. Activation of the c-fos gene by the HPV16 oncoproteins depends upon the camp response element at -60. *J Biol Chem.* (1994) Jul 15;**269**(28):18434-40.

Muller H, Helin K. The E2F transcription factors: key regulators of cell proliferation. *Biochim Biophys Acta.* (2000) Feb 14;**1470**(1):M1-12.

Muller BM, Kistner U, Kindler S, Chung WJ, Kuhlendahl S, Fenster SD, Lau LF, Veh RW, Huganir RL, Gundelfinger ED, Garner CC. SAP102, a novel postsynaptic protein that interacts with NMDA receptor complexes in vivo. *Neuron.* (1996) Aug;**17**(2):255-65.

Muller BM, Kistner U, Veh RW, Cases-Langhoff C, Becker B, Gundelfinger ED, Garner CC. Molecular characterization and spatial distribution of SAP97, a novel presynaptic protein homologous to SAP90 and the Drosophila discs-large tumor suppressor protein. *J Neurosci.* (1995) Mar;**15**(3 Pt 2):2354-66.

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

Munger K, Phelps WC, Bubb V, Howley PM, Schlegel R. The E6 and E7 genes of the human papillomavirus type 16 together are necessary and sufficient for transformation of primary human keratinocytes. *J Virol.* (1989a) Oct;**63**(10):4417-21.

Munger K, Werness BA, Dyson N, Phelps WC, Harlow E, Howley PM. Complex formation of human papillomavirus E7 proteins with the retinoblastoma tumor suppressor gene product. *EMBO J.* (1989b) Dec 20;**8**(13):4099-105.

Munoz N. A review of the phase III clinical data for the quadrivalent human papillomavirus vaccine (Gardasil((R))). *Eur J Obstet Gynecol Reprod Biol.* (2006) Aug 18; [Epub ahead of print] PMID: 16920247

Munoz N. Efficacy of a quadrivalent HPV (types 6, 11, 16, 18) L1 VLP vaccine against cervical intraepithelial neoplasia grades 1-3 and adenocarcinoma in situ A combined analysis. *Eur J Obstet Gynecol Reprod Biol.* (2006) Aug 25; [Epub ahead of print] PMID: 16935411

Murdoch JN, Henderson DJ, Doudney K, Gaston-Massuet C, Phillips HM, Paternotte C, Arkell R, Stanier P, Copp AJ. Disruption of scribble (Scrb1) causes severe neural tube defects in the circletail mouse. *Hum Mol Genet.* (2003) Jan 15;**12**(2):87-98.

Naisbitt S, Kim E, Tu JC, Xiao B, Sala C, Valtschanoff J, Weinberg RJ, Worley PF, Sheng M. Shank, a novel family of postsynaptic density proteins that binds to the NMDA receptor/PSD-95/GKAP complex and cortactin. *Neuron.* (1999) Jul;**23**(3):569-82.

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

Nakagawa S, Yano T, Nakagawa K, Takizawa S, Suzuki Y, Yasugi T, Huibregtse JM, Taketani Y. Analysis of the expression and localisation of a LAP protein, human scribble, in the normal and neoplastic epithelium of uterine cervix. *Br J Cancer.* (2004) Jan 12;**90**(1):194-9.

Nakahara T, Peh WL, Doorbar J, Lee D, Lambert PF. Human papillomavirus type 16 E1circumflexE4 contributes to multiple facets of the papillomavirus life cycle. *J Virol.* (2005) Oct;**79**(20):13150-65.

Nanda K, McCrory DC, Myers ER, Bastian LA, Hasselblad V, Hickey JD, Matchar DB. Accuracy of the Papanicolaou test in screening for and follow-up of cervical cytologic abnormalities: a systematic review. *Ann Intern Med.* (2000) May 16;**132**(10):810-9.

Nawaz Z, Lonard DM, Smith CL, Lev-Lehman E, Tsai SY, Tsai MJ, O'Malley BW. The Angelman syndrome-associated protein, E6-AP, is a coactivator for the nuclear hormone receptor superfamily. *Mol Cell Biol.* (1999) Feb;**19**(2):1182-9.

Neary K, Horwitz BH, DiMaio D. Mutational analysis of open reading frame E4 of bovine papillomavirus type 1. *J Virol.* (1987) Apr;**61**(4):1248-52.

Neudauer CL, Joberty G, Macara IG. PIST: a novel PDZ/coiled-coil domain binding partner for the rho-family GTPase TC10. *Biochem Biophys Res Commun.* (2001) Jan 19;**280**(2):541-7.

References

Nguyen ML, Nguyen MM, Lee D, Griep AE, Lambert PF. The PDZ ligand domain of the human papillomavirus type 16 E6 protein is required for E6's induction of epithelial hyperplasia in vivo. *J Virol.* (2003) Jun;**77**(12):6957-64.

Niethammer M, Kim E, Sheng M. Interaction between the C terminus of NMDA receptor subunits and multiple members of the PSD-95 family of membrane-associated guanylate kinases. *J Neurosci.* (1996) Apr 1;**16**(7):2157-63.

Niethammer M, Valtschanoff JG, Kapoor TM, Allison DW, Weinberg TM, Craig AM, Sheng M. CRIPT, a novel postsynaptic protein that binds to the third PDZ domain of PSD-95/SAP90. *Neuron.* (1998) Apr;**20**(4):693-707.

Nishimune A, Isaac JT, Molnar E, Noel J, Nash SR, Tagaya M, Collingridge GL, Nakanishi S, Henley JM. NSF binding to GluR2 regulates synaptic transmission. *Neuron.* (1998) Jul;**21**(1):87-97.

Nishimura A, Ono T, Ishimoto A, Dowhanick JJ, Frizzell MA, Howley PM, Sakai H. Mechanisms of human papillomavirus E2-mediated repression of viral oncogene expression and cervical cancer cell growth inhibition. *J Virol.* (2000) Apr;**74**(8):3752-60.

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

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

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

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

Nomine Y, Ristriani T, Laurent C, Lefevre JF, Weiss E, Trave G. A strategy for optimizing the monodispersity of fusion proteins: application to purification of recombinant HPV E6 oncoprotein. *Protein Eng.* (2001b) Apr;**14**(4):297-305.

Nourry C, Grant SG, Borg JP. PDZ domain proteins: plug and play! *Sci STKE.* (2003) Apr 22;**2003**(179):RE7.

