



HAL
open science

ÉTUDE STRUCTURALE DES PROTÉINES DE LA CAPSIDE DE L'ADÉNOVIRUS

Majida El Bakkouri

► **To cite this version:**

Majida El Bakkouri. ÉTUDE STRUCTURALE DES PROTÉINES DE LA CAPSIDE DE L'ADÉNOVIRUS. Biochimie [q-bio.BM]. Université Joseph-Fourier - Grenoble I, 2008. Français. NNT: . tel-00369461

HAL Id: tel-00369461

<https://theses.hal.science/tel-00369461>

Submitted on 19 Mar 2009

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

UNIVERSITÉ JOSEPH FOURIER-GRENOBLE I
ÉCOLE DOCTORALE CHIMIE ET SCIENCES DU VIVANT

THÈSE

Présentée par:

Majida EL BAKKOURI

Pour obtenir le grade de:

DOCTEUR DE L'UNIVERSITÉ JOSEPH FOURIER

Discipline: Biochimie et Biologie Structurale

**ÉTUDE STRUCTURALE DES PROTÉINES DE LA
CAPSIDE DE L'ADÉNOVIRUS**

Soutenue publiquement le 03 juillet 2008 devant le jury:

Dr. Eric J. Kremer	Président du jury
Dr. Saw See Hong	Rapporteur
Pr. Patrice Gouet	Rapporteur
Dr. Pascal Fender	Examineur interne
Pr. Rob W. H. Ruigrok	Codirecteur de thèse
Dr. Guy Schoehn	Directeur de thèse

Thèse préparée au sein de l'UVHCI

Unit of Virus Host Cell Interactions, UMR 5233 UJF-EMBL-CNRS

Remerciements

Beaucoup de personnes ont contribué au bon déroulement de ma thèse, je tiens à les remercier pour leur aide précieuse durant ces trois années.

*Je tiens tout d'abord à remercier chaleureusement le Dr. **Guy Schoehn**, mon directeur de thèse qui m'a soutenue avec patience et compétence. Grâce à lui j'ai pu travailler dans une atmosphère saine et enrichissante me donnant envie de poursuivre dans le monde de la recherche.*

*Je tiens à exprimer ma gratitude au Pr. **Rob Ruigrok** qui m'a donné l'opportunité de faire ma thèse dans son laboratoire dans un environnement scientifique excellent. Je le remercie pour son encouragement et son soutien, et d'être toujours si proche des ses thésards.*

*Je tiens à exprimer ma reconnaissance aux membres de mon jury : le Dr. **Eric J. Kremer** pour avoir accepté de présider mon jury de thèse ; le Dr. **Saw See Hong**, le Pr. **Patrice Gouet** et le Dr **Pascal Fender** pour l'honneur que vous me faites en acceptant de juger mes travaux de thèse.*

*Un grand merci au Pr. **Wim Burmeister**. Merci d'avoir toujours pris le temps de répondre à mes questions dans le domaine de la cristallographie.*

*Un grand merci au directeur de l'EMBL Grenoble, le Dr. **Stephen Cusack**, d'avoir pris sur son temps précieux pour la partie cristallographie de ma thèse. Sans oublier **Elena Seiradake**, **Hassan Belrhali** et **Andrew Mc Carthy**, merci de m'avoir initié à la collecte de données à l'ESRF et à la résolution de structure.*

*Mes remerciements vont aussi aux Dr. **Nicolas Tarbouriech**, Dr. **Carlo Petosa** et Dr. **Thibaut Crepin** pour leur patience, leur gentillesse et le partage de leur grand savoir en cristallographie.*

*Je tiens aussi à adresser ma gratitude au Dr. **Manuel Rosa-Calatrava** pour les clones de la protéine IX, au Dr. **Saw-See Hong** pour le clone de la protéine IIIa, au Dr. **Darren Hart** pour notre collaboration et l'obtention de constructions solubles de la protéine IIIa, au Pr. **Hugues Lortat-Jacob** pour les expériences de Biacore, au Dr. **Pascal Fender** pour les échantillons de la base de penton d'HAD3, au Dr. **Jean Pierre Simorre** pour les expériences en RMN, et bien sûr au Pr. **Marc Jamin** pour les expériences de MALLS-SEC.*

*Un grand merci à **Francine Gérard**, **Lucy Freeman**, **Celine Fabry**, Dr. **Aurélie Albertini** et Dr. **Cyril Dian** pour leur amitié et les bons moments que nous avons partagé ensemble. Sans oublier **Jocelyne Visconti**, **Alexandre Dias**, **Julien Pérard**, **Julien Lupo**, **Coly Pierre-Alain Coly**, **Sophie Torres**, Dr. **Euripedes Ribeiro**, **Ambroise Desfosses**, **Ricardo Pires**, Dr. **Charles Sabin**, **Pierre Metais**, Dr. **Denis Bouvier**, Dr. **Anja Drees**, Dr. **Andreas Hinz**, **Christelle Buon**, **Annie Simon**, **Delphine Guilligay** que je remercie pour leur bonne humeur et la bonne ambiance qui règne à l'UVHCI qui va beaucoup me manquer.*

*Un grand merci au Dr. **Irina Gutsche**, Dr. **Marlyse Buisson**, Dr. **Melina Haupt**, Dr. **Xavier Brazzolotto**, Dr. **Jean Baptiste Artero**, Dr. **Monika Budayova-Spano**, **Sandrine Vignon** et au Dr. **Leandro Estrozi** pour les moments agréables de la pause déjeuner.*

Merci à la région Rhône Alpes d'avoir financé cette thèse.

Merci à mes parents qui m'ont permis de faire de si longues études et qui n'ont jamais cessé de croire en moi, sans oublier le soutien précieux de mes frères et sœurs.

*Enfin je remercie **Charles Calmettes** l'homme de ma vie qui me supporte et me soutient depuis 4 ans déjà...*

A mes parents Fatima et Abdelaziz

SOMMAIRE

1.	INTRODUCTION	6
1.1.	<i>HISTOIRE ET CLASSIFICATION</i>	6
1.2.	<i>EPIDEMIOLOGIE</i>	8
1.3.	<i>MORPHOLOGIE ET STRUCTURE D'ADENOVIRUS</i>	11
1.3.1.	<i>Vue d'Ensemble du Virus</i>	11
1.3.2.	<i>Les Protéines Structurales Majeures</i>	14
a.	<i>L'Hexon (II)</i>	14
b.	<i>Le Penton</i>	16
b.1.	<i>La Base de Penton (III, 571 aa)</i>	17
b.2.	<i>La Fibre (IV, 582 aa pour HAd2)</i>	20
b.3.	<i>Le Dodécaèdre</i>	23
1.3.3.	<i>Les Protéines Structurales Mineures</i>	24
a.	<i>La Protéine IX</i>	25
b.	<i>La Protéine IIIa</i>	30
c.	<i>La Protéine VI</i>	31
d.	<i>La Protéine VIII</i>	32
1.3.4.	<i>Les Protéines du core</i>	33
1.3.5.	<i>La Protéase de l'Adénovirus</i>	34
1.3.6.	<i>Structure et Organisation du Génome d'Adénovirus</i>	36
1.4.	<i>LE CYCLE VIRAL</i>	38
1.4.1.	<i>Les Récepteurs Primaires des Adénovirus</i>	38
a.	<i>Le CAR (Coxsackievirus and Adenovirus Receptor)</i>	39
b.	<i>L'Acide Sialique</i>	41
c.	<i>Le CD46</i>	43
d.	<i>Les CD80 et CD86</i>	44
e.	<i>Le MHC-1 (Major histocompatibility complex class I)</i>	44
f.	<i>Le VCAM-1 (Vascular cell adhesion molecule 1)</i>	45
g.	<i>Les Héparanes Sulfates</i>	45
h.	<i>Les Intégrines $\alpha\beta 2$ et $\alpha L\beta 2$</i>	45
1.4.2.	<i>Fixation au Récepteur Secondaire et Internalisation</i>	46
1.4.3.	<i>La Phase Précoce d'Infection</i>	49
1.4.4.	<i>La Phase Tardive d'Infection</i>	50
1.5.	<i>L'INTERET DES ADENOVIRUS POUR LA THERAPIE GENIQUE</i>	51
	LES OBJECTIF PRINCIPAUX DE LA THESE	56
2.	MATERIELS ET METHODES	57
2.1.	<i>SOUS CLONAGE</i>	57
2.2.	<i>SUREXPRESSON DANS LE SYSTEME BACTERIEN</i>	59
2.3.	<i>SUREXPRESSON DANS LE SYSTEME BACULOVIRUS</i>	61
2.4.	<i>DETERMINATION DE L'ETAT D'OLIGOMERISATION</i>	62
2.4.1.	<i>MALLS-SEC</i>	62
2.4.2.	<i>Pontage Chimique</i>	63
2.5.	<i>LA CRISTALLOGRAPHIE AUX RAYONS X</i>	63
2.5.1.	<i>La Cristallogénèse</i>	64
2.5.2.	<i>Principe de la Diffraction aux Rayons X</i>	65
2.5.3.	<i>Traitement des Données</i>	66
2.5.4.	<i>Facteur de Structure et Densité Electronique</i>	67
2.5.5.	<i>Obtention des Phases</i>	68
2.5.6.	<i>Construction et Affinement du Modèle</i>	70
	RESULTATS ET DISCUSSION	71
3.	LA PROTEINE IX	72
3.1.	<i>LA SOLUBILITE DE LA PROTEINE IX</i>	72
3.2.	<i>ESSAI DE DETERMINATION DE LA STRUCTURE DE LA PROTEINE IX</i>	74
3.3.	<i>LOCALISATION DE LA PROTEINE IX</i>	76
3.3.1.	<i>Etats d'Oligomérisation de la Protéine IX</i>	77
a.	<i>La Région C-terminale de la Protéine IX</i>	77

b.	<i>La Protéine IX ΔSYL-V117D</i>	80
c.	<i>La Protéine IX ΔSYL</i>	82
d.	<i>La Protéine IX ΔSYL-Δala</i>	85
e.	<i>Conclusion</i>	87
3.3.2.	<i>Localisation du Domaine C-terminal de la Protéine IX dans la capsid d'HA5</i>	88
3.3.3.	<i>Localisation de la Partie C-terminale de la Protéine IX dans la Capsid du CAV-2</i>	89
a.	<i>Introduction</i>	89
b.	<i>ARTICLE I: Three-dimensional structure of canine adenovirus serotype 2 capsid</i>	89
c.	<i>Conclusion</i>	102
3.3.4.	<i>Discussion</i>	103
4.	<i>LA PROTEINE IIIA</i>	106
4.1.	<i>INTRODUCTION</i>	106
4.2.	<i>SOLUBILISATION A PARTIR DE CORPS D'INCLUSION</i>	107
4.3.	<i>RECHERCHE DE FRAGMENTS SOLUBLES</i>	109
4.4.	<i>CONCLUSION ET PERSPECTIVES</i>	111
5.	<i>LA BASE DE PENTON DE L'ADENOVIRUS HUMAIN DE SEROTYPE 3</i>	113
5.1.	<i>INTRODUCTION</i>	113
5.2.	<i>EXPRESSION ET PURIFICATION</i>	115
5.3.	<i>CRISTALLISATION DU DODECAEDRE DE L'HADV3</i>	117
5.4.	<i>COLLECTE ET TRAITEMENT DES DONNEES</i>	120
5.5.	<i>CONCLUSION ET PERSPECTIVE</i>	123
6.	<i>LA FIBRE COURTE DE L'ADENOVIRUS AVIAIRE DE SEROTYPE 1</i>	124
6.1.	<i>INTRODUCTION</i>	124
6.2.	<i>ARTICLE II: STRUCTURE OF THE C-TERMINAL HEAD DOMAIN OF THE FOWL ADENOVIRUS TYPE 1 SHORT FIBRE</i>	125
6.3.	<i>CONCLUSION</i>	134
	<i>CONCLUSION GENERALE</i>	136
	<i>REFERENCES BIBLIOGRAPHIQUES</i>	138

Liste des Abréviations:

3D:	Trois Dimensions
Å:	Angström
Ad:	Adénovirus
aa:	acide aminé
CAV:	Adénovirus Canin
CAR:	Coxsackie and Adenovirus Receptor
CELO:	Chicken Embryo Lethal Orphan
CHO:	Chinese Hamster Ovary
DBP:	DNA Binding Protein
E:	Early
ESRF:	European Synchrotron Radiation Facility
FAdV:	Fowl Adenovirus (Adénovirus aviaire)
GFP:	Green Fluorescent Protein
GON:	Group of Nine Hexons
HAd:	Adénovirus Humain
HS:	Héparane Sulfate
IPTG	Isopropyl-beta-D-thiogalactopyranoside
ITR:	Inverted Terminal Repeat
kDa:	kilo Dalton
kpb	kilo paire de bases
L:	Late
MALLS-SEC:	Multiple Angle Laser Light Scattering Size Exclusion Chromatography
MET:	Microscopie Électronique à Transmission
NCS:	Non Crystallographic Symmetry
PAGE:	PolyAcrylamide Gel Electrophoresis
PDB:	Protein Data Banque
PML:	Promyelocytic Leukemia
SDS:	Sodium Dodecyl Surface
SRP:	Surface Plasmon Resonance
TEV:	Tabacco Etch Virus
TP:	Terminal Protein

1. Introduction

1.1. Histoire et Classification

Les adénovirus ont été mis en évidence en 1953 par Rowe et *al.* à partir de fragments d'amygdale d'enfant et en 1954 par Hilleman et Werner chez des militaires souffrant d'infections respiratoires aiguës. Après mise en culture, les tissus dégénéraient en 2 à 3 semaines avec des anomalies morphologiques au niveau des cellules. Un agent viral transmissible a été isolé. Dénommé d'abord AD (**A**denoid **D**egenerating), ARG (**A**cute **R**espiratory **D**isease), RI (**R**espiratory **I**llness) ou encore APC (**A**denoidal **P**harygeal **C**onjunctival). Cet agent pathogène est désigné depuis 1956 sous le nom d'adénovirus (Ad) à cause de sa présence dans les amygdales (« adeno » signifie glande en grec). A l'heure actuelle, plus d'une centaine de souches ont été identifiées. La famille des *Adenoviridae* a pu ainsi être définie (Norby et al., 1976).

L'Ad existe probablement depuis plusieurs millions d'années (Benko and Harrach, 2003). Il a été détecté chez de nombreuses espèces, principalement chez les mammifères, les oiseaux et les amphibiens. La conservation de la région centrale du génome, fait de l'Ad un des meilleurs modèles pour l'étude de l'évolution virale.

Les Ads appartiennent à la famille des *Adenoviridae*, divisée en quatre genres principaux différenciables par leur organisation génomique. Ces quatre genres correspondent également aux quatre origines majeures de l'Ad: Mastadénovirus, Aviadénovirus, Atadénovirus et Siadénovirus. Les Mastadénovirus infectent une large gamme de mammifères incluant l'homme, le chien, le cheval, les ovins et les bovins. Chez l'homme, il n'existe pas moins de 51 sérotypes d'adénovirus subdivisés en 6 sous-groupes. Le genre Aviadénovirus infectant les oiseaux est composé de 3 sous-groupes. Ils se distinguent des autres Ads par la présence de deux fibres liées à la même base de penton sur chacun des 12 sommets du virus. Le sérotype le plus étudié est le FAdV-1 aussi connu sous le nom de CELO (*Chicken Embryo Lethal Orphan*). Les Atadénovirus se caractérisent par un fort pourcentage de A et T dans leur génome d'où leur nom Atadénovirus (Benko and Harrach, 1998). Ils présentent un large

spectre d'hôte et infectent plusieurs classes de vertébrés (reptiles, oiseaux, marsupiaux, mammifères), notamment les serpents, lézards, canards, oies, poules, etc. Pour les Siadénovirus, seuls deux membres sont connus, l'adénovirus de grenouille de type 1 et l'adénovirus de la dinde de type 3. Le génome de ces Ads code pour une sialidase hypothétique à l'origine de leur nom, Siadénovirus. Un seul adénovirus est connu pour infecter le poisson ; plus éloigné génétiquement des 4 premiers genres, il pourrait constituer le dernier genre connu à ce jour (Benko et al., 2002; Davison et al., 2003) (Fig. 1.1).

La classification des Ads de chaque genre en différents sous-groupes dépend de plusieurs propriétés, telles que la séquence, le pourcentage en GC contenu dans le génome, le pouvoir oncogène, les familles d'hôtes, le tropisme, les propriétés d'hémagglutination, ou encore l'organisation génétique de la région précoce E3.

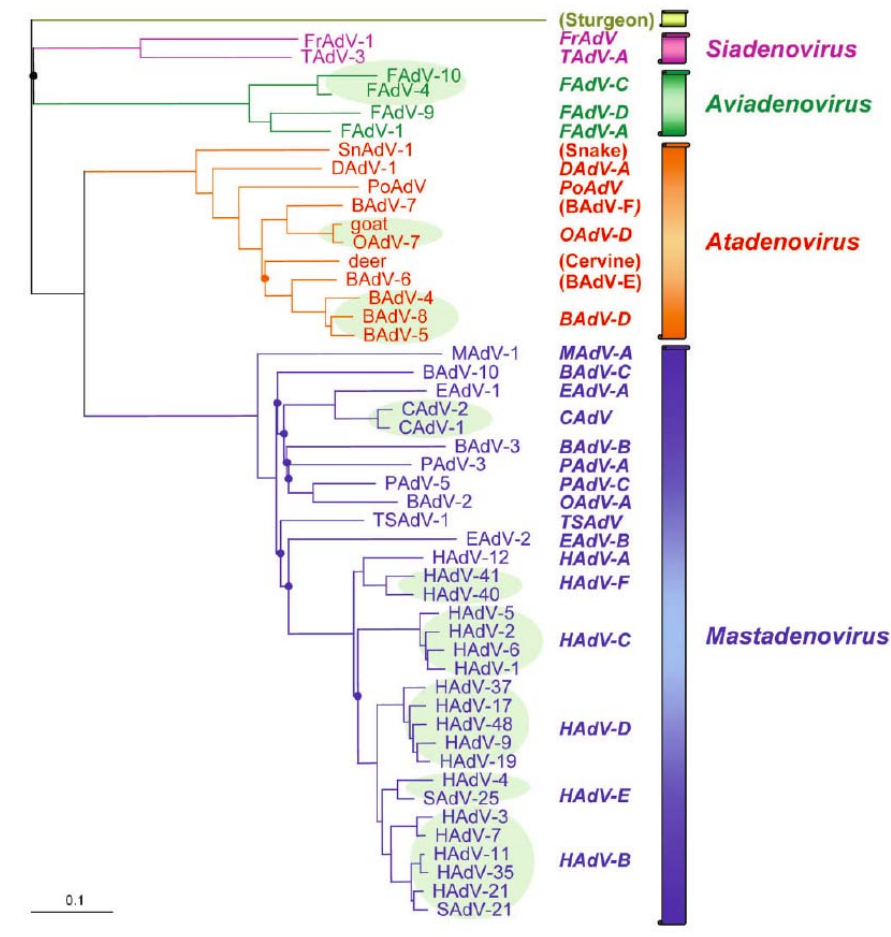


Figure 1.1: Arbre phylogénétique des adénovirus basé sur les séquences d'acides aminés de l'hexon. Les membres des différents genres sont indiqués avec différentes couleurs, et les virus appartenant au même sous-groupe sont regroupés dans un ovale vert. Le nom abrégé du virus est indiqué à la fin de chaque branche avec le nom de l'espèce à droite en italique: B-bovin; C-canin ; D-canard ; E-equine ; F-volaille (fowl) ; Fr-grenouille ; H-humain ; M-murin ; O-ovin ; P-porcine ; Po-opossum; Sn-serpent ; T-dinde (turkey) ; S-primate et TS-scandentien (tree shrew). L'image est reprise de Davison et al. (2003).