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

Ohlenschlager O, Seiboth T, Zengerling H, Briese L, Marchanka A, Ramachandran R, Baum M, Korbas M, Meyer-Klaucke W, Durst M, Gorlach M. Solution structure of the partially folded high-risk human papilloma virus 45 oncoprotein E7. *Oncogene.* (2006) Apr 24; [Epub ahead of print] PMID: 16636661

Okorokov AL, Ponchel F, Milner J. Induced N- and C-terminal cleavage of p53: a core fragment of p53, generated by interaction with damaged DNA, promotes cleavage of the N-terminus of full-length p53, whereas ssDNA induces C-terminal cleavage of p53. *EMBO J.* (1997) Oct 1;**16**(19):6008-17.

Oliner JD, Pietenpol JA, Thiagalingam S, Gyuris J, Kinzler KW, Vogelstein B. Oncoprotein MDM2 conceals the activation domain of tumour suppressor p53. *Nature.* (1993) Apr 29;**362**(6423):857-60.

Olalla L, Aledo JC, Bannenberg G, Marquez J. The C-terminus of human glutaminase L mediates association with PDZ domain-containing proteins. FEBS Lett. 2001 Jan 19;488(3):116-22. **Erratum in:** *FEBS Lett* (2002) Nov 20;**531**(3):570.

Oren M. Decision making by p53: life, death and cancer. *Cell Death Differ.* (2003) Apr;**10**(4):431-42.

PAPANICOLAOU GN. A survey of the actualities and potentialities of exfoliative cytology in cancer diagnosis. *Ann Intern Med.* (1949) Oct;**31**(4):661-74.

Palefsky J. Biology of HPV in HIV infection. *Adv Dent Res.* (2006) Apr 1;**19**(1):99-105.

Palefsky JM, Winkler B, Rabanus JP, Clark C, Chan S, Nizet V, Schoolnik GK. Characterization of in vivo expression of the human papillomavirus type 16 E4 protein in cervical biopsy tissues. *J Clin Invest.* (1991) Jun;**87**(6):2132-41.

Park P, Copeland W, Yang L, Wang T, Botchan MR, Mohr IJ. The cellular DNA polymerase alpha-primase is required for papillomavirus DNA replication and associates with the viral E1 helicase. *Proc Natl Acad Sci U S A.* (1994) Aug 30;**91**(18):8700-4.

Park J, Sun D, Genest DR, Trivijitsilp P, Suh I, Crum CP. Coexistence of low and high grade squamous intraepithelial lesions of the cervix: morphologic progression or multiple papillomaviruses? *Gynecol Oncol.* (1998) Sep;**70**(3):386-91.

References

Parkin DM, Pisani P, Ferlay J. Estimates of the worldwide incidence of 25 major cancers in 1990. *Int J Cancer.* (1999) Mar 15;80(6):827-41.

Parkin DM. The global health burden of infection-associated cancers in the year 2002. *Int J Cancer.* (2006) Jun 15;**118**(12):3030-44.

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

Parks TD, Leuther KK, Howard ED, Johnston SA, Dougherty WG. Release of proteins and peptides from fusion proteins using a recombinant plant virus proteinase. *Anal Biochem.* (1994) Feb 1;**216**(2):413-7.

Patel D, Huang SM, Baglia LA, McCance DJ. The E6 protein of human papillomavirus type 16 binds to and inhibits co-activation by CBP and p300. *EMBO J.* (1999) Sep 15;**18**(18):5061-72.

Patrie KM, Drescher AJ, Goyal M, Wiggins RC, Margolis B. The membrane-associated guanylate kinase protein MAGI-1 binds megalin and is present in glomerular podocytes. *J Am Soc Nephrol.* (2001) Apr;**12**(4):667-77.

Patrie KM, Drescher AJ, Welihinda A, Mundel P, Margolis B. Interaction of two actin-binding proteins, synaptopodin and alpha-actinin-4, with the tight junction protein MAGI-1. *J Biol Chem.* (2002) Aug 16;**277**(33):30183-90.

Pavletich NP, Chambers KA, Pabo CO. The DNA-binding domain of p53 contains the four conserved regions and the major mutation hot spots. *Genes Dev.* (1993) Dec;**7**(12B):2556-64.

Peh WL, Middleton K, Christensen N, Nicholls P, Egawa K, Sotlar K, Brandsma J, Percival A, Lewis J, Liu WJ, Doorbar J. Life cycle heterogeneity in animal models of human papillomavirus-associated disease. *J Virol.* (2002) Oct;**76**(20):10401-16.

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

Pietenpol JA, Tokino T, Thiagalingam S, el-Deiry WS, Kinzler KW, Vogelstein B. Sequence-specific transcriptional activation is essential for growth suppression by p53. *Proc Natl Acad Sci U S A.* (1994) Mar 15;**91**(6):1998-2002.

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

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

Pim D, Thomas M, Banks L. Chimaeric HPV E6 proteins allow dissection of the proteolytic pathways regulating different E6 cellular target proteins. *Oncogene.* (2002) Nov 21;**21**(53):8140-8.

Pim D, Thomas M, Javier R, Gardiol D, Banks L. HPV E6 targeted degradation of the discs large protein: evidence for the involvement of a novel ubiquitin ligase. *Oncogene.* (2000) Feb 10;**19**(6):719-25.

Piotto M, Saudek V, Sklenar V. Gradient-tailored excitation for single-quantum NMR spectroscopy of aqueous solutions. *J Biomol NMR.* (1992) Nov;**2**(6):661-5.

Pisani P, Parkin DM, Bray F, Ferlay J. Estimates of the worldwide mortality from 25 cancers in 1990. *Int J Cancer.* (1999) Sep 24;83(1):18-29. **Erratum in:** *Int J Cancer* (1999) Dec 10;**83**(6):870-3.

Pisani P, Bray F, Parkin DM. Estimates of the world-wide prevalence of cancer for 25 sites in the adult population. *Int J Cancer.* (2002) Jan 1;**97**(1):72-81.

Polakis P. The oncogenic activation of beta-catenin. *Curr Opin Genet Dev.* (1999) Feb;**9**(1):15-21.

Ponting CP. Evidence for PDZ domains in bacteria, yeast, and plants. *Protein Sci.* (1997) Feb;**6**(2):464-8.

Porreco R, Penn I, Droegemueller W, Greer B, Makowski E. Gynecologic malignancies in immunosuppressed organ homograft recipients. *Obstet Gynecol.* (1975) Apr;**45**(4):359-64.

Pray TR, Laimins LA. Differentiation-dependent expression of E1--E4 proteins in cell lines maintaining episomes of human papillomavirus type 31b. *Virology.* (1995) Jan 10;**206**(1):679-85.

Rabson MS, Yee C, Yang YC, Howley PM. Bovine papillomavirus type 1 3' early region transformation and plasmid maintenance functions. *J Virol.* (1986) Nov;**60**(2):626-34.

References

Raghuram V, Mak DD, Foskett JK. Regulation of cystic fibrosis transmembrane conductance regulator single-channel gating by bivalent PDZ-domain-mediated interaction. *Proc Natl Acad Sci U S A.* (2001) Jan 30;**98**(3):1300-5.

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

Reina J, Lacroix E, Hobson SD, Fernandez-Ballester G, Rybin V, Schwab MS, Serrano L, Gonzalez C. Computer-aided design of a PDZ domain to recognize new target sequences. *Nat Struct Biol.* (2002) Aug;**9**(8):621-7.

Reuver SM, Garner CC. E-cadherin mediated cell adhesion recruits SAP97 into the cortical cytoskeleton. *J Cell Sci.* (1998) Apr;**111** (Pt 8):1071-80.

Reynaud C, Fabre S, Jalinot P. The PDZ protein TIP-1 interacts with the Rho effector rhotekin and is involved in Rho signaling to the serum response element. *J Biol Chem.* (2000) Oct 27;**275**(43):33962-8.

Richards RM, Lowy DR, Schiller JT, Day PM. Cleavage of the papillomavirus minor capsid protein, L2, at a furin consensus site is necessary for infection. *Proc Natl Acad Sci U S A.* (2006) Jan 31;**103**(5):1522-7.

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

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

Roberts S, Ashmole I, Gibson LJ, Rookes SM, Barton GJ, Gallimore PH. Mutational analysis of human papillomavirus E4 proteins: identification of structural features important in the formation of cytoplasmic E4/cytokeratin networks in epithelial cells. *J Virol.* (1994a) Oct;**68**(10):6432-45.

Roberts S, Ashmole I, Johnson GD, Kreider JW, Gallimore PH. Cutaneous and mucosal human papillomavirus E4 proteins form intermediate filament-like structures in epithelial cells. *Virology.* (1993) Nov;**197**(1):176-87.

Roberts S, Ashmole I, Rookes SM, Gallimore PH. Mutational analysis of the human papillomavirus type 16 E1--E4 protein shows that the C terminus is dispensable for keratin cytoskeleton association but is involved in inducing disruption of the keratin filaments. *J Virol.* (1997) May;**71**(5):3554-62.

Roberts S, Ashmole I, Sheehan TM, Davies AH, Gallimore PH. Human papillomavirus type 1 E4 protein is a zinc-binding protein. *Virology.* (1994b) Aug 1;**202**(2):865-74.

Roche JK, Crum CP. Local immunity and the uterine cervix: implications for cancer-associated viruses. *Cancer Immunol Immunother.* (1991);**33**(4):203-9.

Roh MH, Makarova O, Liu CJ, Shin K, Lee S, Laurinec S, Goyal M, Wiggins R, Margolis B. The Maguk protein, Pals1, functions as an adapter, linking mammalian homologues of Crumbs and Discs Lost. *J Cell Biol.* (2002) Apr 1;**157**(1):161-72.

Roh MH, Margolis B. Composition and function of PDZ protein complexes during cell polarization. *Am J Physiol Renal Physiol.* (2003) Sep;**285**(3):F377-87.

Ronco LV, Karpova AY, Vidal M, Howley PM. Human papillomavirus 16 E6 oncoprotein binds to interferon regulatory factor-3 and inhibits its transcriptional activity. *Genes Dev.* (1998) Jul 1;**12**(13):2061-72.

Rosenthal AN, Ryan A, Al-Jehani RM, Storey A, Harwood CA, Jacobs IJ. p53 codon 72 polymorphism and risk of cervical cancer in UK. *Lancet.* (1998) Sep 12;**352**(9131):871-2.

Rousset R, Fabre S, Desbois C, Bantignies F, Jalinot P. The C-terminus of the HTLV-1 Tax oncoprotein mediates interaction with the PDZ domain of cellular proteins. *Oncogene.* (1998) Feb 5;**16**(5):643-54.

Ruesch MN, Laimins LA. Human papillomavirus oncoproteins alter differentiation-dependent cell cycle exit on suspension in semisolid medium. *Virology.* (1998) Oct 10;**250**(1):19-29.

Ruff P, Speicher DW, Husain-Chishti A. Molecular identification of a major palmitoylated erythrocyte membrane protein containing the src homology 3 motif. *Proc Natl Acad Sci U S A.* (1991) Aug 1;**88**(15):6595-9.

Ruppert JM, Stillman B. Analysis of a protein-binding domain of p53. *Mol Cell Biol.* (1993) Jun;**13**(6):3811-20.

Sakaguchi K, Herrera JE, Saito S, Miki T, Bustin M, Vassilev A, Anderson CW, Appella E. DNA damage activates p53 through a phosphorylation-acetylation cascade. *Genes Dev.* (1998) Sep 15;**12**(18):2831-41.

Sakamuro D, Sabbatini P, White E, Prendergast GC. The polyproline region of p53 is required to activate apoptosis but not growth arrest. *Oncogene.* (1997) Aug 18;**15**(8):887-98.

References

Sanchez-Perez AM, Soriano S, Clarke AR, Gaston K. Disruption of the human papillomavirus type 16 E2 gene protects cervica carcinoma cells from E2F-induced apoptosis. *J Gen Virol.* (1997) Nov;**78** (Pt 11):3009-18.

Sasson IM, Haley NJ, Hoffmann D, Wynder EL, Hellberg D, Nilsson S. Cigarette smoking and neoplasia of the uterine cervix: smoke constituents in cervical mucus. *N Engl J Med.* (1985) Jan 31;**312**(5):315-6.

Sato T, Irie S, Kitada S, Reed JC. FAP-1: a protein tyrosine phosphatase that associates with Fas. *Science.* (1995) Apr 21;**268**(5209):411-5.

Schagger H, von Jagow G. Tricine-sodium dodecyl sulfate-polyacrylamide gel electrophoresis for the separation of proteins in the range from 1 to 100 kDa. *Anal Biochem.* (1987) Nov 1;**166**(2):368-79.

Scheffner M, Huibregtse JM, Vierstra RD, Howley PM. The HPV-16 E6 and E6-AP complex functions as a ubiquitin-protein ligase in the ubiquitination of p53. *Cell.* (1993) Nov 5;**75**(3):495-505.

Scheffner M, Romanczuk H, Munger K, Huibregtse JM, Mietz JA, Howley PM. Functions of human papillomavirus proteins. *Curr Top Microbiol Immunol.* (1994);**186**:83-99.