1.2. *Epidémiologie*

L'adénovirus humain (HAd) est un agent viral pathogène capable de proliférer efficacement dans une grande variété de types cellulaires. Suivant leur sérotype, les Ads ont un tropisme pour l'appareil respiratoire, l'œil ou le tube digestif. Ils causent le plus fréquemment des infections respiratoires et diverses maladies chez les enfants et personnes immunodéprimées. Les principales maladies ainsi que les sérotypes associés sont référencés ci-dessous et dans le tableau 1.1:

Maladie	Sous-groupe (sérotype)
• Infections respiratoires	B ; C
• Maladies respiratoires aiguës	B (7, 21) ; E (4)
• Maladie des voies respiratoires inférieures	B (3, 7, 21) ; E (4)
• Fièvre pharyngoconjonctivale	B (3, 7,) ; E (4)
• Conjonctivites (des piscines) et kératoconjonctivites épidémiques	D (8, 19a, 37)
• Conjonctivite hémorragique aiguë	B (11)
• Cystites hémorragiques	B (7, 11, 21, 35)
• Gastroentérites	A (31), C (2), F (40, 41)

Les symptômes des maladies respiratoires causées par une infection par l'HAd peuvent être multiples: rhume, pneumonie et bronchite. Les patients immuno-déprimés sont particulièrement sujets à des complications sévères suite à l'infection par l'HAd. Des maladies respiratoires aiguës dues à des Ads ont été diagnostiquées chez des recrues de la deuxième guerre mondiale dans des conditions de promiscuité et de stress.

Les modes de transmission des HAds varient selon leur sérotype, mais tous se transmettent par contact direct ou féco-oral et occasionnellement par immersion dans une eau contaminée. Plusieurs types d'infections par des HAds sont capables de s'établir comme des infections asymptomatiques persistantes des amygdales, des adénoïdes et de l'intestin des hôtes infectés. Parfois une dispersion virale peut survenir pendant des mois ou des années.

Tableau 1.1: Le tropisme des différents sérotypes d'HAAd.

Sous groupe	Sérotipe	Tropisme
A	<u>12</u> , 18, 31	Intestinal
B : 1	3, <u>7</u> , 16, 21, 50	Respiratoire, oculaire
B : 2	<u>11</u> , 14, 34, <u>35</u>	Respiratoire, urinaire, oculaire, rénal
C	<u>1</u> , <u>2</u> , <u>5</u> , 6	Respiratoire, lymphatique
D	8-10, 13, 15, <u>17</u> , 19, 20, 22-30, 32, 33, 36, 37-39, 42-51	Oculaire, intestinal
E	<u>4</u>	Respiratoire, oculaire
F	<u>40</u> , 41	Intestinal

Effets pathologiques selon les sous-groupes d'adénovirus humain:

Sous-groupe A: ce sous-groupe comporte l'HAAd12, HAAd18 et l'HAAd31. Ces virus sont difficiles à isoler et à cultiver bien que les infections qu'ils provoquent soient courantes. Ils sont isolés en majorité chez l'enfant chez lequel ils provoquent des gastroentérites intestinales dans 60% des cas.

Sous-groupe B: ce sous-groupe est subdivisé en deux: B1 (HAAd3, HAAd7, HAAd16, HAAd21 et HAAd50) et B2 (HAAd11, HAAd14, HAAd34 et HAAd35). Les maladies provoquées par les virus du sous-groupe B1 sont le plus souvent d'ordre respiratoire (mais parfois oculaire ou systémique). Les virus du sous-groupe B2 sont responsables d'infections persistantes des voies urinaires (ils peuvent également causer des infections oculaires ou respiratoires légères). Certains sérotypes (HAAd11, l'HAAd34 et l'HAAd35) sont aussi responsables d'infections rénales fatales chez les patients immunodéprimés (patients porteurs du virus HIV) (Kojaoghlanian et al., 2003; Leen and Rooney, 2005; Li et al., 1991).

Sous-groupe C: ce sous-groupe inclut les virus HAAd1, HAAd2, HAAd5 et HAAd6. Ils sont les premiers adénovirus à avoir été identifiés (Rowe et al., 1953). Plus de 80% de la population humaine a déjà été infectée par ces virus comme le montre la présence d'anticorps dirigés contre ces virus (Garnett et al., 2002). Ils sont endémiques dans certaines parties du

monde et généralement la primo-infection date de l'enfance. Les HAd2 et l'HAd5 sont largement étudiés du point de vue structural et biologique du fait de leur faible pouvoir pathogène et de leur facilité de production en culture cellulaire.

Sous-groupe D: ce sous-groupe englobe 32 sérotypes différents cités dans le tableau 1.1. L'HAd8, HAd19 et l'HAd37 ont un tropisme oculaire et sont responsables de kératoconjunctivites épidémiques. Les autres sérotypes étaient rarement rencontrés mais deviennent de plus en plus fréquents chez les patients immunodéficients (HAd9 ; HAd17 ; HAd20 ; HAd22 ; HAd23 ; HAd26 ; HAd27 ; HAd42-49 et HAd51) (De Jong et al., 1999; Hierholzer, 1992; Hierholzer et al., 1988).

Sous-groupe E: ce sous-groupe ne contient qu'un seul sérotype: l'HAd4. Ce virus est responsable de maladies respiratoires aiguës, ainsi que d'une conjonctivite folliculaire (Aoki et al., 1982).

Sous-groupe F: ce sous-groupe comprend les Ads entériques HAd40 et HAd41 qui causent habituellement des gastro-entérites persistantes voire fatales chez l'enfant (Uhnoo et al., 1984). Ces virus sont difficiles à cultiver en culture cellulaire.

Pour certains sérotypes, la maladie associée dépend de la voie d'infection ; par exemple l'inhalation de l'HAd7 est associée à une infection sévère des voies respiratoires inférieures alors que la transmission orale du virus cause typiquement une maladie bénigne.

Un composé anti-Adénovirus (*hexadecyloxypropyl-cidofovir*) a été découvert récemment (Toth et al., 2008). Ce composé agit en réduisant la réplication virale dans les organes ciblés chez le hamster.

Un vaccin a été développé contre l'HAd4 et 7 qui provoquent des maladies respiratoires aiguës (Top et al., 1971), mais ils ne sont disponibles que pour les recrues militaires. La production de ce vaccin a été interrompue en 1996 mais devrait reprendre pour l'armée américaine en cours d'année 2008.

Chez les bovins et les ovins, les Ads sont la cause de maladies respiratoires bénignes. Chez le chien il existe deux sérotypes (CAV-1 et 2). Le CAV-1 provoque une hépatite infectieuse canine aiguë appelée aussi « Hépatite de Rubarth ». Le CAV-2 est responsable d'une gastro-entérite hémorragique mortelle chez le chiot, ainsi que d'une infection

respiratoire bénigne. Un vaccin existe et protègent contre ces deux sérotypes. Il est constitué d'adénovirus canin de type 2 vivant atténué, car les deux adénovirus canins possèdent une parenté antigénique étroite.

1.3. Morphologie et Structure d'Adénovirus

1.3.1. Vue d'Ensemble du Virus

L'adénovirus était le premier échantillon biologique à être observé par microscopie électronique par Horne et ses collaborateurs en 1959. Ces premières observations en coloration négative ont permis de voir que l'adénovirus avait une forme régulière et qu'il était composé de 20 facettes. La première structure tridimensionnelle d'un HAd a été déterminée par Stewart et *al.* en 1991 grâce à l'analyse des images obtenues par cryo-microscopie électronique (cryoMET) (Stewart et al., 1991). Cette structure à 35 Å de résolution a permis d'analyser l'organisation complexe de la capsid. Le nombre de triangulation a en particulier pu être déterminé grâce à la visualisation de la localisation des protéines majeures. Quelques protéines mineures (les protéines IX et IIIa) jouant le rôle de ciment ont également été localisées dans la capsid de l'Ad. En 1993, une reconstruction à 25 Å de résolution a permis le recalage de la structure cristallographique de l'hexon dans la carte de MET, révélant plus d'information sur l'organisation de la capsid (Stewart et al., 1993). En 2005, une reconstruction à 10 Å de résolution a été publiée (Fabry et al., 2005). Elle a permis la construction d'un modèle quasi-atomique après recalage de la structure atomique de l'hexon et de la base de penton. Cette structure a apporté des informations détaillées sur les contacts existants entre les capsomères dans la capsid. La qualité de la reconstruction a également permis l'observation des boucles flexibles de l'hexon et de la base du penton qui n'avaient pas été résolues par cristallographie. Le calcul d'une carte de différence en soustrayant le modèle quasi-atomique de la reconstruction de cryoMET a révélé la position des protéines mineures XI, IIIa et VIII. La dernière reconstruction a été obtenue à 6 Å de résolution (Saban et al., 2006). Celle-ci a permis la construction d'un modèle quasi-atomique plus précis que le précédent avec les protéines majeures. L'analyse de la carte de différence a permis l'observation d'hélices α dans la capsid. En se basant sur la carte de différence et sur des prédictions de structures secondaires des protéines mineures, leurs positions respectives ont été attribuées. La localisation de la protéine IIIa en particulier a été mise à jour ainsi que celle de la protéine IX.

Les adénovirus sont des virus non enveloppés à symétrie icosaédrique, d'environ 90 nanomètres de diamètre sans la fibre (Fig. 1.2). Une particule comporte plus de 2700 polypeptides, avec une masse molaire d'environ 150 MDa. Le génome viral est composé d'un ADN double brin linéaire non segmenté. L'acide nucléique représente 13% de la masse totale du virus et les protéines 87%. Le génome entier a été séquencé (Roberts et al., 1984) et la stœchiométrie des différents polypeptides a été déterminée (van Oostrum and Burnett, 1985). Les protéines de l'Ad sont numérotées de II à X suivant leur migration décroissante (et donc leur masse molaire) sur un gel électrophorétique (Fig. 1.3).

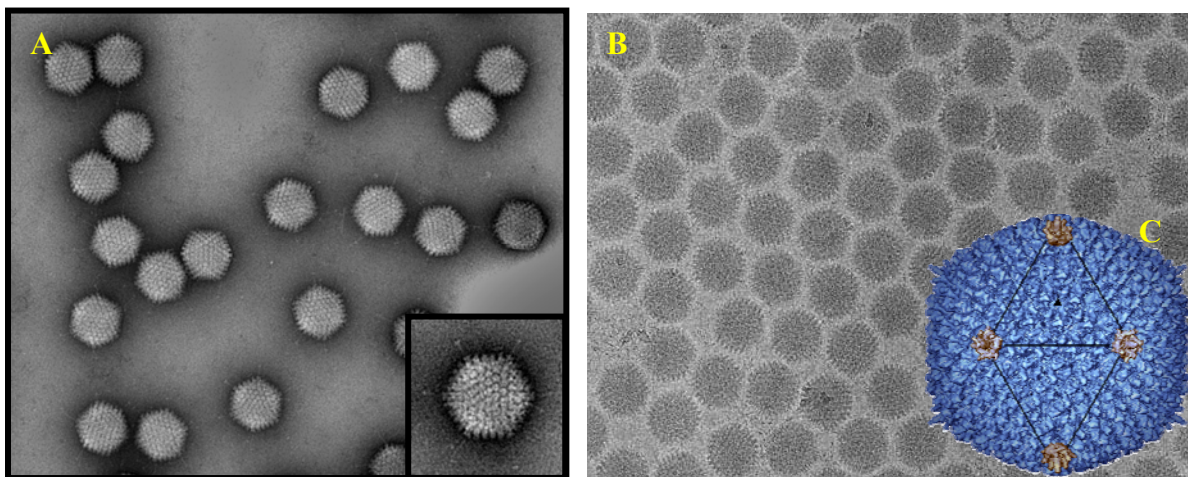


Figure 1.2: L'adénovirus vu par microscopie électronique.

A, L'HAd41 observé par microscopie électronique en coloration négative.

B, L'HAd5 observé en cryo-microscopie électronique.

C, Reconstruction 3D de l'HAd5 calculée à partir d'images de cryoMET. La fibre est flexible et n'apparaît pas dans les reconstructions de cryoMET (images et construction produites par Guy Schoehn).

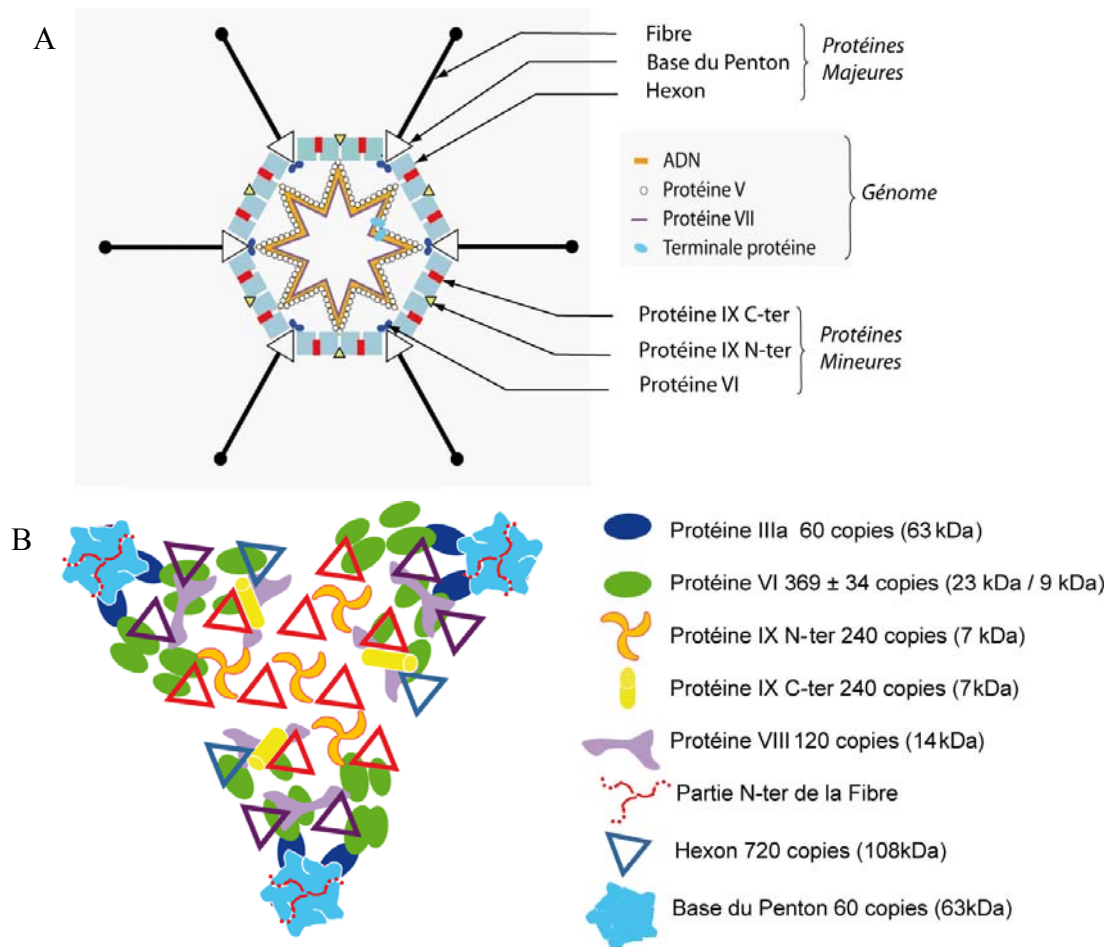


Figure 1.3: Vue schématique de l'adénovirus.

A, vue d'ensemble de la capside montrant la localisation des protéines majeures (Hexon, Base de penton et Fibre), des protéines mineures (protéine IX et VI) et les protéines du associées au génome.

B, Représentation d'une facette de l'icosaèdre avec en bleu la base de penton (protéine pentamérique) à chaque extrémité de la facette triangulaire. Les triangles représentent l'hexon (protéine trimérique). Les neuf triangles rouges au centre de la facette forment avec la protéine IX (sous forme de quatre trimères) un complexe stable nommé GONs (Group Of Nine hexons). Les triangles violets représentent les hexons péripentaux. La protéine IX est représentée en orange (le domaine N-terminal) et en jaune (le domaine C-terminal). La protéine VIII est en violet ; la protéine VI est en vert ; la protéine IIIa est en bleu foncé. La protéine IIIa ainsi que la protéine VI interagissent avec la base du penton.

1.3.2. *Les Protéines Structurales Majeures*

La capsid virale est constituée de 3 protéines dites majeures: l'hexon, la base de penton et la fibre (II, III, et IV respectivement) (Fig. 1.3). Elles sont dites « majeures » car elles constituent la majeure partie de la capsid.

a. *L'Hexon (II)*

L'hexon est la protéine la plus abondante du virus (720 copies), composée de 951 aa (pour HAd5), elle est aussi la plus grosse en taille, représentant 63% de la masse totale des protéines du virus. Les 720 monomères de la protéine hexon forment 240 hexons trimériques qui vont constituer les 20 facettes triangulaires de l'icosaèdre. Chacune des facettes du virus est constituée de 12 hexons trimériques (Fig. 1.3).

La structure atomique de l'hexon a été déterminée, par cristallographie, pour l'HAd2 (Roberts et al., 1986), l'HAd5 (Rux and Burnett, 2000) (Fig. 1.4), l'adénovirus aviaire (Xu et al., 2007) et l'adénovirus du chimpanzé (Pichla-Gollon et al., 2007). La séquence de l'hexon est très conservée parmi les Ads connus et séquencés, les structures sont donc très similaires. Chaque monomère est constitué de 2 tonneaux β à 8-brins (Fig. 1.4). Chaque trimère contient donc 6 tonneaux β , ce qui lui donne son apparence pseudo hexagonale aussi bien à basse résolution en cristallographie (Fig. 1.4B) (Burnett et al., 1985) que sur les hexons isolés observés en microscopie électronique (van Oostrum et al., 1987). Ce repliement en tonneau β ou '*viral jelly roll*' est caractéristique de certains virus. En effet, l'hexon présente un repliement similaire à celui de la protéine P3 du bactériophage PRD1 (phage des bactéries gram négatif) qui est vraisemblablement un cousin éloigné de l'Adénovirus (Benson et al., 2002).

Comme le montre la figure 1.4A, l'hexon est constitué de plusieurs régions (Rux et al., 2003):

- La région N-terminale (NT) est formée d'une hélice, d'une boucle et assure la liaison avec les deux autres monomères dans la partie basale de l'hexon à l'intérieur de la capsid.