Scheffner M, Werness BA, Huibregtse JM, Levine AJ, Howley PM. The E6 oncoprotein encoded by human papillomavirus types 16 and 18 promotes the degradation of p53. *Cell.* (1990) Dec 21;**63**(6):1129-36.

Schiffman M, Castle PE. Human papillomavirus: epidemiology and public health. *Arch Pathol Lab Med.* (2003) Aug;**127**(8):930-4.

Schiffman M, Castle PE. The promise of global cervical-cancer prevention. *N Engl J Med.* (2005) Nov 17;**353**(20):2101-4.

Schiller JT, Vass WC, Vousden KH, Lowy DR. E5 open reading frame of bovine papillomavirus type 1 encodes a transforming gene. *J Virol.* (1986) Jan;**57**(1):1-6.

Schneider-Gadicke A, Kaul S, Schwarz E, Gausepohl H, Frank R, Bastert G. Identification of the human papillomavirus type 18 E6 and E6 proteins in nuclear protein fractions from human cervical carcinoma cells grown in the nude mouse or in vitro. *Cancer Res.* (1988) Jun 1;**48**(11):2969-74.

Schrager LK, Friedland GH, Maude D, Schreiber K, Adachi A, Pizzuti DJ, Koss LG, Klein RS. Cervical and vaginal squamous cell abnormalities in women infected with human immunodeficiency virus. *J Acquir Immune Defic Syndr.* (1989);**2**(6):570-5.

Schultz J, Copley RR, Doerks T, Ponting CP, Bork P. SMART: a web-based tool for the study of genetically mobile domains. Nucleic Acids Res. 2000 Jan 1;28(1):231-4.

Schultz J, Milpetz F, Bork P, Ponting CP. SMART, a simple modular architecture research tool: identification of signalling domains. *Proc Natl Acad Sci U S A.* (1998) May 26;**95**(11):5857-64.

Schwalbach G, Sibler AP, Choulier L, Deryckere F, Weiss E. Production of fluorescent single-chain antibody fragments in Escherichia coli. *Protein Expr Purif.* (2000) Mar;**18**(2):121-32.

Schwarz E, Freese UK, Gissmann L, Mayer W, Roggenbuck B, Stremlau A, zur Hausen H. Structure and transcription of human papillomavirus sequences in cervical carcinoma cells. *Nature.* (1985) Mar 7-13;**314**(6006):111-4.

Sedman SA, Barbosa MS, Vass WC, Hubbert NL, Haas JA, Lowy DR, Schiller JT. The full-length E6 protein of human papillomavirus type 16 has transforming and trans-activating activities and cooperates with E7 to immortalize keratinocytes in culture. *J Virol.* (1991) Sep;**65**(9):4860-6.

Selivanova G, Iotsova V, Okan I, Fritsche M, Strom M, Groner B, Grafstrom RC, Wiman KG. Restoration of the growth suppression function of mutant p53 by a synthetic peptide derived from the p53 C-terminal domain. *Nat Med.* (1997) Jun;**3**(6):632-8.

Selivanova G, Ryabchenko L, Jansson E, Iotsova V, Wiman KG. Reactivation of mutant p53 through interaction of a C terminal peptide with the core domain. *Mol Cell Biol.* (1999) May;**19**(5):3395-402.

Shaulsky G, Goldfinger N, Ben-Ze'ev A, Rotter V. Nuclear accumulation of p53 protein is mediated by several nuclear localization signals and plays a role in tumorigenesis. *Mol Cell Biol.* (1990) Dec;**10**(12):6565-77.

Shaw P, Freeman J, Bovey R, Iggo R. Regulation of specific DNA binding by p53: evidence for a role for O-glycosylation and charged residues at the carboxy-terminus. *Oncogene.* (1996) Feb 15;**12**(4):921-30.

Sheng M, Kim E. The Shank family of scaffold proteins. *J Cell Sci.* (2000) Jun;**113** (Pt 11):1851-6.

Sheng M, Sala C. PDZ domains and the organization of supramolecular complexes. *Annu Rev Neurosci.* (2001);**24**:1-29.

References

Sherman L, Alloul N. Human papillomavirus type 16 expresses a variety of alternatively spliced mRNAs putatively encoding the E2 protein. *Virology.* (1992) Dec;**191**(2):953-9.

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

Sherman L, Schlegel R. Serum- and calcium-induced differentiation of human keratinocytes is inhibited by the E6 oncoprotein of human papillomavirus type 16. *J Virol.* (1996) May;**70**(5):3269-79.

Sherr CJ. The Pezcoller lecture: cancer cell cycles revisited. *Cancer Res.* (2000) Jul 15;**60**(14):3689-95. Review.

Shieh BH, Zhu MY. Regulation of the TRP Ca2+ channel by INAD in Drosophila photoreceptors. *Neuron.* (1996) May;**16**(5):991-8.

Shih YP, Kung WM, Chen JC, Yeh CH, Wang AH, Wang TF. High-throughput screening of soluble recombinant proteins. *Protein Sci.* (2002) Jul;**11**(7):1714-9.

Shin K, Fogg VC, Margolis B. Tight Junctions and Cell Polarity. *Annu Rev Cell Dev Biol.* (2006) Jun 13; [Epub ahead of print] PMID: 16771626

Shin K, Straight S, Margolis B. PATJ regulates tight junction formation and polarity in mammalian epithelial cells. *J Cell Biol.* (2005) Feb 28;**168**(5):705-11.

Shirasawa H, Jin MH, Shimizu K, Akutsu N, Shino Y, Simizu B. Transcription-modulatory activity of full-length E6 and E6*I proteins of human papillomavirus type 16. *Virology.* (1994) Aug 15;**203**(1):36-42.

Shoji H, Tsuchida K, Kishi H, Yamakawa N, Matsuzaki T, Liu Z, Nakamura T, Sugino H. Identification and characterization of a PDZ protein that interacts with activin type II receptors. *J Biol Chem.* (2000) Feb 25;**275**(8):5485-92.

Simonson SJ, Difilippantonio MJ, Lambert PF. Two distinct activities contribute to human papillomavirus 16 E6's oncogenic potential. *Cancer Res.* (2005) Sep 15;**65**(18):8266-73.

Sinal SH, Woods CR. Human papillomavirus infections of the genital and respiratory tracts in young children. *Semin Pediatr Infect Dis.* (2005) Oct;**16**(4):306-16.

Singh M, Krajewski M, Mikolajka A, Holak TA. Molecular determinants for the complex formation between the retinoblastoma protein and LXCXE sequences. *J Biol Chem.* (2005) Nov 11;**280**(45):37868-76.