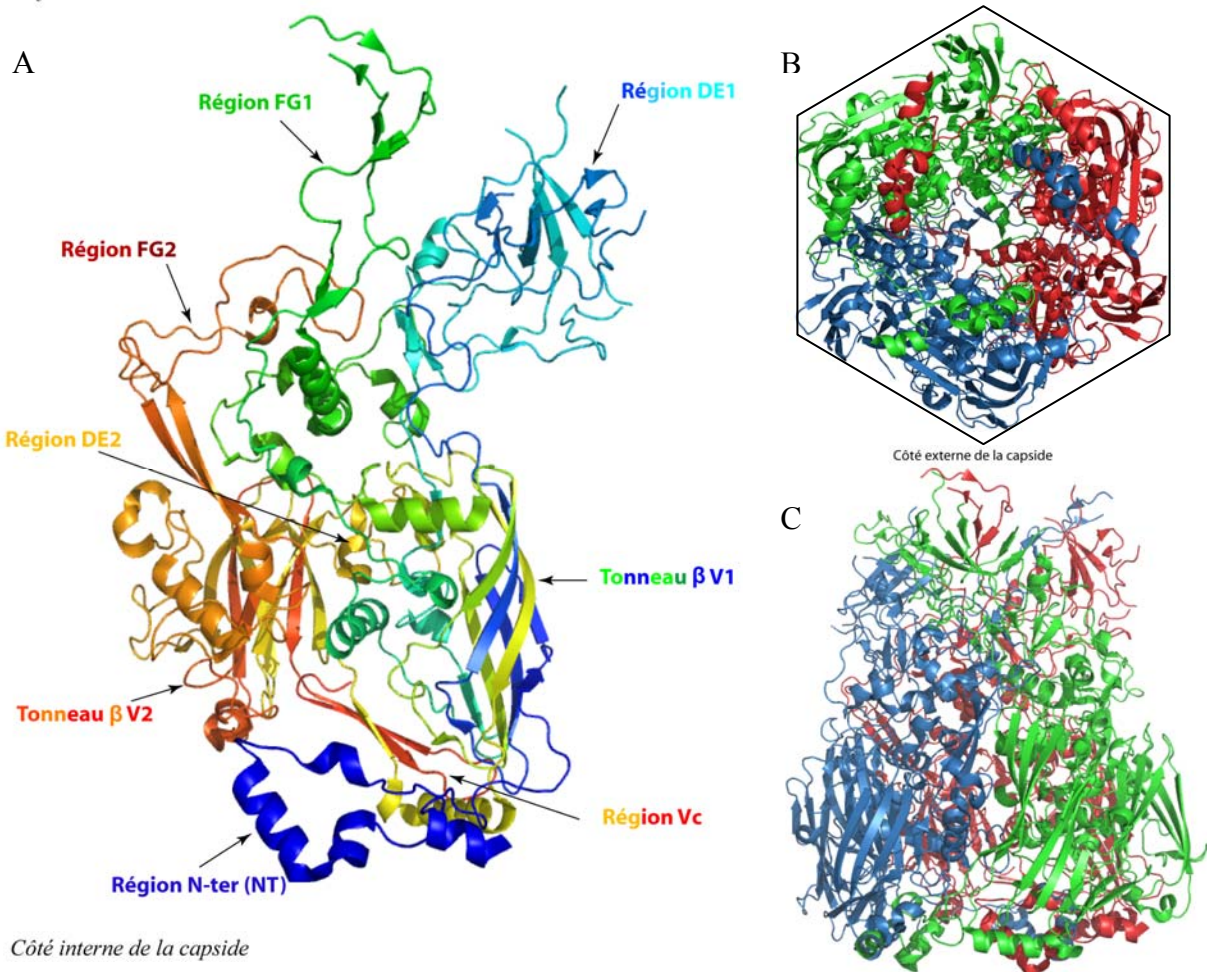
- V1 est le plus grand des deux tonneaux β , DE1 et FG1 sont deux grandes boucles insérées entre les brins β D/E et F/G respectivement du domaine V1. La boucle DE1 constitue la région la moins conservée de l'hexon. La boucle FG1 est plus petite que la boucle DE1, elle assure la liaison avec la boucle DE1 d'un monomère adjacent au sein du trimère.

- **V2** est le plus petit des deux tonneaux. Comme pour le domaine V1, les deux domaines DE2 et FG2 (insertion entre les brins β D/E et F/G) sont issus de ce tonneau.

- **VC** est le domaine de connexion entre les deux tonneaux β V1 et V2. Il est constitué d'une hélice et de trois feuillets permettant de maintenir une structure tridimensionnelle stable entre les deux tonneaux β du côté interne de la capside.

La présence des deux tonneaux β est probablement due à une duplication de gène (Athappilly et al., 1994). Cette propriété permet de distinguer les virus à deux tonneaux β tel que l'adénovirus et le bactériophage PRD1 des virus à un seul tonneau β comme les picornavirus (Bamford et al., 2002).

Surface du virion



Côté interne de la capside

Figure 1.4: Structure cristallographique de l'hexon de l'HA2.

A, Présentation d'un monomère, les domaines DE1, FG1, FG2, sont du côté externe de la capside ; le domaine VC et la boucle DE2 sont indiqués ; ils assurent la liaison entre les deux tonneaux β V1 et V2. La région N-terminale (NT) est située dans la partie basale de l'hexon, du côté interne de la capside.

B, Représentation d'un trimère vu de dessous selon l'axe d'ordre 3. La forme pseudo-hexagonale est soulignée par la présence d'un hexagone noir.

C, Vue de côté d'un trimère montrant que contrairement à la base du penton, l'hexon est plus large dans sa partie basale et plus étroite dans sa partie apicale (Rux et al., 2003).

Des études fonctionnelles et structurales de l'hexon de l'HAd2 ont montré que ce dernier est capable d'interagir avec la DPPC (*dipalmitoyl phosphatidylcholine*), un composé de la surface pulmonaire (Balakireva et al., 2003). Il est possible que l'hexon intervienne dans une nouvelle voie d'internalisation du virus au niveau du tractus respiratoire. Au cours de la réplication, l'hexon est important dans le transport du virus du cytoplasme vers le noyau (Suomalainen et al., 1999). Récemment, des travaux publiés sur l'HAd5 ont révélé une nouvelle fonction de l'hexon: celui-ci est capable d'interagir avec le facteur de coagulation FX permettant ainsi l'infection du foie sans implication de l'interaction fibre-CAR (*Coxsackie and Adenovirus Receptor*). Ceci est une propriété intéressante car elle permet une voie de transduction alternative lorsque l'interaction avec le CAR n'est pas valide (exemple de transmission intravasculaire) (Waddington et al., 2008).

b. Le Penton

Le penton est localisé au niveau des 12 sommets de l'icosaèdre, il est constitué par l'assemblage non covalent de deux protéines (Fig. 1.5): la base du penton et la fibre. La base de penton est un pentamère de la protéine III, alors que la fibre est un trimère de la protéine IV. Le penton contient tous les éléments nécessaires pour assurer l'attachement et l'internalisation du virus: la fibre permet la reconnaissance et l'attachement du virion à un récepteur cellulaire ; la base du penton induit l'internalisation du virion grâce à l'interaction avec les intégrines cellulaires.

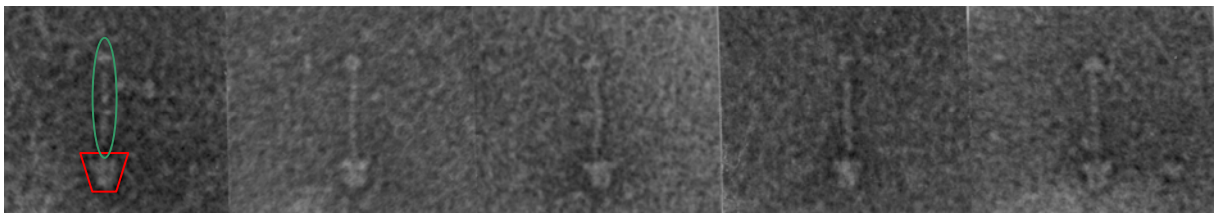


Figure 1.5: Images de microscopie électronique en coloration négative du penton. Sur la première image, la base de penton est encadrée en rouge et la fibre en vert (Image de Guy Schoehn).

b.1. La Base de Penton (III, 571 aa)

C'est une protéine pentamérique qui est localisée au niveau des 12 sommets de l'icosaèdre. Elle joue non seulement un rôle majeur dans la structure et la géométrie des particules icosaédriques mais permet également la fixation et l'exposition de la fibre à la surface de la particule virale. Le rôle physiologique de la base de penton n'est pour l'instant pas très clair. La base de penton semble être impliquée dans la libération du virus des endosomes (Greber et al., 1993), ses partenaires cellulaires sont encore inconnus. La seule interaction connue de la base de penton avec la cellule est celle de son motif RGD (Arg-Gly-Asp) avec les intégrines cellulaires $\alpha\beta3$ ou $\alpha\beta5$, il survient pendant l'internalisation des particules virales. La base de penton contient deux motifs polyprolines (xPPxY) très conservés dans sa partie N-terminale. Ces séquences sont capables d'interagir avec plusieurs protéines *via* le domaine WW de ces dernières. Ces domaines hydrophobes sont impliqués dans de multiples interactions et leur présence permet à la protéine de servir d'adaptateur dans la formation des complexes fonctionnels constitués de plusieurs composants. Parmi les protéines à domaine WW identifiées qui interagissent avec la base du penton, trois contiennent le domaine HECT caractéristique de la ligase d'ubiquitine (Galinier et al., 2002).

La séquence primaire de la base de penton est hautement conservée au sein de tous les Ads humains comme non humains (typiquement on trouve 70% d'identité). Les différences majeures sont localisées dans deux boucles indiquées sur la figure 1.6A. La première est appelée la boucle variable, la deuxième, appelée la boucle RGD, est caractérisée par des longueurs qui diffèrent d'un sérotype à l'autre. La boucle RGD contient le motif RGD hautement conservé chez les Ads humains excepté l'HAd40 et 41 (Albinsson and Kidd, 1999; Zubieta et al., 2005). Le motif RGD permet la liaison du virus à la surface de la cellule *via* les intégrines α_v avant l'internalisation de l'Ad (Bai et al., 1993; Murakami et al., 2007). Parmi les Ads humains, seuls ceux du sous-groupe F (HAd40-41) ne possèdent pas cette séquence. Malgré sa conservation, une délétion de ce motif n'induit pas une perte totale du pouvoir d'infection de l'adénovirus. Il a été montré que pour l'internalisation de l'HAd5, l'interaction fibre-récepteur est suffisante bien que moins efficace que lorsque l'interaction RGD-intégrines est fonctionnelle (Rux and Burnett, 2004). L'internalisation après interaction avec les intégrines n'est donc pas la seule voie de pénétration des Ads dans les cellules.

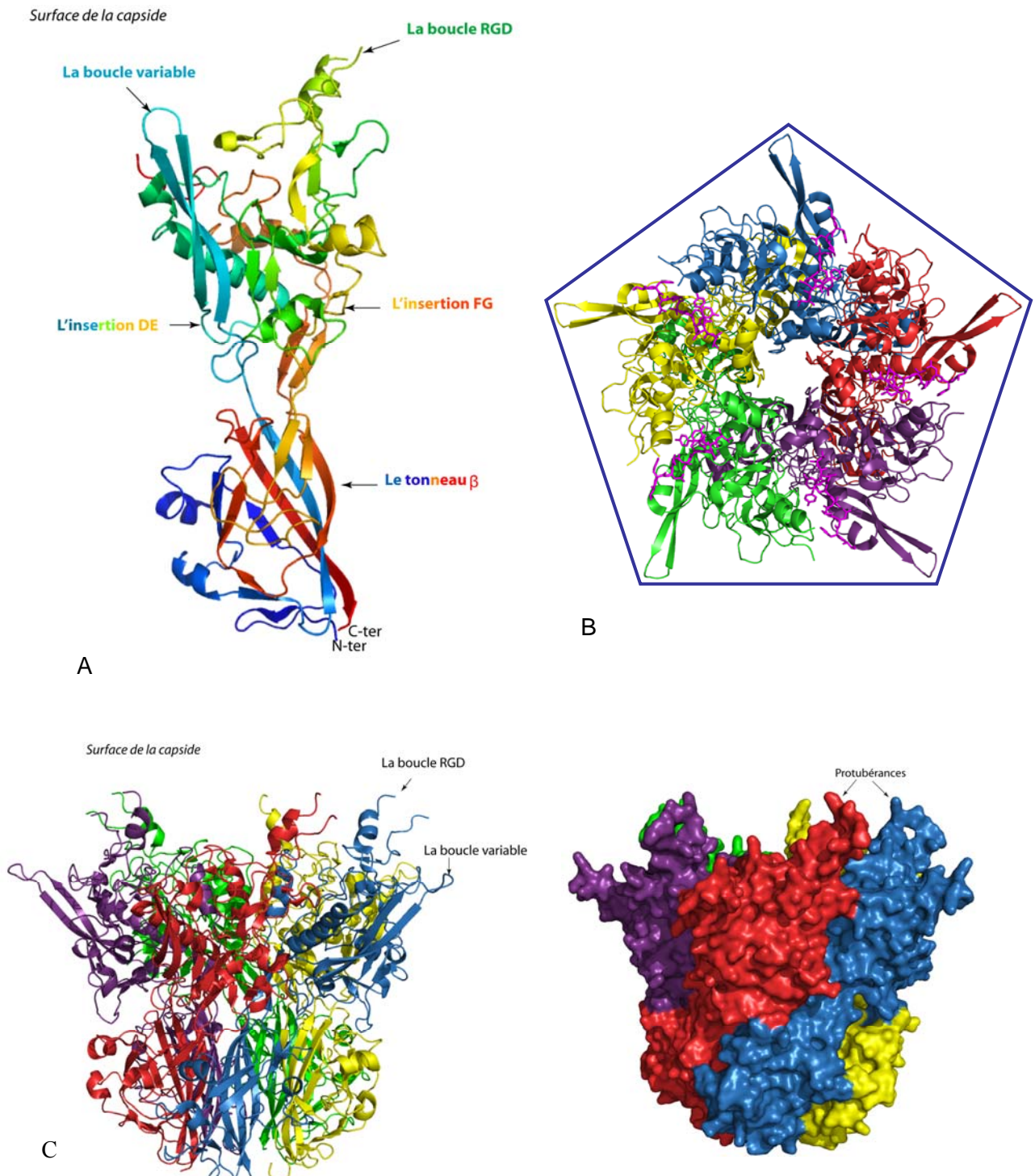


Figure 1.6: Structure cristallographique de la base du penton de l'HAd2 (Zubieta et al., 2005).

A, Architecture générale du monomère avec le tonneau β en bas et les deux insertions FG et DE en haut. La boucle RGD est indiquée par une flèche sur la figure, elle est absente dans la structure probablement en raison d'un clivage ou d'une forte flexibilité.

B, Structure atomique de la base du penton vue du dessus. Chacun des monomères est représenté en une couleur différente. Le peptide mimant la partie N-terminale de la fibre (FNPVYPY) et qui a été co-cristallisé avec la base de penton est représentée en rose.

C, Représentation de structure secondaire (à gauche) ou de surface (à droite) d'un pentamère de la base du penton vu de côté. Chaque monomère est coloré différemment. L'orientation dans la capsid est indiquée ainsi que la boucle RGD et la boucle variable.

La structure atomique de la base de penton de l'HA2 a été déterminée récemment au laboratoire (Zubieta et al., 2005). Elle révèle un arrangement circulaire de 5 monomères de base de penton créant une cavité centrale d'un diamètre d'environ 30Å formée du côté apical. La partie supérieure de la base du penton est assez plane, seule la boucle contenant le motif RGD forme une protubérance plus ou moins longue (Fig 1.6). Chaque monomère est composé par deux domaines distincts. Le domaine basal, constitué de 8 brins β antiparallèles organisés en tonneau β , est situé du côté interne de la capsid. Le domaine externe est constitué de deux longues extensions insérées entre les brins E-D et F-G du tonneau β . Cette topologie est très similaire à celle de l'hexon et semble indiquer une origine commune des structures, probablement par duplication de gène. L'insertion ED contient la boucle RGD ainsi que la boucle variable de la base de penton. Ces deux boucles sont exposées à la surface du virus et ne contribuent pas aux contacts avec l'hexon. Elles sont probablement les plus exposées à des mutations au cours de l'évolution. Par conséquent, c'est probablement ce type de régions qui fait la spécificité de chaque sérotype.

Dans le penton entier, la fibre se lie à la base grâce à un motif conservé de sept acides aminés (FNPVYPY) du côté N-terminal (la queue). La structure de la base de penton d'HA2 a été également résolue en complexe avec un peptide de 20 acides aminés mimant la partie N-terminale universellement conservée de la fibre (Fig. 1.6B). Seuls les acides aminés 11 à 19 ont pu être visualisés dans la structure cristallographique. Dans le complexe base du penton – peptide, cinq peptides occupent des sites équivalents dans chacun des monomères. Dans le virus, seul trois des cinq sites sont occupés. Deux des monomères de la fibre se fixent probablement sur deux sites voisins et le troisième occuperait le site opposé laissant deux sites libres. Dans le cas des Ads aviaires dont les bases du penton portent deux fibres, ce modèle n'explique pas comment deux fibres peuvent se fixer à une base de penton. Il faut noter que les fibres aviaires ne contiennent pas le motif conservé dans la partie N-terminale de la fibre présent chez les adénovirus humain.

Pour quelques sérotypes, 12 pentons sont capables de s'associer pour former un complexe symétrique appelé dodécaèdre (Fender et al., 2005; Fender et al., 1997; Fender et al., 2008). Ces dodécaèdres peuvent être détectés dans les cellules infectées. Cependant, ils se forment également lorsque la base seule ou la base plus la fibre sont exprimées dans le système baculovirus/cellules d'insectes. Ces dodécaèdres avec ou sans fibre sont capables de rentrer dans la cellule et de s'accumuler au niveau de la membrane nucléaire.

b.2. La Fibre (IV, 582 aa pour HAd2)

La fibre est un homotrimère de la protéine IV projetée vers l'extérieur sur les 12 sommets de l'icosaèdre. La fibre est composée de trois domaines structuraux: le domaine N-terminal appelé la queue de la fibre ; la région centrale allongée appelée la tige et le domaine C-terminal globulaire appelé la tête de la fibre (Fig. 1.7). La queue permet la liaison non covalente de la fibre à la base de penton *via* son motif conservé de 7 acides aminés (11-FNPVYPY). La tige à une structure allongée mesurant de 9 nm (HAd3) à 36 nm (HAd40) pour les adénovirus humain (Tableau 1.2). Dans le cas des virus non humains, elle peut mesurer jusqu'à 77.5 nm. La tête de la fibre est responsable de l'attachement de l'adénovirus à la surface de la cellule *via* le récepteur primaire.

Le rôle majeur de la tête de la fibre est la reconnaissance du récepteur primaire à la surface de la cellule lors de l'infection par l'adénovirus (*cf.* paragraphe sur les récepteurs). Par conséquent, elle détermine le tropisme de l'adénovirus pour un tissu cellulaire spécifique. Des travaux sont menés par de nombreux groupes dans le but de savoir si cette protéine joue un autre rôle durant l'infection. De nombreuses protéines virales sont synthétisées sous forme de précurseur donnant naissance à la protéine mature après clivage par la protéase de l'Ad (Anderson, 1990; Anderson et al., 1973). Des études ont montré que ce processus de clivage pourrait être défectueux si la fibre était délétée (Chee-Sheung and Ginsberg, 1982; D'Halluin et al., 1980; Falgout and Ketner, 1988).

Dans le cadre d'études fondamentales de l'Ad et du développement de vecteurs en thérapies géniques, de multiples études structurales ont été menées et ont débouchées sur la résolution de la structure de nombreuses têtes de fibre à l'échelle atomique. La première structure de tête de fibre de l'HAd a été résolue en 1994 avec celle de l'HAd5 (Xia et al., 1994) suivie par celle de l'HAd2 (van Raaij et al., 1999a) ; HAd12 (Bewley et al., 1999); HAd3 (Durmort et al., 2001); HAd37 ; la fibre courte de l'HAd41 (Seiradake and Cusack, 2005); HAd11 (Persson et al., 2007) et finalement celle de l'HAd35 (Wang et al., 2007). Parmi les structures connues de têtes de fibres d'adénovirus infectant les animaux, on trouve celle de l'adénovirus canin (Seiradake et al., 2006) et les deux fibres de l'adénovirus aviaire, la longue (Guardado-Calvo et al., 2007) et la courte (*cf.* partie résultats: fibre). Ces différentes têtes de fibres forment toutes un trimère adoptant un repliement similaire malgré de grandes différences au niveau la séquence. Ces structures sont constituées de deux feuillets β de 4 à 5

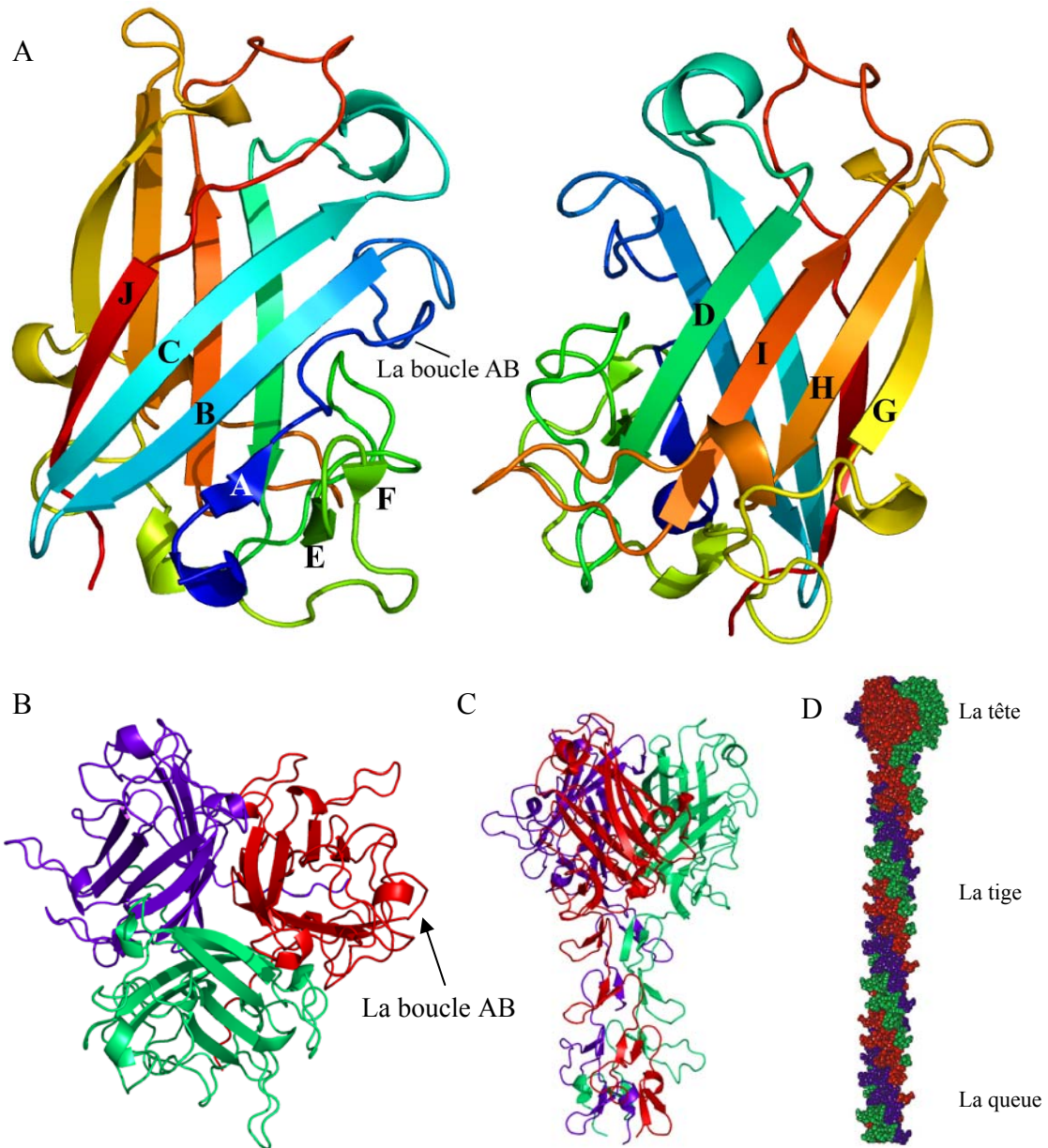


Figure 1.7: La structure de la fibre de l'HAd2.

A, Structure cristallographique d'un monomère de la tête de la fibre de l'HAd2 déterminée à 1.51 Å de résolution (code PDB 1QHV). Le monomère vu de deux côtés est coloré du bleu vers le rouge (N-terminal vers C-terminal).

B, Un trimère de la tête de fibre d'HAd2 vu de dessous selon son axe 3, chaque chaîne est colorée différemment.

C, Structure cristallographique déterminée à 2.4 Å de résolution d'une partie de la fibre de l'HAd2 contenant la tête et 4 motifs de la tige montrant l'organisation de celle-ci en triple spirale bêta (code PDB 1QIU).

D, Modèle de la fibre complète contenant les 23 motifs de la tige proposées par Van Raaij et al. (1999b).

brins antiparallèles arrangés en un sandwich β (Fig. 1.7). Dans le trimère, le feuillet β constitué par les brins DIHG est exposé au solvant alors que le feuillet ABCJ est situé à l'interface du trimère ; les deux feuillets sont liés *via* un cœur hydrophobe riche en résidus aromatiques. Les brins E et F qui ne sont pas toujours présents dans les structures des fibres sont parfois remplacés par des boucles (Xia et al., 1994). La boucle AB est l'une des régions les plus conservées d'un point de vue structural, elle est impliquée dans l'interaction avec le récepteur majeur des adénovirus nommé CAR (plus de détails dans les parties introduction: cycle viral et résultats : fibre).

La tige est caractérisée par la répétition d'un motif hydrophobe d'une quinzaine d'acides aminés contenant une glycine et une proline hautement conservées à des endroits spécifiques (Ruigrok et al., 1990; Ruigrok et al., 1994). La seule structure atomique d'une partie de la tige de la fibre de l'HAd2 a été déterminée en 1999 (van Raaij et al., 1999b) (Fig. 1.7C et 1.8). Cette molécule cristallisée était en effet issue de la fibre de l'HAd2 et comprenait la tête plus 4 motifs répétés de la tige. La structure est caractérisée par un repliement spécifique en « triple spirale bêta » où chaque répétition (ou motif) est constituée d'un brin β orienté vers la tête de la fibre et parallèle à l'axe de la tige suivie d'un coude contenant une glycine ou une proline. Le coude est suivi d'un autre brin β orienté à l'opposé du premier et formant un angle de 45° avec l'axe de la tige. Le nombre de motifs (qui détermine aussi la longueur de la fibre) varie en fonction du sous-groupe, entre 5 (HAd3 sous-groupe B1) et 23 motifs (HAd2 sous-groupe C). Le motif répété minimal est composé de 15 acides aminés mais il existe des variations pouvant aller jusqu'à 20 acides aminés. Ces variations induisent la formation de coude permettant une flexibilité de la tige (Schoehn et al., 2008). Il a été montré que la flexibilité de la fibre facilitait l'interaction entre la tête de la fibre et son récepteur à la surface cellulaire (Wu et al., 2003).

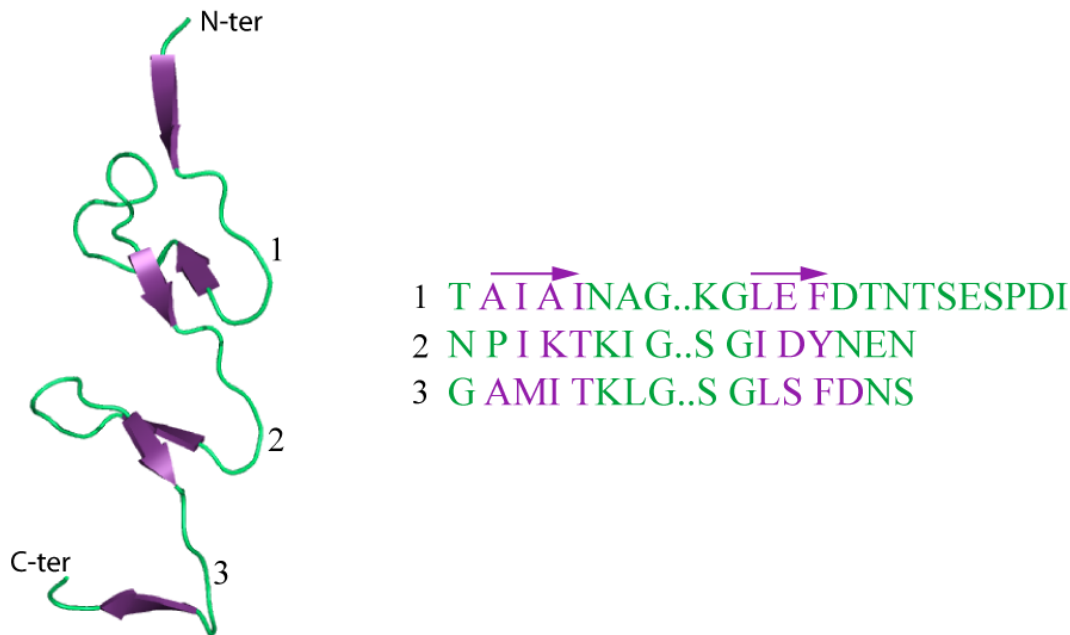


Figure 1.8: La Structure de la tige de la fibre. Exemple de 3 motifs dans le repliement en « triple spirale bêta » de la tige de fibre d'HA2. 1, 2 et 3 correspondent au motifs 19, 20 et 21 dans la structure cristallographique de la tête de fibre d'HA2 (Fig. 1.7C).

b.3. Le Dodécaèdre

C'est un complexe protéique composé par l'association symétrique non covalente de 12 pentons qui interagissent entre eux par l'intermédiaire de la partie inférieure de leur base (Fig. 1.9). Les bases seules (sans la fibre) peuvent s'associer pour donner naissance à des dodécaèdres sans fibre. Les dodécaèdres ont notamment été observés dans des extraits de cellules infectées par l'HA3, HA4, HA7, HA9 et HA15 (Fender et al., 1997) mais leur rôle reste pour l'instant inconnu. Ces complexes sont capables de s'attacher et d'entrer dans les cellules humaines selon un processus d'endocytose proche de celui utilisé par les Ads. La base du penton seule est suffisante pour déclencher ces mécanismes d'endocytose. Ces caractéristiques en font un candidat potentiel pour un vecteur de transfert de gènes. Cette particule possède en effet une cavité de 80 Å et offre une efficacité de transfection comparable aux vecteurs adénoviraux existant actuellement (Fender et al., 1997; Schoehn et al., 1996). Les dodécaèdres étant de plus dépourvus de génome, ils présentent l'avantage d'être naturellement déficients pour la réplication. Une seule (ou deux) protéine(s) entrent dans leurs compositions ce qui pourrait limiter les risques de réactions immunitaires induites avec les vecteurs viraux classiques.

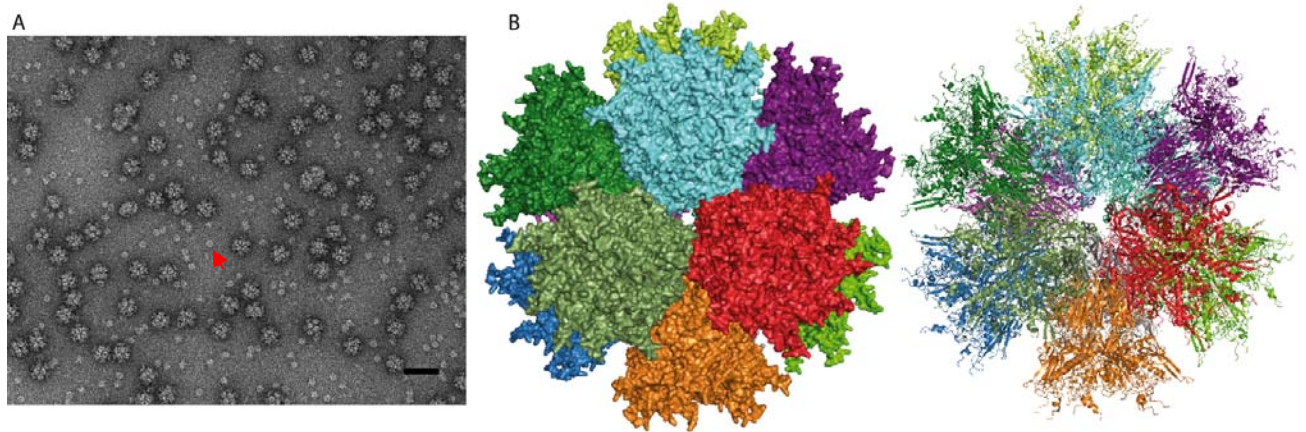


Figure 1.9: L'architecture du dodécaèdre.

A, Image de MET en coloration négative des particules dodécaédriques de la base du penton de l'HA3, l'arrière plan montre des bases du penton libres indiquées par une flèche. La barre d'échelle représente 50 nm.

B, Structure atomique du dodécaèdre de l'HA2: représentation de la surface à gauche (vue selon l'axe de symétrie d'ordre 3 de la particule); les structures secondaires sont présentées à droite.

Il a été montré récemment que le dodécaèdre, contrairement au virus entier, pouvait interagir avec les héparanes sulfates protéoglycanes (Fender et al., 2008; Vives et al., 2004). Ces sucres qui sont omniprésents à la surface de la cellule, permettent d'augmenter la fixation des dodécaèdres sur les cellules par rapport aux virus. Par conséquent, l'étape de pénétration dans la cellule est également accrue (Fender et al. 2008). Les structures des dodécaèdres avec et sans fibres de l'HA3 ont été déterminées par microscopie électronique (Fuschiotti et al. 2006). Contrairement à l'HA3, des dodécaèdres de pentons de l'HA2 n'ont jamais été observés *in vivo*. Cependant, les conditions de cristallogénèse utilisées pour obtenir des cristaux des bases d'HA2 ont induit la formation de ce type de particule (Zubieta et al., 2006; Zubieta et al., 2005) (Fig. 1.9). La structure atomique de la base du penton de l'HA2 a pu être résolue à 3.3 Å de résolution grâce à la reconstruction de cryo-microscopie électronique qui a servi de modèle pour le phasage des données. La structure atomique de la base du penton de l'HA3 n'est pour l'instant pas résolue.

1.3.3. Les Protéines Structurales Mineures

Les protéines mineures (IX, IIIa, VI, et VIII ; Fig. 1.3) ont fait l'objet de beaucoup moins d'études structurales que les protéines majeures, leurs structures et leurs fonctions sont donc moins connues. Ces protéines jouent un rôle de ciment dans la capsid et assurent l'étanchéité de celle-ci. Elles interviennent également durant l'assemblage et le désassemblage de la capsid (Burnett, 1985). Aucune donnée cristallographique n'est

actuellement disponible pour les protéines mineures. De plus, elles sont généralement trop petites pour être directement visibles par microscopie électronique. Seules les reconstructions tridimensionnelles de virus entier obtenues par cryo-microscopie électronique combinées avec les données cristallographiques existantes sur les protéines majeures ont permis, par différence d'images, de localiser des densités supplémentaires dues à ces protéines mineures. L'attribution de chacune des protéines mineures aux différentes densités supplémentaires détectées s'est avérée difficile. Elle a été surtout basée sur les données concernant la stœchiométrie de ces protéines au sein de la capsid (Fabry et al., 2005; Saban et al., 2006; Stewart et al., 1991; Stewart et al., 1993).

a. La Protéine IX

De toutes les protéines mineures, la protéine IX est la plus étudiée du point de vue de sa fonction et de sa localisation aussi bien au sein de la capsid que dans les cellules infectées. Avec ses 14.4 kDa (140 résidus), la protéine IX est la plus petite des protéines mineures de l'adénovirus. Contrairement aux autres protéines mineures, la protéine IX est spécifique aux Mastadénovirus. On peut dénombrer 12 copies de la protéine IX par facette pour les particules icosaédriques (Fig. 1.3 ; 1.10 et 1.11). La localisation de ces protéines IX a été possible grâce à une caractérisation de la capsid de l'adénovirus. En effet, lorsque l'on traite ce virus avec de la pyridine, les capsides se dissocient en capsomères mais également en une structure appelée le groupe des neuf hexons ou « *group of nine hexons* » (GON) (Colby and Shenk, 1981; Everitt et al., 1973; Prage et al., 1970). Ces GON représentent la partie centrale de la facette de l'HAd (Fig. 1.10A). Une combinaison d'analyse d'image de coloration négative de ces GONs, de STEM (*Scanning Transmission Electron Microscope*) et de cristallographie aux rayons X (structure atomique de l'hexon) a permis la localisation des protéines IX au sein des GONs et donc de la capsid (Furcinitti et al., 1989). Les 240 protéines IX présentes dans la capsid se répartissent en groupe de 4 trimères sur chacune des 20 faces du virus (Fabry et al., 2005; Furcinitti et al., 1989; Stewart et al., 1991; Stewart et al., 1993). Jusqu'en 2006, la communauté scientifique admettait que les triskélions présents entre les hexons du GON constituaient l'intégralité de la protéine IX. Une reconstruction 3D de l'adénovirus à 6Å de résolution a permis d'avoir plus de précision sur la position des domaines N et C-terminaux de la IX (Saban et al., 2006). Comme le montre la figure 1.11, le triskélion que l'on attribuait à la totalité de la protéine IX serait uniquement sa partie N-terminale alors que plusieurs domaines C-terminaux s'associeraient pour donner naissance à la densité cylindrique

initialement proposée pour être la protéine IIIa. Ces attributions ont été faites en se basant sur les hélices α visibles dans la carte de MET et sur les prédictions informatiques de structures secondaires. Ce modèle sera discuté dans la partie résultat: protéine IX.

Le rôle majeur de la protéine IX semble être un rôle structural. En effet, la stabilité du virus augmente en présence de la protéine IX: un mutant d'HAAd5 délété de la protéine IX est capable de proliférer de façon quasiment aussi efficace que le sauvage, mais ne forme plus de GONs ; le virion est également plus sensible à la chaleur (Boulanger et al., 1979; Colby and Shenk, 1981). La protéine IX est identifiée comme une protéine essentielle pour la stabilisation optimale du virus, mais elle n'est pas requise pour la formation de virus viable.