Skelton NJ, Koehler MF, Zobel K, Wong WL, Yeh S, Pisabarro MT, Yin JP, Lasky LA, Sidhu SS. Origins of PDZ domain ligand specificity. Structure determination and mutagenesis of the Erbin PDZ domain. *J Biol Chem.* (2003) Feb 28;**278**(9):7645-54.

Slee EA, O'Connor DJ, Lu X. To die or not to die: how does p53 decide? *Oncogene.* (2004) Apr 12;**23**(16):2809-18.

Smith A, Ashworth A. Cancer predisposition: where's the phosphate? *Curr Biol.* (1998) Mar 26;**8**(7):R241-3.

Smotkin D, Prokoph H, Wettstein FO. Oncogenic and nononcogenic human genital papillomaviruses generate the E7 mRNA by different mechanisms. *J Virol.* (1989) Mar;**63**(3):1441-7.

Song S, Pitot HC, Lambert PF. The human papillomavirus type 16 E6 gene alone is sufficient to induce carcinomas in transgenic animals. *J Virol.* (1999) Jul;**73**(7):5887-93.

Songyang Z, Fanning AS, Fu C, Xu J, Marfatia SM, Chishti AH, Crompton A, Chan AC, Anderson JM, Cantley LC. Recognition of unique carboxyl-terminal motifs by distinct PDZ domains. *Science.* (1997) Jan 3;**275**(5296):73-7

Soussi T, Caron de Fromentel C, May P. Structural aspects of the p53 protein in relation to gene evolution. *Oncogene.* (1990) Jul;**5**(7):945-52.

Soussi T, May P. Structural aspects of the p53 protein in relation to gene evolution: a second look. *J Mol Biol.* (1996) Aug 2;**260**(5):623-37.

Srivastava S, Osten P, Vilim FS, Khatri L, Inman G, States B, Daly C, DeSouza S, Abagyan R, Valtschanoff JG, Weinberg RJ, Ziff EB. Novel anchorage of GluR2/3 to the postsynaptic density by the AMPA receptor-binding protein ABP. *Neuron.* (1998) Sep;**21**(3):581-91.

Stacey SN, Jordan D, Williamson AJ, Brown M, Coote JH, Arrand JR. Leaky scanning is the predominant mechanism for translation of human papillomavirus type 16 E7 oncoprotein from E6/E7 bicistronic mRNA. *J Virol.* (2000) Aug;**74**(16):7284-97.

References

Steenbergen RD, Parker JN, Isern S, Snijders PJ, Walboomers JM, Meijer CJ, Broker TR, Chow LT. Viral E6-E7 transcription in the basal layer of organotypic cultures without apparent p21cip1 protein precedes immortalization of human papillomavirus type 16- and 18-transfected human keratinocytes. *J Virol.* (1998) Jan;**72**(1):749-57.

Steger G, Corbach S. Dose-dependent regulation of the early promoter of human papillomavirus type 18 by the viral E2 protein. *J Virol.* (1997) Jan;**71**(1):50-8.

Sterling JC, Skepper JN, Stanley MA. Immunoelectron microscopical localization of human papillomavirus type 16 L1 and E4 proteins in cervical keratinocytes cultured in vivo. *J Invest Dermatol.* (1993) Feb;**100**(2):154-8.

Sterlinko Grm H, Banks L. HPV proteins as targets for therapeutic intervention. *Antivir Ther.* (2004) Oct;**9**(5):665-78.

Stoler MH, Rhodes CR, Whitbeck A, Wolinsky SM, Chow LT, Broker TR. Human papillomavirus type 16 and 18 gene expression in cervical neoplasias. *Hum Pathol.* (1992) Feb;**23**(2):117-28.

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

Sudol M. Structure and function of the WW domain. *Prog Biophys Mol Biol.* (1996);**65**(1-2):113-32.

Suzuki T, Ohsugi Y, Uchida-Toita M, Akiyama T, Yoshida M. Tax oncoprotein of HTLV-1 binds to the human homologue of Drosophila discs large tumor suppressor protein, hDLG, and perturbs its function in cell growth control. *Oncogene.* (1999) Oct 28;**18**(44):5967-72.

Takeichi M. Morphogenetic roles of classic cadherins. *Curr Opin Cell Biol.* (1995) Oct;**7**(5):619-27.

Takizawa S, Nagasaka K, Nakagawa S, Yano T, Nakagawa K, Yasugi T, Takeuchi T, Kanda T, Huibregtse JM, Akiyama T, Taketani Y. Human scribble, a novel tumor suppressor identified as a target of high-risk HPV E6 for ubiquitin-mediated degradation, interacts with adenomatous polyposis coli. *Genes Cells.* (2006) Apr;**11**(4):453-64.

Tan SH, Gloss B, Bernard HU. During negative regulation of the human papillomavirus-16 E6 promoter, the viral E2 protein can displace Sp1 from a proximal promoter element. *Nucleic Acids Res.* (1992) Jan 25;**20**(2):251-6.

Tan SH, Leong LE, Walker PA, Bernard HU. The human papillomavirus type 16 E2 transcription factor binds with low cooperativity to two flanking sites and represses the E6 promoter through displacement of Sp1 and TFIID. *J Virol.* (1994) Oct;**68**(10):6411-20.

The Human Papillomavirus Compendium 1995 (Online): www.stdgen.lanl.gov/stdgen/virus/hpv/compendium/htdocs/

Thierry F, Benotmane MA, Demeret C, Mori M, Teissier S, Desaintes C. A genomic approach reveals a novel mitotic pathway in papillomavirus carcinogenesis. *Cancer Res.* (2004) Feb 1;**64**(3):895-903.

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

Thomas MC, Chiang CM. E6 oncoprotein represses p53-dependent gene activation via inhibition of protein acetylation independently of inducing p53 degradation. *Mol Cell.* (2005) Jan 21;17(2):251-64.

Thomas M, Glaunsinger B, Pim D, Javier R, Banks L. HPV E6 and MAGUK protein interactions: determination of the molecular basis for specific protein recognition and degradation. *Oncogene.* (2001) Sep 6;**20**(39):5431-9.

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

Thomas M, Matlashewski G, Pim D, Banks L. Induction of apoptosis by p53 is independent of its oligomeric state and can be abolished by HPV-18 E6 through ubiquitin mediated degradation. *Oncogene.* (1996) Jul 18;**13**(2):265-73.

Thut CJ, Chen JL, Klemm R, Tjian R. p53 transcriptional activation mediated by coactivators TAFII40 and TAFII60. *Science.* (1995) Jan 6;**267**(5194):100-4.