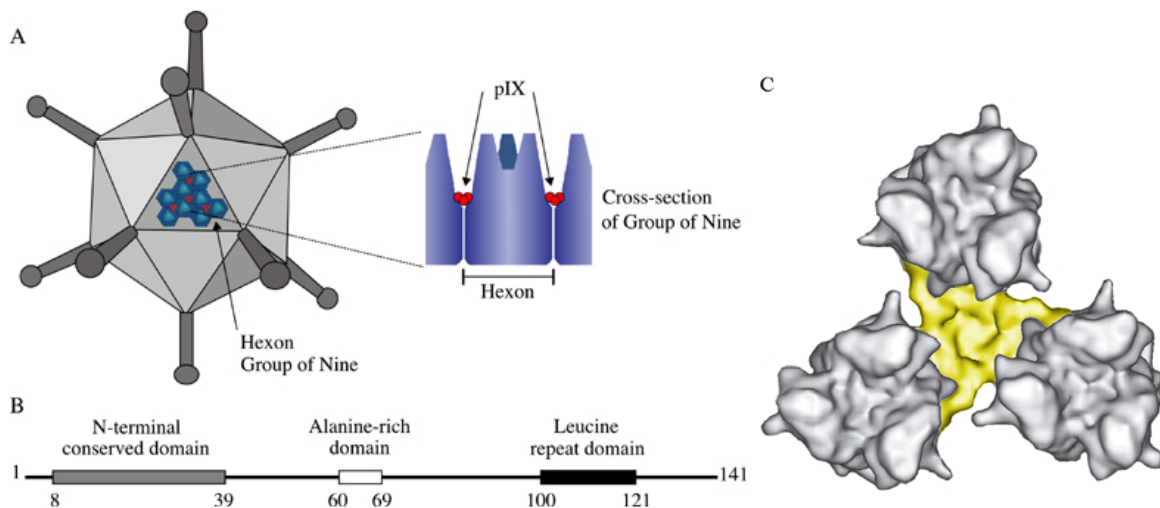


Figure 1.10: La position de la protéine IX dans l'adénovirus.

A, Schéma de la capsid de l'adénovirus. Chaque facette de l'icosaèdre est composée de 12 hexons, un groupe de 9 hexons constitue le GON au centre de la facette (en bleu et voir figure 1.3). Le GON est stabilisé par l'association de 4 trimères de protéines IX (en rouge). La protéine IX stabilise les interactions hexon-hexon (image (Parks, 2005)).

B, L'organisation de la protéine IX en trois domaines.

C, Détail de la structure du centre d'une face de l'adénovirus déterminée par analyse d'images de cryoMET et reconstruction 3D. Trois hexons (gris) et un trimère de protéines IX (jaune) sont représentés (Fabry et al., 2005).

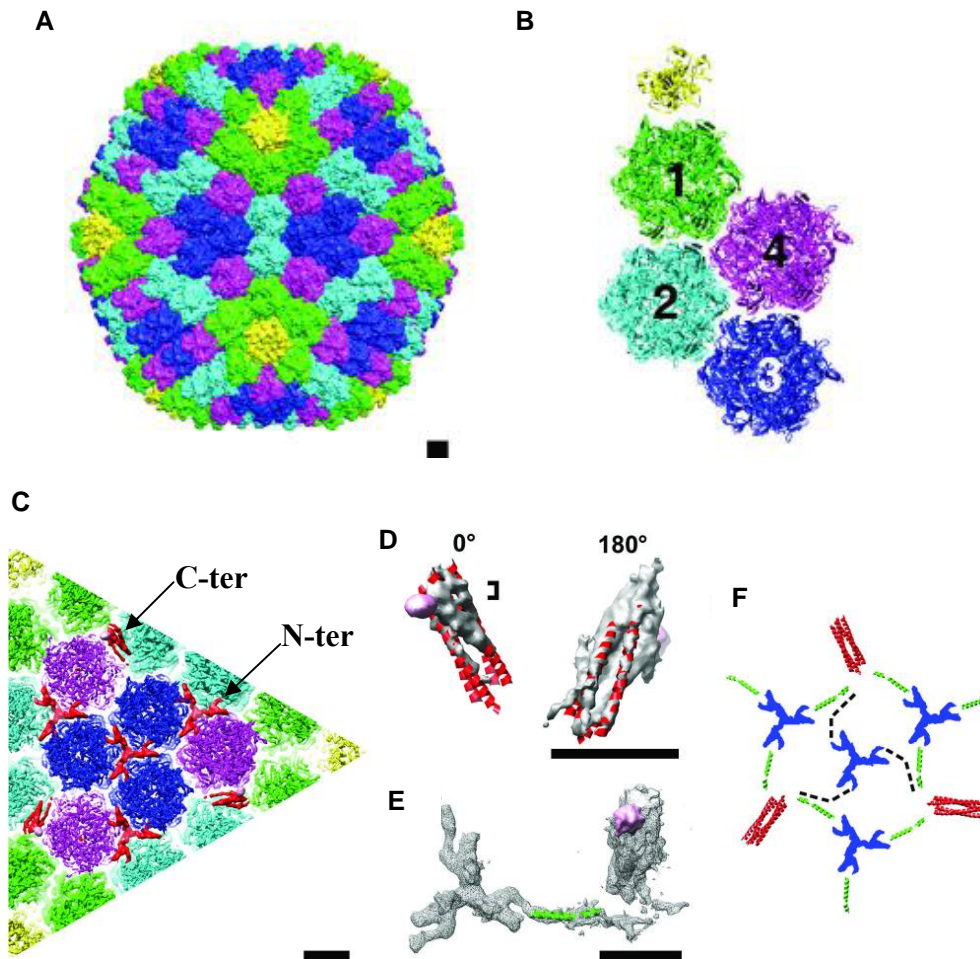


Figure 1.11: Modèle d'organisation de la protéine IX à la surface de la capside proposé par Saban et al. (2006).

A, Un modèle quasi-atomique d'HAd35 obtenu après recalage des structures cristallographiques de l'hexon et de la base de penton dans la densité électronique issue de la reconstruction obtenue par cryoMET. La barre d'échelle, 50 Å.

B, L'unité asymétrique de l'icosaèdre avec 4 hexons colorés en vert, bleu ciel, bleu, violet et un monomère de la base de penton en jaune.

C, Détail d'une facette du virus avec en rouge la carte de différence entre le modèle quasi-atomique et la reconstruction 3D issue de la microscopie électronique, mettant en évidence la localisation des protéines mineures. Ces densités en rouge ont été attribuées à la protéine IX, les flèches indiquent les domaines C- et N-terminaux de cette protéine.

D, Vue détaillée de la densité attribuée à la région C-terminale de la protéine IX. Des densités cylindriques pouvant être occupées par des hélices α sont visibles en 4 exemplaires dans cette densité.

E, Vue détaillée de la densité totale attribuée à la protéine IX. Le triskélion à gauche est un trimère constitué par 3 domaines N-terminaux de la protéine IX. Ce domaine se poursuit par deux hélices (domaine central) et finie par l'association de 4 domaines C-terminaux (à droite).

F, Vue schématique de l'organisation de la protéine IX à la surface de la capside. Les 4 densités trimériques en bleu sont assimilées à 12 régions N-terminales (amino-acide 1 à 56), l'hélice α en vert (aa 57 à 91) constitue la partie centrale de la protéine. Enfin le domaine constitué par 4 hélices α est attribué au domaine C-terminal de la protéine IX (aa 92 à 131), les auteurs proposent qu'une des 4 hélices de la région C-terminale provienne d'une protéine IX de la facette adjacente. Ce schéma est en accord avec les prédictions de structures secondaires concernant la protéine IX.

En plus de son rôle structural, la protéine IX joue de multiples fonctions dans la biologie du virus. Elle possède notamment une activité transcriptionnelle pour les gènes majeurs tardifs d'adénovirus (Lutz et al., 1997). De plus, elle contribue activement à la réorganisation des protéines nucléaires de la cellule hôte lors d'une infection par l'Ad dans le but de fournir un environnement plus favorable à la réplication du virus. En effet, durant l'infection, la protéine IX s'accumule dans le noyau de la cellule et forme des structures nucléaires amorphes observées en microscopie électronique et appelées inclusion (Rosa-Calatrava et al., 2001).

Il a été montré par mutagenèse dirigée que la protéine IX était constituée de plusieurs domaines fonctionnels. Ainsi, le domaine N-terminal, hautement conservé entre les différents sérotypes, est essentiel pour l'incorporation de la protéine IX dans la capsid. C'est ce domaine qui interagit avec les hexons. Le domaine C-terminal, riche en leucine et prédit pour former un *coiled-coil*, serait responsable de la trimérisation de cette protéine (Rosa-Calatrava et al., 2001; Vellinga et al., 2005) et potentiellement impliqué dans l'interaction avec d'autres protéines partenaires (Rosa-Calatrava et al., 2001). Cette propriété est utile dans l'activité transactivatrice nécessitant une interaction avec des partenaires cellulaires. L'intégrité du domaine *coiled-coil* est également essentielle à la formation des inclusions pour la réorganisation nucléaire lors de l'infection. La zone centrale riche en alanine ainsi que le domaine C-terminal sont utiles pour l'activité transcriptionnelle de la protéine, mais ne sont pas utiles pour l'incorporation de la protéine IX dans la capsid. En conclusion, l'incorporation de la protéine IX au sein de la capsid ne dépend pas de sa trimérisation mais plutôt de l'intégrité de son domaine N-terminal (Rosa-Calatrava et al., 2003; Vellinga et al., 2005).

L'activité transcriptionnelle de la protéine IX reste à prouver: des résultats récents ont montré que la protéine IX ne possède qu'une modeste activité sur les promoteurs E1A et E4 et que cette protéine n'est pas essentielle pour la réplication adénovirale (Sargent et al., 2004a). Ces contradictions proviennent probablement des différences dans les systèmes cellulaires utilisés par différents groupes.

Après infection par l'Ad, l'ultrastructure du noyau cellulaire change. Des inclusions contenant la protéine IX sont formées dans le noyau. Il a été prouvé par mutation que le domaine C-terminal (une région leucine-zipper) de la protéine IX est impliquée dans la formation de ces corps d'inclusion (Rosa-Calatrava et al., 2001; Souquere-Besse et al., 2002). La protéine PML (*cellular promyelocytic leukaemia*), impliquée dans la régulation de la réponse cellulaire antivirale est détectée dans les inclusions au cours de la phase tardive de

l'infection. Sa séquestration dans ces inclusions va donc aider le virus à se reproduire (Rosa-Calatrava et al., 2003).

Il a également été montré que la protéine IX joue un rôle dans l'encapsidation de l'ADN par les HAdS. Un virion possédant une protéine IX intacte est capable d'encapsider 1.5 à 2.0 kpb d'ADN en plus de son génome alors qu'un virion dont le génôme est délété du gène de la protéine IX perd environ 2 kpb en capacité d'encapsidation (Sargent et al., 2004b). Pour ce virus mutant l'encapsidation du génome sauvage est également moins efficace (Ghosh-Choudhury et al., 1987).

Modification de la Protéine IX

La protéine IX est accessible à la surface du virus en de nombreux exemplaires. Ceci fait d'elle un candidat de choix à modifier dans le but d'un reciblage spécifique des vecteurs d'Ad avec des perspectives d'application pratique en thérapie génique (*cf.* paragraphe thérapie génique). Comme le domaine C-terminal est plus accessible que le domaine N-terminal (Akalu et al., 1999) et que sa modification n'influe pas sur son intégration dans la capsid, c'est sur cette partie qu'ont été concentrés les essais de modification. Ces modifications ont pour but de changer le tropisme cellulaire des adénovirus et plusieurs voies ont été explorées pour obtenir une preuve de faisabilité:

- (i) Addition d'une séquence poly-lysines pouvant interagir avec les héparanes sulfates présents à la surface de la cellule ou addition d'un FLAG (un octapeptide DYKDDDDK) qui a permis de prouver l'accessibilité du domaine C-terminal grâce à l'anticorps anti-FLAG (Dmitriev et al., 2002).
- (ii) Addition d'un motif RGD reconnu par les intégrines cellulaires (Wu et al., 2002).
- (iii) Addition d'une hélice α de longueur variable placée entre un ligand et l'extrémité C-terminale de la protéine IX. La longueur de l'hélice peut être directement corrélée avec l'accessibilité du ligand à la surface de la capsid (Vellinga et al., 2004).
- (iv) Addition de la GFP (*green fluorescent protein*) (Le et al., 2004).
- (v) Addition d'un peptide accepteur de la biotine qui permettra la liaison d'un ligand biotinylé à la surface de la capsid (Campos et al., 2004).

b. La Protéine IIIa

La protéine IIIa est présente en 60 copies dans la capside de l'adénovirus (van Oostrum and Burnett, 1985). Elle est synthétisée sous forme de précurseur de 67 kDa et le clivage de celui-ci par une protéase virale durant la maturation du virus donne naissance à une protéine mature de 570 acides aminés (63.5 kDa pour l'HA2). Sa présence dans la capside est biologiquement significative ; plusieurs mutants thermosensibles affectant la protéine IIIa montrent un défaut d'assemblage et les particules relarguées s'accumulent vides à des températures non permissives (Chroboczek et al., 1986; D'Halluin et al., 1982). Le défaut d'assemblage peut être dû à un clivage défectueux qui bloque la maturation de la protéine IIIa pour des températures non permissives (Boudin et al., 1980).

Pendant longtemps il a été admis que la protéine IIIa avait une forme allongée et faisait la jonction entre deux hexons appartenant à des GONs différents (San Martin and Burnett, 2003; Stewart et al., 1993). Dans le modèle actuel, cette position est attribuée au domaine C-terminal de la protéine IX (Saban et al., 2006). Ce même modèle montre que la protéine IIIa est associée à la base de penton du côté interne de la capside (Fig. 1.12). Elle est d'une part en contact avec la base du penton et de l'autre avec l'ADN. La localisation interne de la protéine IIIa est confirmée. Une étiquette six-histidine a été incorporée dans le domaine N-terminal de la protéine IIIa. L'insertion dans le gène codant a été vérifiée par PCR et aucune évidence sur l'exposition de l'étiquette à la surface de particules intactes n'a été mise en évidence (Curiel, D. T. (2003), *Capsid-modified recombinant adenovirus and methods of use*. United States patent 6555368). Des expériences de co-immunoprécipitation ont montré que la protéine IIIa interagissait avec la protéine VII qui est localisée du côté interne de la capside (Boudin et al., 1980; Stewart et al., 1993).

Au moins une partie de la protéine IIIa est phosphorylée (San Martin and Burnett, 2003; Tsuzuki and Luftig, 1983) ; l'impact physiologique de cette phosphorylation est pour l'instant inconnu. Il a été montré qu'une expression de la protéine IIIa durant la phase précoce de l'infection induisait une baisse significative de synthèse des protéines virales de la phase tardive ainsi qu'un blocage modéré de la réplication de l'ADN virale. Ceci indique un rôle de la protéine IIIa dans l'expression de gène tardif, cependant le mécanisme est encore inconnu (Molin et al., 2002).

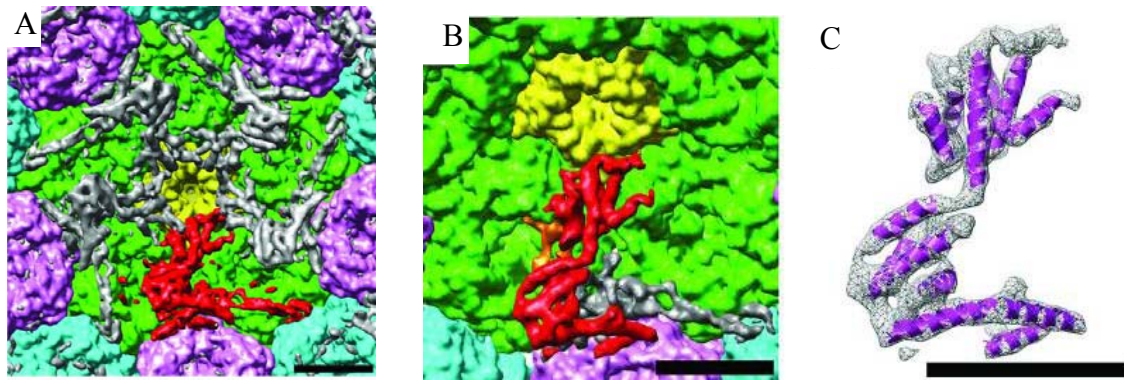


Figure 1.12: La position attribuée à la protéine IIIa dans une reconstruction en microscopie électronique à 6 Å de résolution (Saban et al., 2006).

A, Vue du côté interne de la capside au niveau d'un sommet de l'icosaèdre (la base de penton). En rouge et gris la carte de différence entre le modèle quasi atomique et la construction d'Ad par cryoMET. La carte de différence correspondant ainsi aux protéines mineures. Une densité en rouge correspond à un monomère de la protéine IIIa plus un monomère de la protéine VIII.

B, Vue d'un monomère de la protéine IIIa en rouge, avec une copie de la protéine VIII en gris ainsi qu'une densité orange non attribuée et qui pourrait correspondre à la protéine VI.

C, Un recalage de 13 hélices α dans la densité correspondante à la protéine IIIa. La barre d'échelle, 50 Å.

c. La Protéine VI

La protéine mature de 22 kDa est générée après un clivage du précurseur des côtés N et C-terminaux par la protéase de l'adénovirus (Lehmborg et al., 1999). Cette protéine est localisée du côté interne de la capside probablement entre deux hexons péripentonaux et permet la connexion de la capside avec la nucléocapside du virion. L'association de la protéine VI et l'hexon a été mise en évidence (Matthews and Russell, 1994), ainsi que la liaison à l'ADN (Matthews and Russell, 1995). Elle est présente en 369 ± 34 copies par particule (Lehmborg et al., 1999) et prédite être repliée en deux longues hélices α . Le précurseur VI joue plusieurs rôles biologiques dans différentes étapes de propagation et de maturation du virus. Elle est à la fois un substrat et un activateur de la protéase virale. En effet, le peptide clivé en C-terminal lors de la maturation de cette protéine sert de cofacteur pour la protéase de l'Ad (Mangel et al., 1993; Webster et al., 1993). La protéine VI régule également l'import de l'hexon dans le noyau durant l'assemblage de l'Ad (Wodrich et al., 2003). Le rôle de la protéine VI dans l'import et l'export nucléaire est dû à la présence de deux séquences signal de localisation nucléaire (NLS *nuclear localisation sequences*) et deux séquences signal d'export nucléaire (NES *nuclear export sequences*). Wodrich et al. proposent que les deux séquences NLS2 et NES2 sont importantes dans la navigation de la protéine VI entre le noyau et le cytoplasme. La séquence NES2 est masquée lorsque la protéine VI est liée à l'hexon, ce

qui permet son import dans le noyau. Par la suite, lorsque l'hexon est dans le noyau la séquence NES2 redevient accessible et permet le passage de la protéine VI vers le cytoplasme. Les deux séquences NES2 et NLS2 sont protéolysées lors de la maturation du virus. La protéine VI a également un rôle dans la perturbation de la membrane de l'endosome après l'entrée dans la cellule lors du désassemblage du virus (Wiethoff et al., 2005). Wiethoff et ses collaborateurs ont montré que la protéine VI possédait une activité de lyse membranaire, permettant la libération du virus des endosomes. Cependant les changements conformationnels au sein de la capsid virale nécessaires pour permettre la libération de la protéine VI ne sont pas encore clairs.

d. La Protéine VIII

L'analyse du mutant thermolabile d'adénovirus suggère que la protéine VIII joue un rôle dans la stabilisation structurale du virion (Liu et al., 1985). Comme les protéines IIIa et VI, la protéine VIII est synthétisée sous forme de précurseur (24.7 kDa) donnant naissance à la protéine mature après clivage par la protéase virale. C'est une protéine de ~15 kDa (140 résidus) présente en 127 ± 3 copies par virion (Stewart et al., 1993) et qui n'a été localisée que du côté interne de la capsid (Everitt et al., 1975).

Le précurseur de la protéine VIII n'est détecté que dans les capsides vides mal formées, mais absent dans des particules virales entières. Cette propriété est d'ailleurs utilisée pour identifier les capsides vides lors de la préparation de vecteurs adénoviraux (Sutjipto et al., 2005; Takahashi et al., 2006; Vellekamp et al., 2001). La protéine VIII mature est riche en proline (8%) et en résidus basiques (arginine et lysine 11 %) donc prédite comme étant largement non repliée. Le repliement doit avoir lieu lorsqu'elle est en contact avec ses partenaires viraux. Dans notre laboratoire, la reconstruction par microscopie électronique à 10 Å de résolution d'HAd5 a été obtenue, ce qui a permis grâce à la comparaison avec un modèle quasi atomique de la capsid de l'adénovirus de localiser la protéine VIII du côté interne de la capsid (Fabry et al., 2005). Cette densité respecte le volume de la protéine VIII ainsi que le nombre de copies. Ainsi deux positions non équivalentes ont été attribuées à la protéine VIII, la première assure le contact entre le GON et les hexons péripentonaux autour de l'axe d'ordre 5, et la deuxième position se trouve au sein du GON autour de l'axe d'ordre 3 (Fig. 1.13). Une reconstruction 3D d'un adénovirus à 6 Å de résolution a été publiée en 2006 et confirme le positionnement des 120 protéines VIII dans la capsid (Saban et al., 2006).

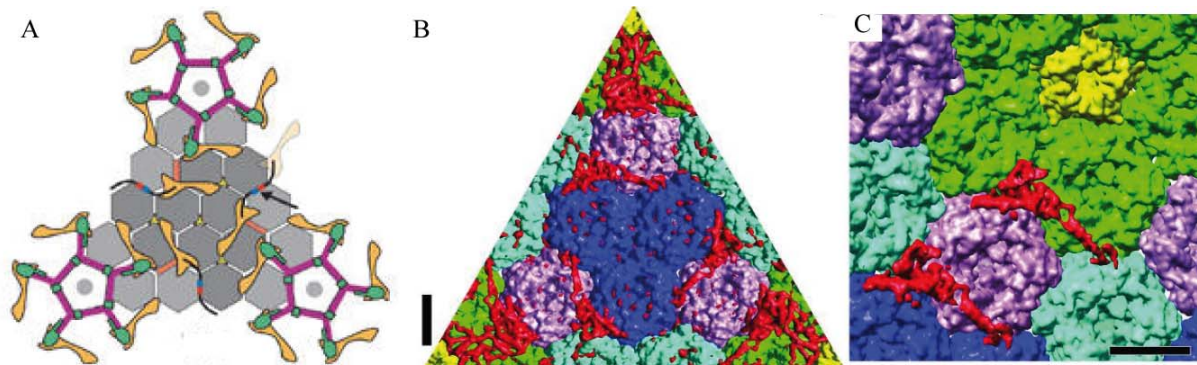


Figure 1.13: La position de la protéine VIII au sein de la capside.

A, Vue schématique du côté interne d'une facette de l'HAD5 (Fabry et al., 2005). En orange, la position attribuée à la protéine VIII, avec deux copies par unité asymétrique dans deux positions non équivalentes. Chaque copie de la protéine VIII interagit de manière similaire (à cette résolution) avec 4 hexons malgré des positions différentes dans la capside.

B, Vue interne d'une facette de l'HAD35 obtenu par cryoMET à 6 Å de résolution (Saban et al., 2006). En rouge la carte de différence entre le modèle quasi atomique et la reconstruction 3D de l'HAD35 et qui correspond aux protéines mineures. La densité rouge du centre est attribuée aux protéines VIII et les densités rouges vers les sommets de l'icosaèdre correspondent aux protéines VIII, aux parties N-terminales de la base du penton et aux protéines VI et IIIa.

C, Les densités identiques existantes dans la carte de différences représentée en B correspondant à la protéine VIII. Elles ont été isolées de manière informatique et représentées dans leur position au sein de la capside pour visualiser les zones d'interaction avec les hexons. La barre d'échelle, 50 Å.

1.3.4. Les Protéines du core

Le génome des adénovirus humains est constitué par un ADN double brin linéaire d'environ 36 kpb, enchâssé dans la capside. Cet ADN est encadré par deux séquences nécessaires à l'initiation de la réplication appelés ITRs (*inverted terminal repeats*). Le génome est associé avec 4 protéines structurales: la pV, pVII, pX (ou μ) et à la protéine de liaison à l'ADN: pTP (*terminal protein*), l'ensemble est appelé nucléoïde:

La Protéine TP

C'est une protéine de 55 kDa associée de façon covalente aux deux extrémités 5' du génome viral, elle forme un dimère ce qui a pour effet de circulariser le génome. Cette protéine est impliquée dans la réplication initiale de l'ADN (Mysiak et al., 2004; Rekosh et al., 1977).

La Protéine V

La protéine V (~42 kDa ; 368 aa) est spécifique au Mastadénovirus, elle est présente en 157 copies environ par virion (van Oostrum and Burnett, 1985). En plus de sa liaison à l'ADN, la protéine V est capable d'interagir avec la base de penton, la protéine mineure VI, la

nucléoprotéine VII et la protéine μ permettant ainsi une liaison entre la capsidite et le cœur du virion (Chatterjee et al., 1985; Everitt et al., 1975). Durant l'infection, la protéine V est impliquée dans le transport de l'ADN viral dans le noyau cellulaire (Matthews and Russell, 1998).

La Protéine VII

Avec plus de 800 copies par virion, la protéine VII est la protéine majeure du cœur du virion (Rux and Burnett, 2004). La protéine VII mature d'environ 17 kDa (174 aa) est produite à des phases tardives de la maturation après clivage du précurseur pVII de 23 acides aminés du côté N-terminal par la protéase virale. Durant la phase tardive de l'infection virale, la protéine VII se lie fortement à l'ADN non encapsidé. Par la suite, ce complexe sera encapsidé dans la particule virale et s'organisera comme la chromatine suivant un mécanisme encore inconnu (Zhang and Arcos, 2005).

La Protéine X (μ)

La protéine μ est la plus petite protéine du nucléoïde (~5 kDa ; 19 aa), son rôle est encore inconnu mais elle pourrait être impliquée dans l'encapsidation du génome viral. Il a été montré que durant l'import du génome viral à travers le pore nucléaire, les protéines VII et μ restent associées à l'ADN alors que la protéine V se détache avant ou immédiatement après la translocation du génome dans le noyau (Chatterjee et al., 1986).

Les protéines V, VII, et μ sont riches en arginine et ressemblent donc aux histones qui sont capables de neutraliser la charge du phosphate dans l'ADN permettant ainsi la condensation de l'ADN viral.

1.3.5. La Protéase de l'Adénovirus

La région L3 code pour une protéase de 23 kDa essentielle pour la maturation de la capsidite virale. Elle joue également un rôle dans le processus du désassemblage du virus au niveau de la membrane nucléaire lors de l'internalisation (Greber et al., 1996). La protéase est une protéine absolument requise pour la synthèse de virus infectieux. Elle existe en 10-30 copies par virion et permet le clivage de plusieurs précurseurs de protéines virales notamment la TP, μ , V, VI, VII et VIII (Hannan et al., 1983; Mirza and Weber, 1980). La protéase est activée dans la phase tardive d'infection après l'assemblage du virus, elle permet la maturation de la capsidite grâce au clivage des précurseurs. L'activité de la protéase augmente

en présence de l'ADN viral et du domaine C-terminal de la protéine pVI (Mangel et al., 2003; Mangel et al., 1993). Le précurseur pVI est à la fois un substrat et un cofacteur de la protéase. En effet, après clivage du pVI en protéine VI par la protéase, le domaine C-terminal de 11 acides aminés libéré devient le cofacteur de la protéase. La structure cristallographique de cette protéase en complexe avec ces 11 acides aminés a été déterminée (Ding et al., 1996). La structure de ce complexe est présentée dans la figure 1.14 (le brin β en rose représente le cofacteur).

La structure tridimensionnelle de cette protéase est unique, le site actif est constitué de trois acides aminés l'His45, Glu71 et Cys122 avec un arrangement similaire à celui de la papaïne suggérant un mécanisme catalytique semblable. Un rôle de régulateur de l'activité de la protéase par le cofacteur a été suggéré par Ding et al. (1996). La protéase est relativement inactive avant et pendant l'assemblage du virus, mais une fois le virion assemblé la protéase clive la protéine pVI et s'associe à son domaine C-terminal (pVIc). La protéase est aussi activée en présence d'ADN viral, celui-ci lui sert probablement de guide pour sa diffusion dans la cellule et ainsi le clivage d'autres précurseurs des protéines (Ding et al., 1996).

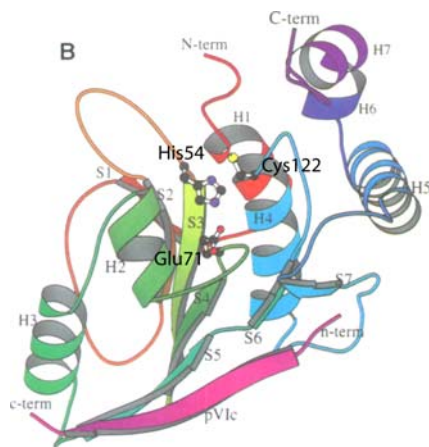


Figure 1.14: Structure cristallographique de la protéase de l'HAD2. Les 11 acides aminés représentant le cofacteur (en rose) ainsi que la triade catalytique sont indiqués.

Une trentaine d'autres protéines non structurales sont connues pour l'adénovirus. La majorité d'entre elles possède un rôle catalytique dans le cycle viral. Parmi elles la DBP (*DNA binding protein*) codée par l'unité de transcription E2, une phosphoprotéine de taille variable selon le sous-groupe (473-529 aa). Elle est multifonctionnelle mais surtout impliquée dans la réplication virale et différents processus du métabolisme de l'ADN viral (van Breukelen et al., 2003).

1.3.6. Structure et Organisation du Génome d'Adénovirus

L'organisation du génome de l'adénovirus est illustrée dans la figure 1.15. Le génome est organisé en plusieurs unités transcriptionnelles réparties dans les deux sens de lecture. Ces unités sont exprimées de façon différentielle durant le cycle viral. On distingue les gènes précoces (E pour *early*) et tardifs (L pour *late*) transcrits respectivement avant et après la réplication virale. Les gènes codant pour des protéines exprimées au même moment du cycle viral ont généralement des fonctions similaires et se trouvent sur la même unité de transcription. En général, la région E1A code pour deux protéines responsables d'une activité transcriptionnelle ainsi qu'une induction des cellules vers la phase S. La région E1B code pour deux protéines qui inhibent l'apoptose. La région E2 code pour trois protéines impliquées directement dans la réplication de l'ADN viral: la polymérase, la pTP et la DBP. La région E3 code pour des protéines impliquées dans l'évasion du virus vis-à-vis de la réponse immunitaire. La région E4 code pour des protéines impliquées dans la réplication, la transcription de l'ARNm, le transport membranaire, la traduction et la résistance à l'apoptose. Les gènes tardifs codent pour les protéines structurales qui permettent la construction de la capsid (Russell, 2000).

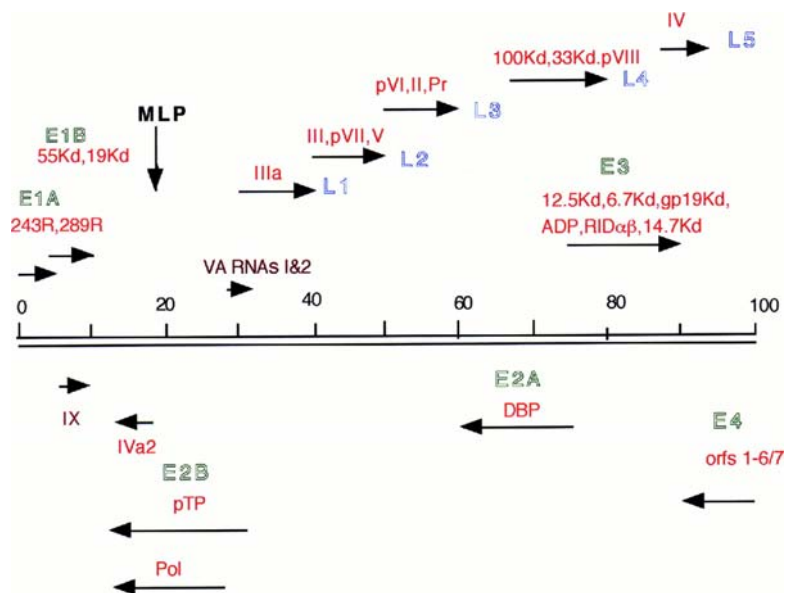


Figure 1.15: Illustration schématique de l'organisation des unités transcriptionnelles du génome de l'adénovirus. Le génome est représenté par une ligne horizontale noire. Les régions codantes sont indiquées par des flèches dans le sens de la transcription. En vert, les transcrits précoces (Early), les régions E1, E4 et E2 sont présentes dans tous les genres, alors que la région E3 est spécifique des Mastadénovirus et des Siadénovirus. En bleu les transcrits tardifs (Late, L1, L2, L3, L4 et L5). En marron la localisation du gène des VA RNAs. En noir le promoteur tardif majeur MLP (Major Late Promoter) (image (Russell, 2000)).

a- Mastadénovirus

Plusieurs génomes de Mastadénovirus ont été entièrement séquencés. Ils ont une taille intermédiaire qui varie entre 30.3 et 36.5 kpb avec une teneur en GC dans l'ADN de 40.8 à 63.8%. Les ITRs (Inverted Terminal Repeat) sont relativement longs (93 à 371 bp) et complexes par rapport aux autres genres, ils contiennent une variété de sites de liaison à des facteurs cellulaires. Les adénovirus humains du sous-groupe C sont les mieux étudiés de point de vue de l'organisation de leur génome et de la réplication virale. Cependant les différentes découvertes s'appliquent généralement à tous les Mastadénovirus excepté pour l'organisation des régions E3 et E4. Ces deux régions précoces sont différentes chez les adénovirus animaux. L'ADN des Mastadénovirus code pour des protéines spécifiques à ce genre: les protéines V et IX. Leur absence dans les autres genres n'est pour l'instant pas expliquée.

b- Aviadénovirus

Le génome des Aviadénovirus est le plus long parmi les adénovirus connus (20 à 45% de longueur en plus). Deux génomes ont été entièrement séquencés, le FAdV-1 et FAdV-9 (*Fowl adenovirus type 1 et 9*) et leur taille est respectivement de 43.8 et 45.1 kpb. La teneur en GC est de 53.8 à 59% et les séquences ITRs sont plus courtes que celles des Mastadénovirus (entre 54 et 71 pb). L'Aviadénovirus est caractérisé par l'omniprésence de deux fibres par penton. Les deux fibres du FAdV-1 (ou CELO) sont codées par deux gènes différents et possèdent deux longueurs de tiges différentes alors que chez les autres FAdVs la fibre est codée par un seul gène (Ojkic and Nagy, 2000). Les FAdV-1 (CELO), 9 et 10 sont largement étudiés en vue d'une utilisation potentielle comme vecteur en thérapie génique.

c- Atadénovirus

Les Atadénovirus infectent plusieurs classes de vertébrés (reptiles, oiseaux, marsupiaux et mammifères). La taille des génomes isolés et séquencés varie entre 29.6 (OAdV-7 ; *Ovine adenovirus type 7*) et 33.2 kpb (DAdV-1 ; *Duck adenovirus type 1*) avec un ITR de 46 (OAdV-7) à 59 kpb (BAdV-4 ; *Bovine adenovirus type 4*). Les Atadénovirus ont une teneur en GC relativement faible d'environ 34-43% (OAdV-7/DAdV-1). Le pourcentage élevé en AT qui caractérise ces adénovirus a donné naissance à leur nom de genre: les ATadénovirus.

La région centrale du génome des Atadénovirus est similaire à celle des Mastadénovirus, alors que les deux extrémités sont différentes. Plusieurs protéines sont

uniques aux Adénovirus, notamment la protéine structurale p32K qui serait présente dans tous les membres de ce genre étudiés jusqu'à présent et dont la fonction est encore méconnue (ÉIÓ et al., 2002). Aucune étude structurale n'a encore été effectuée sur ces virus car le virus n'a pas encore été produit en quantité suffisante.

d- Siadénovirus

Le genre Siadénovirus est constitué uniquement de deux membres pour l'instant, le FrAdV-1 (*Frog adenovirus type 1*) et le TAdV-3 (*Turkey adenovirus type 3*). Le génome de ces deux virus a été entièrement séquencé. Il représente le plus petit génome d'Ad connu aujourd'hui avec 26.2 et 26.3 kpb pour le FrAdV-1 et le TAdV-3 respectivement. Le pourcentage en GC est de 38 et 35% avec une taille des ITRs de 36 et 39 pb de longueur. Leur génome code pour une sialidase hypothétique d'où leur nom Siadénovirus.

1.4. Le Cycle Viral

La majorité des travaux menés sur les étapes de réplication des Ad sont réalisés sur les Ads humain du sous groupe C (HAd2, et HAd5). Cependant, les étapes clés qui ont été élucidées pour ce sous-groupe s'appliquent aux autres Ads. Les différentes étapes du cycle viral des Ads sont illustrées dans la figure 1.20.

Les deux premières étapes de l'infection d'une cellule par un adénovirus sont la reconnaissance et la fixation du virus à un type cellulaire spécifique, suivies par l'internalisation du virus. Ces deux étapes se font respectivement par le biais de récepteurs primaire et secondaire.

1.4.1. Les Récepteurs Primaires des Adénovirus

L'adsorption primaire du virus à la surface de la cellule se fait par l'intermédiaire d'une liaison de haute affinité entre la tête de la fibre et son récepteur cellulaire. Plusieurs récepteurs primaires ont déjà été identifiés:

a. Le CAR (Coxsackievirus and Adenovirus Receptor)

Il est connu depuis 1976 que les adénovirus humain du sous-groupe C et les coxsackievirus du groupe B ont un récepteur en commun (Lonberg-Holm et al., 1976). En 1997, la nature de ce récepteur a été déterminée pour l'HA_{D2} et l'HA_{D5} et appelé CAR (Bergelson et al., 1997). Plus tard, il a été montré que des sérotypes appartenant à cinq des six sous-groupes (A, C, D, E, F) sont capables de se lier au CAR et l'utilisent comme récepteur fonctionnel à la surface de la cellule (Roelvink et al., 1998). Cela n'empêche pas plusieurs sérotypes appartenant en particulier au sous-groupe D d'interagir avec d'autres récepteurs comme le résume le tableau 1.2.