Titolo S, Pelletier A, Sauve F, Brault K, Wardrop E, White PW, Amin A, Cordingley MG, Archambault J. Role of the ATP-binding domain of the human papillomavirus type 11 E1 helicase in E2-dependent binding to the origin. *J Virol.* (1999) Jul;**73**(7):5282-93.

Tochio H, Mok YK, Zhang Q, Kan HM, Bredt DS, Zhang M. Formation of nNOS/PSD-95 PDZ dimer requires a preformed beta finger structure from the nNOS PDZ domain. *J Mol Biol.* (2000) Oct 27;**303**(3):359-70.

References

Tommasino M, Adamczewski JP, Carlotti F, Barth CF, Manetti R, Contorni M, Cavalieri F, Hunt T, Crawford L. HPV16 E7 protein associates with the protein kinase p33CDK2 and cyclin A. *Oncogene.* (1993) Jan;**8**(1):195-202.

Tong X, Boll W, Kirchhausen T, Howley PM. Interaction of the bovine papillomavirus E6 protein with the clathrin adaptor complex AP-1. *J Virol.* (1998) Jan;**72**(1):476-82.

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

Trus BL, Roden RB, Greenstone HL, Vrhel M, Schiller JT, Booy FP. Novel structural features of bovine papillomavirus capsid revealed by a three-dimensional reconstruction to 9 A resolution. *Nat Struct Biol.* (1997) May;**4**(5):413-20.

Tsai TC, Chen SL. The biochemical and biological functions of human papillomavirus type 16 E5 protein. *Arch Virol.* (2003) Aug;**148**(8):1445-53.

Tsunoda S, Sierralta J, Sun Y, Bodner R, Suzuki E, Becker A, Socolich M, Zuker CS. A multivalent PDZ-domain protein assembles signalling complexes in a G-protein-coupled cascade. *Nature.* (1997) Jul 17;**388**(6639):243-9.

Tsunoda S, Sun Y, Suzuki E, Zuker C. Independent anchoring and assembly mechanisms of INAD signaling complexes in Drosophila photoreceptors. *J Neurosci.* (2001) Jan 1;**21**(1):150-8.

Tu JC, Xiao B, Naisbitt S, Yuan JP, Petralia RS, Brakeman P, Doan A, Aakalu VK, Lanahan AA, Sheng M, Worley PF. Coupling of mGluR/Homer and PSD-95 complexes by the Shank family of postsynaptic density proteins. *Neuron.* (1999) Jul;**23**(3):583-92.

Ullmer C, Schmuck K, Figge A, Lubbert H. Cloning and characterization of MUPP1, a novel PDZ domain protein. *FEBS Lett.* (1998) Mar 6;**424**(1-2):63-8.

Unger T, Nau MM, Segal S, Minna JD. p53: a transdominant regulator of transcription whose function is ablated by mutations occurring in human cancer. *EMBO J.* (1992) Apr;**11**(4):1383-90.

Ushikai M, Lace MJ, Yamakawa Y, Kono M, Anson J, Ishiji T, Parkkinen S, Wicker N, Valentine ME, Davidson I, et al. trans activation by the full-length E2 proteins of human papillomavirus type 16 and bovine papillomavirus type 1 in vitro and in vivo: cooperation with activation domains of cellular transcription factors. *J Virol.* (1994) Oct;**68**(10):6655-66.

Vaccaro P, Dente L. PDZ domains: troubles in classification. *FEBS Lett.* (2002) Feb 13;**512**(1-3):345-9.

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

Van Epps HL. Peyton Rous: father of the tumor virus. *J Exp Med.* (2005) Feb 7;**201**(3):320.

van Ham M, Hendriks W. PDZ domains-glue and guide. *Mol Biol Rep.* (2003) Jun;**30**(2):69-82.

Vande Pol SB, Brown MC, Turner CE. Association of Bovine Papillomavirus Type 1 E6 oncoprotein with the focal adhesion protein paxillin through a conserved protein interaction motif. *Oncogene.* (1998) Jan 8;**16**(1):43-52.

Varshavsky A. The N-end rule and regulation of apoptosis. *Nat Cell Biol.* (2003) May;**5**(5):373-6.

Veldman T, Liu X, Yuan H, Schlegel R. 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.* (2003) Jul 8;**100**(14):8211-6.

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

Walker KK, Levine AJ. Identification of a novel p53 functional domain that is necessary for efficient growth suppression. *Proc Natl Acad Sci U S A.* (1996) Dec 24;**93**(26):15335-40.

Walma T, Spronk CA, Tessari M, Aelen J, Schepens J, Hendriks W, Vuister GW. Structure, dynamics and binding characteristics of the second PDZ domain of PTP-BL. *J Mol Biol.* (2002) Mar 8;**316**(5):1101-10.

Wang NM, Chang JG. Are aberrant transcripts of FHIT, TSG101, and PTEN/MMAC1 oncogenesis related? *Int J Mol Med.* (1999) May;**3**(5):491-5.

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

References

Wang Y, Prives C. Increased and altered DNA binding of human p53 by S and G2/M but not G1 cyclin-dependent kinases. *Nature.* (1995) Jul 6;**376**(6535):88-91.

Wang XW, Vermeulen W, Coursen JD, Gibson M, Lupold SE, Forrester K, Xu G, Elmore L, Yeh H, Hoeijmakers JH, Harris CC. The XPB and XPD DNA helicases are components of the p53-mediated apoptosis pathway. *Genes Dev.* (1996) May 15;**10**(10):1219-32.

Wang S, Yue H, Derin RB, Guggino WB, Li M. Accessory protein facilitated CFTR-CFTR interaction, a molecular mechanism to potentiate the chloride channel activity. *Cell.* (2000) Sep 29;**103**(1):169-79.

Watson RA, Thomas M, Banks L, Roberts S. Activity of the human papillomavirus E6 PDZ-binding motif correlates with an enhanced morphological transformation of immortalized human keratinocytes. *J Cell Sci.* (2003) Dec 15;**116**(Pt 24):4925-34.

Waugh DS. Making the most of affinity tags. *Trends Biotechnol.* (2005) Jun;**23**(6):316-20.

Wazer DE, Liu XL, Chu Q, Gao Q, Band V. Immortalization of distinct human mammary epithelial cell types by human papilloma virus 16 E6 or E7. *Proc Natl Acad Sci U S A.* (1995) Apr 25;**92**(9):3687-91.

Webster K, Parish J, Pandya M, Stern PL, Clarke AR, Gaston K. The human papillomavirus (HPV) 16 E2 protein induces apoptosis in the absence of other HPV proteins and via a p53-dependent pathway. *J Biol Chem.* (2000) Jan 7;**275**(1):87-94.