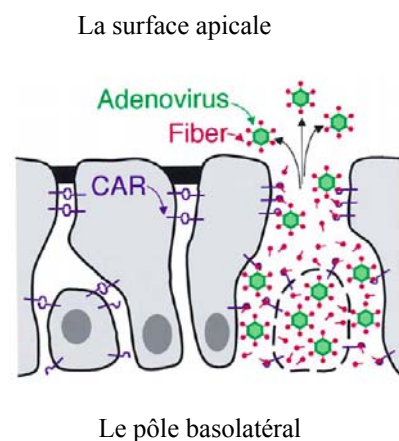
La protéine CAR est une protéine transmembranaire de type 1 appartenant à la famille des immunoglobulines (Ig). Cette protéine de 46 kDa est omniprésente à la surface des tissus cellulaires humains. On la retrouve notamment dans le cœur, le pancréas, le système nerveux, la prostate, les testicules, les poumons, le foie, la rate et les intestins, mais de manière plus rare dans les lymphocytes B et T, ainsi que dans les cellules musculaires adultes et plusieurs cellules malignes (Meier and Greber, 2004). La protéine CAR est constituée d'un domaine intracellulaire et d'un domaine extracellulaire séparés par un domaine transmembranaire. Les domaines transmembranaire et cytoplasmique ne sont pas nécessaires pour l'infection, ils constituent simplement le point d'ancrage de CAR dans les cellules (Wang et al., 1999). Le domaine extracellulaire est formé par deux domaines Ig, D1 et D2 (Meier and Greber, 2004), D1 étant le domaine de liaison à l'adénovirus *via* la tête de la fibre. La protéine CAR joue un rôle très important dans les cellules en assurant les jonctions intercellulaires au niveau du pôle basolatéral des cellules épithéliales. Elle co-immunoprécipite avec des protéines du complexe de la jonction apicale (ZO-1 et β catenin). Cette fonction d'adhésion est accomplie par une dimérisation entre deux molécules situées entre deux cellules voisines (Fig. 1.16).

La structure atomique du domaine D1 de CAR a été déterminée par cristallographie et RMN, seule ou en complexe avec les têtes de fibre de l'HA_{D12} (Bewley et al., 1999), de l'HA_{D37} et enfin de l'adénovirus canin de sérotype 2 (CAV-2) (Seiradake et al., 2006). Ces structures cristallographiques montrent que trois monomères D1 interagissent avec une tête de fibre trimérique. Le site d'interaction de la tête des fibres avec le CAR est toujours localisé dans la même zone malgré le faible taux de conservation de la séquence. Le site de liaison est situé majoritairement dans les parties flexibles de la tête de fibres. Chaque monomère de la protéine CAR interagit principalement avec les boucles AB, DE et le feuillet β F d'un monomère de la tête de la fibre et la boucle FG du monomère adjacent (Fig. 1.17). L'analyse

de mutants de la tête de fibre d'HA5 a montré le rôle majeur des boucles AB et DE dans la liaison avec la protéine CAR (Kirby et al., 2000). L'affinité de la tête pour la protéine CAR va également varier suivant la composition de ces boucles: deux acides aminés présents dans la tête de l'HA5 sont absents chez l'HA12: (la Ser⁴⁰⁸ dans la boucle AB et la Tyr⁴⁷⁷ dans la boucle DE) ce qui entraîne une baisse d'affinité d'un facteur 8 (Howitt et al., 2003).

Le rôle de la protéine CAR dans le cycle de vie de l'Ad est encore sujet à caution. En effet, il semblerait que CAR ne soit pas uniquement un récepteur cellulaire pour ce virus. Il a été montré qu'après infection d'un épithélium respiratoire cilié humain par le pôle basolatéral à l'aide d'un adénovirus, les virus néo-synthétisés pouvaient se retrouver à la surface de cet épithélium (Walters et al., 2002). Les cellules de surface de l'épithélium ne sont pas compétentes vis à vis de l'adénovirus. La présence de virus à la surface de l'épithélium est donc très étonnante puisque les particules virales libérées des cellules basales ne peuvent théoriquement pas accéder à la surface en raison de l'étanchéité de cet épithélium, assuré par les jonctions cellulaires. Des études de cristallographie et de biochimie suggèrent que la tête de la fibre des Ads, qui possède une forte affinité pour la protéine CAR, empêche la dimérisation de celui-ci en s'y liant (Bewley et al., 1999; van Raaij et al., 2000). La tête de la fibre de l'adénovirus peut donc perturber les interactions CAR-CAR ce qui a pour effet d'augmenter la perméabilité de l'épithélium (Walters et al., 2002) (Fig.1.16). La tête de la fibre de l'adénovirus joue probablement en plus de son rôle de reconnaissance cellulaire un rôle dans le passage des virus au travers des barrières physiques existantes dans les tissus (Walters et al., 2002). Cette hypothèse est confirmée pour l'HA37. Ce virus n'utilise pas la protéine CAR comme récepteur, mais la tête de la fibre s'y attache tout de même avec une grande affinité, probablement pour permettre aux virus néo-synthétisés de franchir la barrière épithéliale pour infecter d'autres cellules (Seiradake et al., 2006).

Figure 1.16: Modèle de mécanisme de passage des adénovirus vers la surface apicale de l'épithélium : l'adénovirus entier ou les fibres de celui-ci pourraient rompre les contacts intercellulaires dus à la protéine CAR et ainsi diminuer l'étanchéité de la barrière épithéliale (Walters et al., 2002).



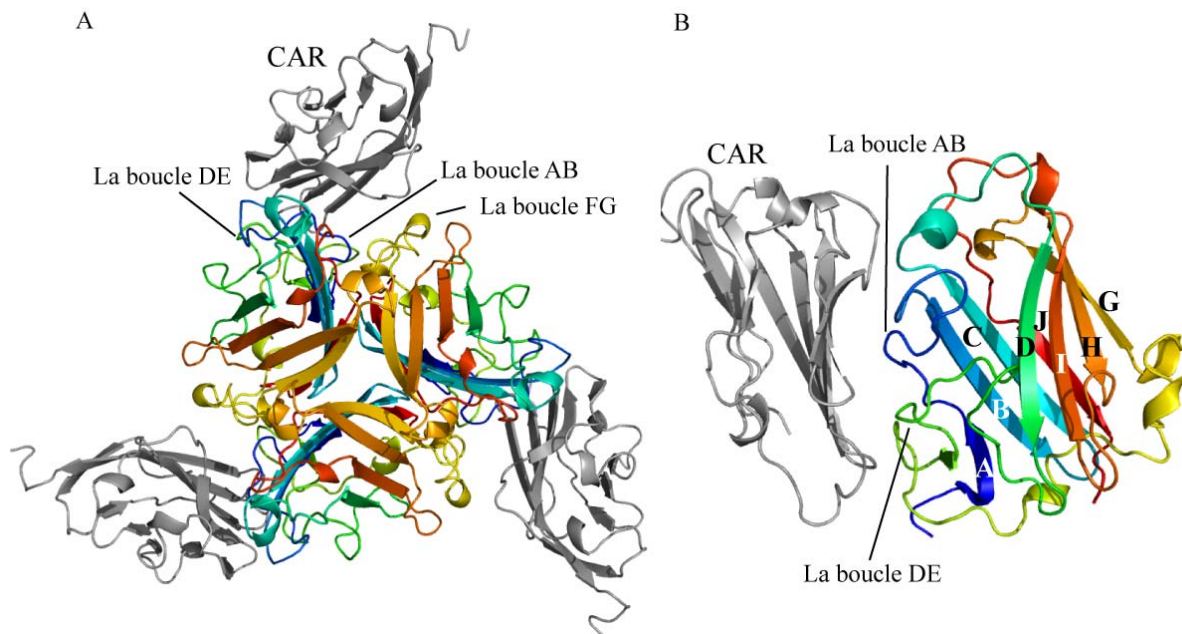


Figure 1.17: La structure de la tête de la fibre d’HAd12 en complexe avec le domaine D1 du CAR.
A, La tête de la fibre est colorée du bleu au rouge depuis la partie N-terminale jusqu’à la partie C-terminale et la protéine CAR est représentée en gris.
B, L’interface d’interaction entre un monomère de la protéine CAR et un monomère de la tête de la fibre de l’HAd12. La protéine CAR interagit principalement avec les boucles AB et DE de la tête de la fibre.

D’autres exemples d’interaction entre un virus et des protéines d’adhésion intercellulaire existent ; le réovirus qui reconnaît la protéine JAM1 (*Junctional Adhesion Molecule*) (Barton et al., 2001b) ; le rhinovirus qui interagit avec la molécule d’adhésion ICAM-1 (*Intercellular Adhesion Molecule 1*) (Greve et al., 1989; Staunton et al., 1989) et l’alphaherpesvirus qui se lie à la protéine Prr1 (*Poliovirus receptor-related*) (Geraghty et al., 1998).

Quelques Ads, comme l’HAd41 ou le virus CELO, possèdent deux fibres, une longue et une courte (Favier et al., 2002; Pieniazek et al., 1990). La fibre longue interagit avec la protéine CAR, alors que le récepteur de la fibre courte est encore inconnu (Seiradake and Cusack, 2005). Cette dernière pourrait être responsable du tropisme de l’HAd41 pour le tube digestif (Allard and Wadell, 1988; Croyle et al., 1998).

b. L’Acide Sialique

Plusieurs sérotypes d’Ads appartenant au sous-groupe D utilisent l’acide sialique plutôt que la protéine CAR comme récepteur (HAd8, HAd19a et HAd37). Tous ces virus sont associés avec des kératoconjunctivites épidémiques. Du point de vue structural et

biophysique, leurs têtes de fibres sont caractérisées par des points isoélectriques élevés ($pI=9.1$) (Arnberg et al., 1997). Le point isoélectrique de l'acide sialique étant bas ($pI=2.3$), l'interaction qui existe entre les deux est très probablement charge dépendante. Les structures cristallographiques des têtes de fibre des HAd37 et HAd19p en complexe avec l'acide sialique ont été déterminées par Burmeister et al. (2004). Les structures de ces complexes confirment que le site d'interaction de l'acide sialique sur la tête de la fibre est effectivement chargé positivement. Le site de liaison localisé dans la partie apicale de la tête près du centre du trimère est stabilisé par des liaisons hydrogènes mettant en jeu des résidus conservés chez les HAdS du sous-groupe D (la Tyr³¹² et la Lys³⁴⁵ interagissant avec l'acide sialique, et la Tyr³⁰⁸, Pro³¹⁷ et la Val³²² se trouvent dans le site actif) (Fig. 1.18).

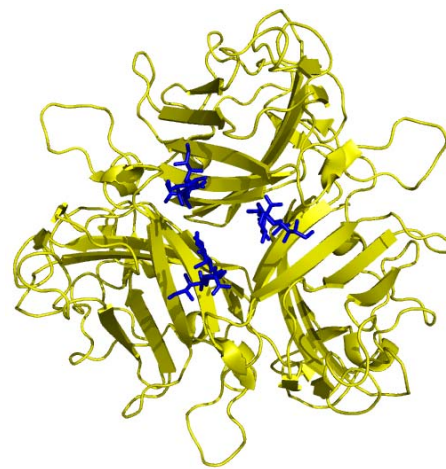


Figure 1.18: Représentation de la tête de fibre d'HAd37 (en jaune) en complexe avec l'acide sialique (en bleu) vue de dessus.

Contrairement à l'HAd37, l'HAd19p se fixe à l'acide sialique mais ne l'utilise pas comme récepteur. Seuls deux acides aminés sont différents entre les deux têtes de fibre: les Lys²⁴⁰ et Asn³⁴⁰ sont remplacés par une Glu et Asp dans l'HAd19p. La mutation Lys²⁴⁰Glu est la plus importante car elle est suffisante pour empêcher la liaison de l'HAd19p aux cellules du tissu conjonctif (Huang et al., 1999). Ces deux mutations ponctuelles se traduisent par une baisse du pI (8.8) et probablement par un changement de surface électrostatique (Arnberg et al., 2002) ce qui entraîne une légère baisse d'affinité de l'acide sialique pour la tête de la fibre d'HAd19p ($k_d=5mM$ pour l'HAd37 et $7mM$ pour l'HAd19p) (Burmeister et al., 2004). Toutefois ces changements sont trop mineurs pour expliquer l'incapacité de l'HAd19p à utiliser ce récepteur et la différence de tropisme entre les deux virus. Il est envisageable qu'*in vivo*, l'HAd19p se lie soit à un acide sialique modifié plus complexe soit utilise en plus un co-récepteur qui interagit avec une autre protéine de la capsid (Burmeister et al., 2004).

D'autres virus utilisent l'acide sialique comme récepteur, notamment les orthomyxovirus, paramyxovirus, picornavirus, papovavirus, coronavirus, réovirus et parvovirus (Alexander and Dimock, 2002; Barton et al., 2001a; Bauer et al., 1999; Cahan and

Paulson, 1980; Jin et al., 1994; Kaludov et al., 2001; Krempl et al., 2000; Shah and Lipton, 2002; Stoner et al., 1973; Suzuki et al., 2001; Suzuki et al., 2000; Uncapher et al., 1991; Utagawa et al., 1982; Zhou et al., 2000). L'affinité de ces virus pour l'acide sialique est comparable à celle des adénovirus (de l'ordre du mM).

c. Le CD46

Les adénovirus du sous-groupe B n'ont pas le même récepteur que les autres Ads : ces virus ne rentrent en effet pas en compétition avec les Ads du sous-groupe C lors de l'attachement aux cellules (Stevenson et al., 1995). De nombreuses études ont été nécessaires pour découvrir le récepteur cellulaire des virus appartenant à ce sous-groupe. En 2003, le CD46 a été identifié comme récepteur cellulaire pour l'HAd11 (sous-groupe B) (Segerman et al., 2003). Depuis, il a été montré que plusieurs sérotypes du sous-groupe B cités dans le tableau 1.2 se lient ou utilisent la protéine CD46 comme récepteur pour se fixer à la cellule hôte (Gaggar et al., 2003; Marttila et al., 2005; Sirena et al., 2004). La protéine CD46 est une glycoprotéine transmembranaire exprimée dans la majorité des tissus cellulaires humains. Sa partie extracellulaire (N-terminale) est constituée par 4 répétitions du domaine SCRs (SCR1-SCR4 ; *Short Consensus Repeats*). Les domaines SCR1 et SCR2 sont nécessaires et suffisants pour une liaison optimale des l'HAd3, 7, 11 et 35 à la surface de la cellule (Fleischli et al., 2007; Fleischli et al., 2005).

La structure cristallographique des domaines SCR1 et SCR2 du CD46 a été déterminée seule (Casasnovas et al., 1999) et en complexe avec la tête de la fibre d'HAd11 (Persson et al., 2007). Le nombre de contacts et la surface impliquée dans cette interaction indique une forte affinité (2 nM) (Persson et al., 2007). Dans la structure du complexe (Fig. 1.19), les deux domaines de CD46 adoptent une conformation linéaire en se liant sur le côté de la tête de la fibre « à cheval » sur deux des trois monomères. La partie SCR1 et l'interface SCR1-SCR2 interagissent avec un monomère de la fibre, alors que le domaine SCR2 va interagir avec le monomère adjacent. Les interactions entre les deux constituants du complexe font intervenir des liaisons de type *Van Der Waals* (pour SCR1) et des liaisons hydrogènes (pour SCR2). L'interaction entre la fibre et l'interface SCR1-SCR2 impliquant les boucles HI et DG de la fibre semble très importante: une mutation ponctuelle au niveau de la tête de la fibre (Arg279Gln) inhibe totalement l'interaction fibre-CD46 (Gustafsson et al., 2006). Il a également été montré que la liaison de CD46 à la tête de la fibre d'HAd11 induisait des changements conformationnels importants sur les domaines SCR1 et SCR2 du CD46.

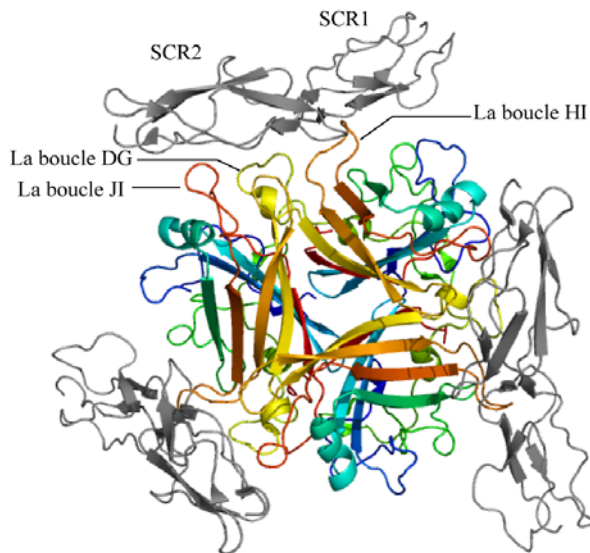


Figure 1.19: La structure cristallographique du complexe tête de fibre d'HA11 (coloré) et son récepteur CD46 (gris).

La protéine CD46, comme l'acide sialique et la protéine CAR, est également utilisée comme récepteur par de nombreux pathogènes humains (virus et bactéries) comme la rougeole (Naniche et al., 1993), l'herpès (Greenstone et al., 2002) et le streptocoque (Giannakis et al., 2002). Les zones d'interaction entre la protéine CD46 et son site de reconnaissance varient suivant les pathogènes.

Il a été montré que l'HA37 interagissait aussi avec le CD46 (Wu et al., 2004).

d. Les CD80 et CD86

Les protéines CD80 et CD86 appartiennent à la superfamille des Igs. Elles sont connues pour leur fonction d'activation des lymphocytes T (Greenwald et al., 2005). Des travaux récents ont montré que l'HA3 interagissait avec les protéines CD80 et CD86 (Short et al., 2004) mais pas avec la protéine CD46. Les auteurs suggèrent qu'il existe probablement un autre récepteur pour l'HA3, au moins dans les cellules Hela, car une neutralisation de CD80 et CD86 par des anticorps ne réduit l'infection que de 50%. D'autres études généralisent les protéines CD80 et CD86 comme étant des récepteurs pour les adénovirus appartenant au sous-groupe B (Short et al., 2006).

e. Le MHC-1 (Major histocompatibility complex class I)

Le domaine C-terminal $\alpha 2$ du complexe majeur d'histocompatibilité de classe 1 MHC-1 est reconnu par les HA2 et HA5 et sert de récepteur dans les cellules Hela et Daudi exprimant cette protéine (Hong et al., 1997). Cependant, d'autres études ont montré que

l'HA5 ne peut pas se lier à des cellules d'hamster exprimant la protéine MHC-1 (Davison et al., 1999). Encore une fois, les données sont contradictoires, d'autres données sont nécessaires pour éclaircir le rôle exact de MHC-1 dans l'infection virale par les HA2 et 5.

f. *Le VCAM-1 (Vascular cell adhesion molecule 1)*

Egalement connue sous le nom de CD106, cette protéine joue un rôle important dans le système de réponse immunitaire. Comme la protéine CAR, la protéine VCAM-1 est une immunoglobuline qui peut être reconnue par l'HA5 (Chu et al., 2001).

g. *Les Héparanes Sulfates*

Les héparanes sulfates (HS) sont des polysaccharides sulfatés exprimés de façon ubiquitaire à la surface des cellules de mammifères et dans les matrices extracellulaires. Ces molécules jouent un rôle central dans de nombreux processus biologiques tels que la prolifération cellulaire, l'adhésion cellulaire, la chémo-attraction, l'inflammation, la coagulation, l'assemblage de matrice, le développement embryonnaire, etc. Les héparanes sulfates sont liés à un cœur protéique et forment alors des protéoglycanes (HSPGs), le domaine protéique de liaison aux HS peut être utilisé par un grand nombre de micro-organismes pathogènes, en particulier les virus, pour s'ancrer à la surface cellulaire. L'HA2, l'HA5 interagissent avec les héparanes sulfates (Dehecchi et al., 2001; Dehecchi et al., 2000), cette interaction se fait probablement *via* la tige de la fibre (Smith et al., 2003). La base de penton d'HA3 interagit avec les HS, toutefois, l'HA3 entier ne semble pas s'y fixer (Fender et al., 2008; Vives et al., 2004).

h. *Les Intégrines $\alpha M\beta 2$ et $\alpha L\beta 2$*

L'attachement du virus à la surface de la cellule peut parfois dépendre directement d'une interaction entre la base de penton et les intégrines. L'HA2 dépourvu de fibre est capable d'infecter des monocytes qui n'expriment pas de protéine CAR. L'infection est provoquée par un attachement primaire aux intégrines $\alpha M\beta 2$ et $\alpha L\beta 2$ suivi par une interaction avec l'intégrine α_v nécessaire pour l'internalisation du virion (Huang et al., 1996).