Weinman EJ, Steplock D, Donowitz M, Shenolikar S. NHERF associations with sodium-hydrogen exchanger isoform 3 (NHE3) and ezrin are essential for cAMP-mediated phosphorylation and inhibition of NHE3. *Biochemistry.* (2000) May 23;**39**(20):6123-9.

Weinman EJ, Steplock D, Wang Y, Shenolikar S. Characterization of a protein cofactor that mediates protein kinase A regulation of the renal brush border membrane Na(+)-H+ exchanger. *J Clin Invest.* (1995) May;**95**(5):2143-9.

Weis K, Griffiths G, Lamond AI. The endoplasmic reticulum calcium-binding protein of 55 kDa is a novel EF-hand protein retained in the endoplasmic reticulum by a carboxyl-terminal His-Asp-Glu-Leu motif. *J Biol Chem.* (1994) Jul 22;**269**(29):19142-50.

Wells SI, Francis DA, Karpova AY, Dowhanick JJ, Benson JD, Howley PM. Papillomavirus E2 induces senescence in HPV-positive cells via pRB- and p21(CIP)-dependent pathways. *EMBO J.* (2000) Nov 1;**19**(21):5762-71.

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

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

Widmann M, Christen P. Comparison of folding rates of homologous prokaryotic and eukaryotic proteins. *J Biol Chem.* (2000) Jun 23;**275**(25):18619-22.

Wiedemann U, Boisguerin P, Leben R, Leitner D, Krause G, Moelling K, Volkmer-Engert R, Oschkinat H. Quantification of PDZ domain specificity, prediction of ligand affinity and rational design of super-binding peptides. *J Mol Biol.* (2004) Oct 22;**343**(3):703-18.

Willott E, Balda MS, Fanning AS, Jameson B, Van Itallie C, Anderson JM. The tight junction protein ZO-1 is homologous to the Drosophila discs-large tumor suppressor protein of septate junctions. *Proc Natl Acad Sci U S A.* (1993) Aug 15;**90**(16):7834-8.

Wilson R, Ryan GB, Knight GL, Laimins LA, Roberts S. The full-length E1;E4 protein of human papillomavirus type 18 modulates differentiation-dependent viral DNA amplification and late gene expression. *Virology.* (2007) Feb 13; [Epub ahead of print] PMID: 17303206

Wodarz A. Tumor suppressors: linking cell polarity and growth control. *Curr Biol.* (2000) Sep 7;**10**(17):R624-6.

Wolf G, Trub T, Ottinger E, Groninga L, Lynch A, White MF, Miyazaki M, Lee J, Shoelson SE. PTB domains of IRS-1 and Shc have distinct but overlapping binding specificities. *J Biol Chem.* (1995) Nov 17;**270**(46):27407-10.

Wood JD, Yuan J, Margolis RL, Colomer V, Duan K, Kushi J, Kaminsky Z, Kleiderlein JJ, Sharp AH, Ross CA. Atrophin-1, the DRPLA gene product, interacts with two families of WW domain-containing proteins. *Mol Cell Neurosci.* (1998) Jun;**11**(3):149-60.

Woods DF, Bryant PJ. The discs-large tumor suppressor gene of Drosophila encodes a guanylate kinase homolog localized at septate junctions. *Cell.* (1991) Aug 9;**66**(3):451-64.

Woods DF, Bryant PJ. ZO-1, DlgA and PSD-95/SAP90: homologous proteins in tight, septate and synaptic cell junctions. *Mech Dev.* (1993) Dec;**44**(2-3):85-9.

References

Woods DF, Hough C, Peel D, Callaini G, Bryant PJ. Dlg protein is required for junction structure, cell polarity, and proliferation control in Drosophila epithelia. *J Cell Biol.* (1996) Sep;**134**(6):1469-82.

Wu X, Bayle JH, Olson D, Levine AJ. The p53-mdm-2 autoregulatory feedback loop. *Genes Dev.* (1993) Jul;**7**(7A):1126-32

Wu X, Hepner K, Castelino-Prabhu S, Do D, Kaye MB, Yuan XJ, Wood J, Ross C, Sawyers CL, Whang YE. Evidence for regulation of the PTEN tumor suppressor by a membrane-localized multi-PDZ domain containing scaffold protein MAGI-2. *Proc Natl Acad Sci U S A.* (2000a) Apr 11;**97**(8):4233-8.

Wu YS, Domaille PJ. 1H, 15N and 13C assignments of the catalytic domain of E6-associated protein (E6AP). *J Biomol NMR.* (2001) Nov;**21**(3):285-6.

Wu Y, Dowbenko D, Spencer S, Laura R, Lee J, Gu Q, Lasky LA. Interaction of the tumor suppressor PTEN/MMAC with a PDZ domain of MAGI3, a novel membrane-associated guanylate kinase. *J Biol Chem.* (2000b) Jul 14;**275**(28):21477-85.

Wyszynski M, Valtschanoff JG, Naisbitt S, Dunah AW, Kim E, Standaert DG, Weinberg R, Sheng M. Association of AMPA receptors with a subset of glutamate receptor-interacting protein in vivo. *J Neurosci.* (1999) Aug 1;**19**(15):6528-37.

Yamada KM, Geiger B. Molecular interactions in cell adhesion complexes. *Curr Opin Cell Biol.* (1997) Feb;**9**(1):76-85.

Yamamoto Y, Huibregtse JM, Howley PM. The human E6-AP gene (UBE3A) encodes three potential protein isoforms generated by differential splicing. *Genomics.* (1997) Apr 15;**41**(2):263-6.

Yap AS, Brieher WM, Gumbiner BM. Molecular and functional analysis of cadherin-based adherens junctions. *Annu Rev Cell Dev Biol.* (1997);**13**:119-46.

Yew PR, Berk AJ. Inhibition of p53 transactivation required for transformation by adenovirus early 1B protein. *Nature.* (1992) May 7;**357**(6373):82-5.

Yuan F, Auborn K, James C. Altered growth and viral gene expression in human papillomavirus type 16-containing cancer cell lines treated with progesterone. *Cancer Invest.* (1999);**17**(1):19-29.

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

Zarrinpar A, Lim WA. Converging on proline: the mechanism of WW domain peptide recognition. *Nat Struct Biol.* (2000) Aug;**7**(8):611-3.

Zerfass-Thome K, Zwerschke W, Mannhardt B, Tindle R, Botz JW, Jansen-Durr P. Inactivation of the cdk inhibitor p27KIP1 by the human papillomavirus type 16 E7 oncoprotein. *Oncogene.* (1996) Dec 5;**13**(11):2323-30.

Zhang J, Rose BR, Thompson CH, Jarrett C, Russell P, Houghton RS, Cossart YE. Associations between oncogenic human papillomaviruses and local invasive patterns in cervical cancer. *Gynecol Oncol.* (1995) May;**57**(2):170-7.

Zhou MM, Huang B, Olejniczak ET, Meadows RP, Shuker SB, Miyazaki M, Trub T, Shoelson SE, Fesik SW. Structural basis for IL-4 receptor phosphopeptide recognition by the IRS-1 PTB domain. *Nat Struct Biol.* (1996) Apr;**3**(4):388-93.

Zhou MM, Ravichandran KS, Olejniczak EF, Petros AM, Meadows RP, Sattler M, Harlan JE, Wade WS, Burakoff SJ, Fesik SW. Structure and ligand recognition of the phosphotyrosine binding domain of Shc. *Nature.* (1995) Dec 7;**378**(6557):584-92.

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

zur Hausen H. Papillomavirus infections--a major cause of human cancers. *Biochim Biophys Acta.* (1996) Oct 9;**1288**(2):F55-78.

zur Hausen H. Papillomaviruses causing cancer: evasion from host-cell control in early events in carcinogenesis. *J Natl Cancer Inst.* (2000) May 3;**92**(9):690-8.

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

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

SUPPLEMENTARY DATA

Publication 7:
Zanier K, Nomine Y, Charbonnier S, Ruhlmann C, Schultz P, Schweizer J, Trave G.
Formation of well-defined soluble aggregates upon fusion to MBP is a generic property of E6 proteins from various human papillomavirus species. *Protein Expr Purif.* (2007) Jan;**51**(1):59-70.

Publication 8:
Nomine Y, Masson M, Charbonnier S, Zanier K, Ristriani T, Deryckere F, Sibler AP, Desplancq D, Atkinson RA, Weiss E, Orfanoudakis G, Kieffer B, Trave G. Structural and functional analysis of E6 oncoprotein: insights in the molecular pathways of human papillomavirus-mediated pathogenesis. *Mol Cell.* (2006) Mar 3;**21**(5):665-78.

Publication 9:
Lagrange M, Charbonnier S, Orfanoudakis G, Robinson P, Zanier K, Masson M, Lutz Y, Trave G, Weiss E, Deryckere F. Binding of human papillomavirus 16 E6 to p53 and E6AP is impaired by monoclonal antibodies directed against the second zinc-binding domain of E6. *J Gen Virol.* (2005) Apr;**86**(Pt 4):1001-7.

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

Katia Zanier, Yves Nominé, **Sebastian Charbonnier,** Christine Ruhlmann, Patrick Schultz, Johannes Schweizer and Gilles Travé

Structural and Functional Analysis of E6 Oncoprotein: Insights in the Molecular Pathways of Human Papillomavirus-Mediated Pathogenesis

Yves Nominé, Murielle Masson, **Sebastian Charbonnier,** Katia Zanier, Tutik Ristriani, François Deryckère, Annie-Paule Sibler, Dominique Desplancq, Robert Andrew Atkinson, Etienne Weiss, Georges Orfanoudakis, Bruno Kieffer et Gilles Travé

Molecular Cell, 2006, Vol. 21, Pages 665-678

Pages 665 à 678 :

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

Pour les utilisateurs ULP, il est possible de consulter cette publication sur le site de l'éditeur :
http://dx.doi.org/10.1016/j.molcel.2006.01.024

Il est également possible de consulter la thèse sous sa forme papier ou d'en faire une demande via le service de prêt entre bibliothèques (PEB), auprès du Service Commun de Documentation de l'ULP: peb.sciences@scd-ulp.u-strasbg.fr

Binding of human papillomavirus 16 E6 to p53 and E6AP is impaired by monoclonal antibodies directed against the second zinc-binding domain of E6

Magali Lagrange, **Sebastian Charbonnier**, Georges Orfanoudakis, Philip Robinson, Katia Zanier, Murielle Masson, Yves Lutz, Gilles Trave, Etienne Weiss et François Deryckere

Titre: Étude structurale et cinétique des interactions entre l'oncoprotéine E6 du virus du papillome humain et les domaines PDZ.

RÉSUMÉ : Les virus du papillome humain (HPV) infectent certaines cellules épithéliales et causent une multitude de lésions, allant de la verrue aux néoplasies, et pouvant conduire à la formation d'un cancer. Deux protéines virales, E6 et E7, participent en synergie à de nombreux processus cellulaires, pouvant conduire à l'immortalisation, voir la transformation des cellules. Dans ce contexte, les fonctions les plus citées sont la dégradation de l'antioncogène p53 et de la protéine pRb par respectivement E6 et E7. L'oncoprotéine E6 semble plutôt être impliquée dans les étapes tardives de progression maligne vers un cancer invasif que dans les étapes de promotion. Ceci semble corréler avec l'activité de dégradation d'une famille de protéines contenant des domaines PDZ par E6. Ces protéines PDZ sont souvent localisées au niveau de jonctions cellulaires et participent à la formation de complexes multiprotéiques de jonction, de signalisation ou de transport membranaire. Une dérégulation de ces protéines induit souvent une perte de polarité apico-basale et d'inhibition de contact, favorisant des phénotypes invasifs. L'objectif de mon travail de doctorat s'est articulé autour de l'étude de l'interaction entre une protéine PDZ précise, MAGI-1 et la protéine E6. Une première étape a été de produire et purifier le domaine PDZ1 de MAGI-1 ainsi que le domaine C-terminal de E6. Ensuite, j'ai participé à la mise au point de deux techniques de mesure d'interactions protéine-protéine, l'une basée sur la technologie Biacore et l'autre sur la rétention différentielle par chromatographie d'affinité. Enfin, j'ai étudié le complexe par RMN. Le calcul de structure de MAGI-1 PDZ1 complexé au peptide C-terminal de E6 sera achevé prochainement.

DISCIPLINE : Biologie Moléculaire

MOTS CLEFS : HPV, oncoprotéine E6, cancer du col de l'utérus, MAGI-1, PDZ, interaction, cinétique, Biacore, holdup assay, structure, RMN, dynamique

---

INTITULÉ ET ADRESSE DU LABORATOIRE : UMR 7175 – LC1 ; CNRS, Équipe Oncoprotéines ; ESBS ; 1, Bd Sébastien Brandt ; BP 10413 ; 67412 Illkirch Cedex