La diversité des récepteurs utilisés par les Ads pour se fixer aux cellules, permet l'infection de plusieurs types de tissus cellulaires. Le travail d'identification de ces récepteurs n'est pas encore terminé, d'autres récepteurs d'adénovirus restent probablement à identifier.

Tableau 1.2: Récapitulatif des récepteurs connus pour les HAds. Les complexes récepteur-tête de fibre dont la structure cristallographique est connue sont indiqués en rouge. Les sérotypes dont la structure cristallographique de la tête de fibre est connue sont soulignés.

Sous groupe	Sérotipe	Récepteur	Références	Longueur de la fibre (nm)	Tropisme
A	<u>12</u> ; 31	CAR	(Bewley et al., 1999)	28-31	Intestinal
	18	?			
B1	<u>3</u>	CD46 CD80/86	(Fleischli et al., 2007) (Short et al., 2004)	9-11	Respiratoire Oculaire Rénale
	16 ; 21 ; 50	CD46/80/86			
	7	CD46?			
B2	<u>11</u> ; <u>35</u>	CD46	(Persson et al., 2007) (Fleischli et al., 2007)		
	14 ; 34	CD46/80/86			
C	<u>2</u> ; <u>5</u>	CAR, Héparane sulfate MHC-1	(Bergelson et al., 1997) (Dechecchi et al., 2000) (Dechecchi et al., 2001) (Hong et al., 1997)	23-31	Respiratoire lymphatique
	1 ; 6	?			
D	8 ; 19a	Acide sialique		12-13	Oculaire Intestinal
	15 ; <u>19p</u>	Acide sialique CAR	(Burmeister et al., 2004)		
	<u>37</u>	Acide sialique	(Burmeister et al., 2004) (Seiradake et al., 2006)		
	9 10 13 17 20 22-30 32 33 36 38 39 42-49 51	CD46 ?	(Wu et al., 2001)		
E	4	CAR		17	Respiratoire Oculaire
F	<u>41</u>	CAR (longue)		33	Intestinal
	40	?		28	

1.4.2. Fixation au Récepteur Secondaire et Internalisation

L'attachement de l'Ad à la surface de la cellule *via* la tête de fibre est suivi par une interaction entre les intégrines et la base de penton (Mathias et al., 1994; Wickham et al., 1993). L'assemblage d'intégrines à la surface de la cellule induit une multitude de signaux facilitant l'internalisation du virus par la voie d'endocytose (Greber, 2002; Meier and Greber, 2004). La majorité des bases de penton des adénovirus interagissent avec les intégrines *via* le motif RGD (Arg-Gly-Asp). Parmi ces intégrines, nous retrouvons $\alpha_v\beta_3$, $\alpha_v\beta_5$ (Wickham et al., 1993), $\alpha_v\beta_1$ (Li et al., 2001), $\alpha_5\beta_1$ (Davison et al., 1997), $\alpha M\beta_2$ et $\alpha L\beta_1$ (Huang et al.,

1996). Les intégrines $\alpha_v\beta_3$ et $\alpha_v\beta_5$ sont capables de déclencher l'endocytose, $\alpha_v\beta_5$ permet une meilleure perméabilité de la membrane et par conséquent la libération des particules de l'endosome (Wang et al., 2000; Wickham et al., 1994).

L'interaction base de penton-intégrine active de nombreuses voies de signalisation cellulaire. Ainsi, l'activation de la voie des kinases PI-3K (*phosphoinositide-3-OH kinase*) déclenchant la famille des GTPase Rho, la polymérisation et la réorganisation de l'actine pour faciliter l'endocytose (Li et al., 1998; Meier and Greber, 2004). L'endocytose des vésicules recouvertes de clathrines est la voie majoritaire d'entrée des Ads dans la cellule. Cependant, il est évident qu'au moins une autre voie existe. Meier et al. (2002) ont montré que la liaison de l'HA2 à la surface des cellules épithéliales déclenchait une macropinocytose de manière simultanée avec l'endocytose (Meier et al., 2002). La macropinocytose nécessite l'activation des intégrines α_v , la F-actine, la protéine kinase C (PKC) et les petites protéines G de la famille Rho (GTPases).

Le processus de déshabillage séquentiel du virus commence par la dissociation de la fibre de la capsid et finit par la pénétration de l'ADN dans le noyau (Greber et al., 1993; Nakano et al., 2000). La séquence exacte de déshabillage du virus n'est pas connue mais celle-ci débute en même temps que l'acidification de l'endosome avec le relargage de la fibre. Le virus partiellement déshabillé est ensuite transporté jusqu'au pore nucléaire grâce à l'interaction avec des protéines motrices de type dynéine, circulant sur le réseau de microtubules (Greber and Way, 2006; Greber et al., 1993; Nakano et al., 2000). Les virus vont alors s'accumuler au niveau de la membrane nucléaire (Leopold et al., 1998; Suomalainen et al., 1999) où des étapes supplémentaires de déshabillage se produisent (Greber et al., 1997) permettant ainsi l'injection de l'ADN dans le noyau. L'étape de translocation du virus vers le noyau est rapide et elle survient 30 à 60 minutes après l'infection de la cellule. Les premières étapes de transcription peuvent alors commencer. Peu de temps après la pénétration de l'ADN dans le noyau, les premiers ARNm apparaissent dans le cytoplasme (Fig. 1.20).

Le cycle d'infection par l'Ad se déroule en deux phases. La phase précoce, avant la réplication de l'ADN, est caractérisée par le passage du génome dans le noyau, suivi par une expression sélective des gènes précoces. La phase précoce prépare la cellule à la réplication de l'ADN viral permettant ainsi la transcription et la traduction des gènes tardifs. La phase tardive, après la réplication de l'ADN, est marquée par la synthèse des protéines structurales. Enfin, l'assemblage des protéines structurales et la maturation du virus ont lieu dans le noyau.

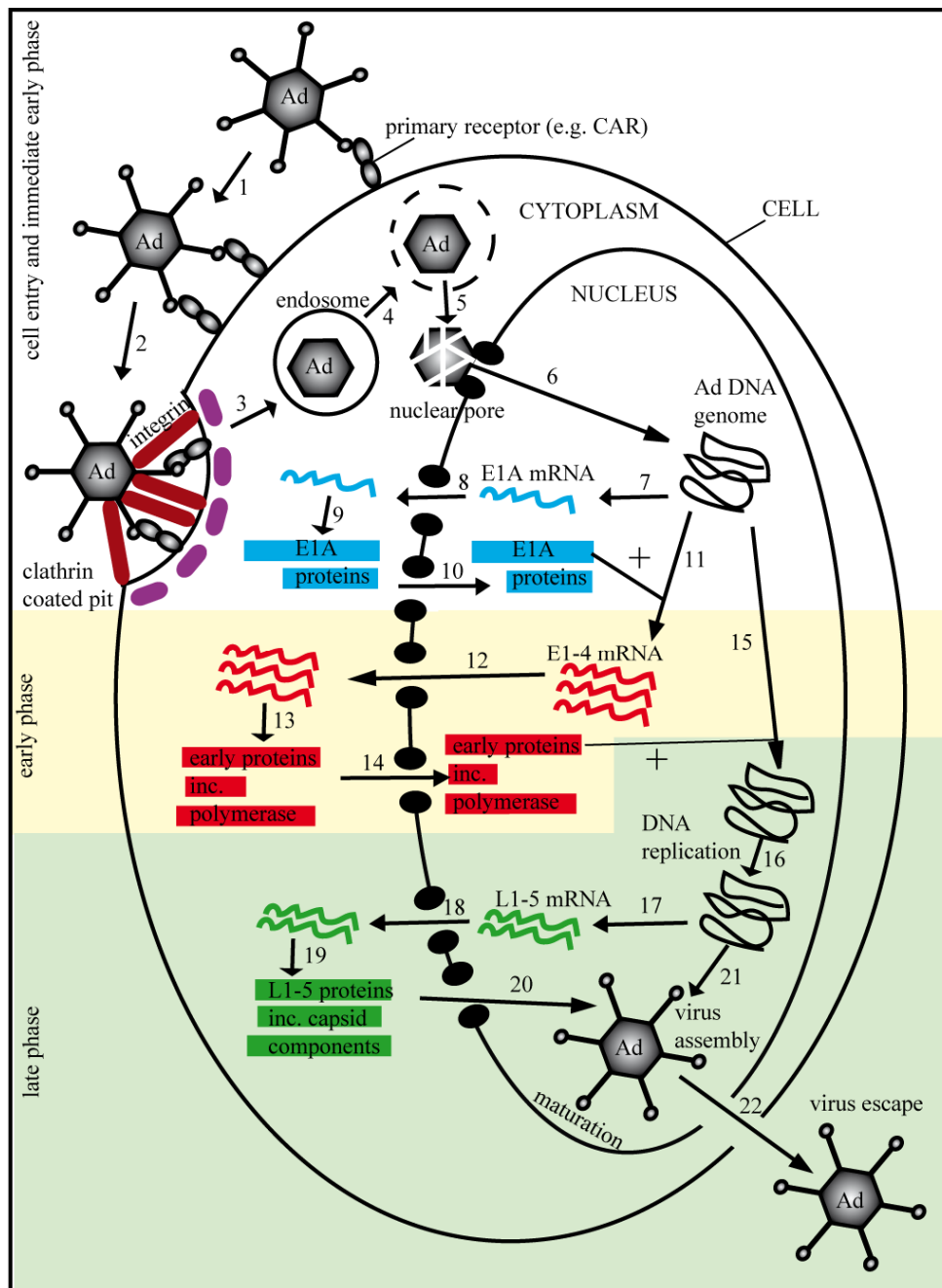


Figure 1.20: Le cycle viral de l'adénovirus basé sur l'étude des HAd2 et HAd5. Les trois couleurs indiquent les phases : très précoce (blanc), précoce (jaune) et tardive (vert). Les flèches sont numérotées et représentent les différentes étapes clés du cycle d'infection: 1. L'attachement de l'Ad au récepteur primaire. 2. L'attachement au récepteur secondaire (intégrines) et l'internalisation par endocytose dans une vésicule à clathrines. 3. L'internalisation dans l'endosome précoce. 4. Relargage de la capside de l'endosome et perte de certaines protéines. 5. Transport vers le pore nucléaire via les microtubules. 6. le génome est injecté dans le noyau. 7. Transcription de la région E1A. 8. Export des ARNm de E1A dans le cytosol. 9. Traduction des ARNm de E1A. 10. Transport des protéines de E1A dans le noyau. 11. Ces protéines induisent la transcription des gènes précoces. 12. Export des ARNm précoces. 13. Traduction des protéines précoces. 14. Import des protéines précoces. 15 et 16. Induction de la réplication virale par les protéines précoces. 17. Transcription des gènes tardifs. 18. Export des ARNm tardifs. 19. Traduction des protéines majeures (surtout les protéines de la capside). 20. Import des protéines majeures et assemblage de la capside dans le noyau. 21. Encapsidation de l'ADN viral. 22. Maturation et relargage de l'Ad. (Le cycle viral est tiré de la thèse d'Elena Seiradake avec son accord).

1.4.3. La Phase Précoce d'Infection

Pendant la phase précoce, les protéines nécessaires à la réplication du virus et à la modulation de la réponse immune de l'hôte sont produites. Toutes ces protéines se trouvent sur des unités de transcription précoces appelées E1, E2, E3 et E4 (E pour *early*).

Après pénétration de l'ADN viral dans le noyau, la phase précoce débute par l'expression du gène très précoce E1A. Les protéines E1A proviennent d'un épissage alternatif du gène précoce E1. E1A comporte deux protéines majeures: 13S (ou 289R) et 12S (ou 243R). Ces protéines sont nécessaires à la modulation du métabolisme de la cellule dans le but de le rendre plus favorable à la réplication virale (elles inhibent la multiplication et la différenciation cellulaire) (Berk, 1986). Les deux ARNm codent pour deux transactivateurs capables de stimuler la transcription des gènes viraux précoces impliqués dans la réplication du génome viral (E1B, E2, E3 et E4) (Brockmann and Esche, 2003), mais aussi de gènes cellulaires impliqués dans la prolifération cellulaire (Zhao et al., 2003). La synthèse des protéines de E1A induit également l'apoptose de la cellule en activant la protéine p53. Pour éviter une mort cellulaire précoce par apoptose (et donc un arrêt du cycle viral) deux protéines codées par la région précoce E1B, 19K et 55K, sont synthétisées. La protéine E1B-19K, un équivalent fonctionnel d'un régulateur anti-apoptotique *Bcl-2*, est capable d'inhiber plusieurs voies d'apoptose (Chiou et al., 1994). La protéine E1B-55K, en complexe avec une autre protéine précoce (E4-ORF6), dirige la protéine p53 vers le protéasome pour une dégradation (Harada et al., 2002). La protéine E1B-19K est inhibée tardivement par la protéine d'E3 qui joue un rôle dans la destruction de la cellule.

Chez les Mastadénovirus, la région E3 code pour 5 à 9 protéines selon les espèces. Toutes ces protéines permettent à la cellule infectée d'échapper au système immunitaire inné (Wold and Gooding, 1989). Un des produits du gène E3 est la protéine ADP (*adenovirus death protein*) qui facilite la lyse cellulaire lors de la libération des virus (Tollefson et al., 1996).

La région E4 code pour au moins 7 protéines différentes, parmi lesquelles les protéines ORF3 et ORF6 qui sont les plus étudiées. Ces deux protéines forment avec la protéine E1B-55K un complexe qui est impliqué dans la perturbation et la réorganisation des structures de la matrice nucléaire nommée PML (*Promyelocytic Leukemia*) (Evans and Hearing, 2003). Le complexe E4-ORF6/E1B-55K est localisé dans le site de réplication de l'Ad où il recrute différents composants associés au PML impliqués dans différentes fonctions (transcription, réparation d'ADN, réplication, ...) (Doucas et al., 1996). Le complexe E4-ORF6/E1B-55K

induit l'expression des gènes tardifs ainsi que le transport nucléaire des ARNm tardifs (Flint and Gonzalez, 2003).

À ce stade du cycle viral, l'Ad transcrit aussi un ensemble d'ARNs appelés VA RNAs. Ces ARNs, transcrits par l'ARN polymérase de type III, ne sont pas traduits mais jouent un rôle dans le combat contre les systèmes de défense cellulaires et dans la traduction efficace de l'ARNm viral (Russell, 2000).

La réplication de l'ADN viral qui commence 6 à 8 heures après le début de l'infection, est nécessaire pour marquer la transition entre les phases précoce et tardive de l'infection. Les protéines nécessaires pour la réplication virale sont codées par la région précoce E2 du génome viral: la polymérase, le précurseur de la protéine TP (pTP *terminal protein*) et la DBP (*DNA binding protein*). La réplication est initiée au niveau des régions ITRs à chaque extrémité du génome qui servent d'origine de réplication. La protéine pTP, se lie à la polymérase et sert d'amorce pour la réplication. Lors de l'élongation, la protéine DBP recouvre l'ADN monocaténaire néo-synthétisé et le protège contre l'activité des nucléases cellulaires. Cette protéine aide aussi le défilement de l'ADN double brin lors de l'élongation (Dekker et al., 1997). La maturation de la protéine pTP après digestion par la protéase virale donne naissance à la protéine TP liée de façon covalente aux extrémités 5' de l'ADN bicaténaire.

1.4.4. La Phase Tardive d'Infection

Les gènes tardifs sont regroupés sous un seul opéron composé de 5 régions: L1, L2, L3, L4 et L5. La transcription de ces gènes est contrôlée par le promoteur majeur tardif (MLP, pour *Major Late Promoter*). Le regroupement de plusieurs gènes sous la régulation d'un seul promoteur permet une expression simultanée de toutes les protéines tardives.

Deux gènes IVa2 et IX sont exprimés lorsque la synthèse de l'ADN viral débute (Binger and Flint, 1984), elles auraient un rôle dans l'activation des gènes tardifs. Le début de la phase tardive est marqué par l'activation du promoteur MLP (12 heures après infection). L'unité de transcription tardive majeure (MLTU *major late transcription unit*) code pour les protéines de la capsid (sauf la IX) ainsi que pour d'autres protéines nécessaires pour l'assemblage, l'encapsidation et la maturation du virion : 100K, 52K, 33K, 22K et la protéase virale. Les protéines structurales sont importées dans le noyau, où l'assemblage de la procapsid a lieu (Perez-Romero et al., 2006). L'assemblage est achevé par l'encapsidation du génome et le positionnement du penton. La capsid subira une étape de maturation par

protéolyse des protéines structurales par la protéase virale encapsidée. Ces événements sont accompagnés par des changements majeurs au niveau de l'infrastructure du noyau et de la perméabilité de la membrane nucléaire (Rao et al., 1996). Ceci facilite la progression du virus depuis le noyau vers le cytoplasme. La désintégration de la membrane plasmique et le relargage des particules virales dans le milieu extracellulaire (environ 10^4 virus par cellule) sont contrôlés par la protéine ADP (Tollefson et al., 1996) et constituent les étapes finales du cycle viral.

1.5. L'Intérêt des Adénovirus pour la Thérapie Génique

La thérapie génique vise à introduire un ou plusieurs gènes dans certaines cellules d'un patient afin de compléter ou de remplacer un gène défectueux. La thérapie génique est un terme général qui englobe plusieurs méthodes et stratégies thérapeutiques. Actuellement, plus de 1300 essais cliniques auraient été réalisés dans le monde, 70% de ces essais se sont focalisés sur le traitement du cancer, environ 20% sur le traitement de maladies héréditaires classiques et 10% sur des infections diverses comme les infections virales (Gardlik et al., 2005). Les vecteurs viraux sont utilisés dans environ 70% des essais cliniques pour délivrer des gènes. L'adénovirus (~25% des essais) et le rétrovirus (~23%) sont les vecteurs viraux les plus utilisés dans les essais. D'autres méthodes de transfert de gènes ont été élaborées pour ne pas recourir aux virus: injection directe d'ADN, *lipofection* (le matériel génétique est contenu dans des liposomes), *gene-gun* et l'électroporation *in vivo* (Gardlik et al., 2005). A l'inverse des vecteurs viraux, ils sont plus faciles à produire, à manipuler et à stocker, mais leur efficacité est bien moindre (<http://www.wiley.co.uk/genmed/clinical/>).

Les Ads sont relativement faciles à produire en grande quantité, ils infectent de nombreux types cellulaires quiescents ou en division et permettent une expression efficace des gènes dans le noyau. Ceci explique leur utilisation fréquente pour le développement des vecteurs de transfert de gènes. Leur génome ne s'intègre pas (dans la plupart des cas) à celui de la cellule hôte ce qui a l'avantage de minimiser le risque de transformation, mais entraîne aussi la perte d'expression du transgène dans des cellules en prolifération. Une expression à long terme nécessitera donc, soit des traitements répétitifs avec le vecteur, soit le ciblage des cellules post-mitotiques.

De nombreux avantages ont été cités quant à l'utilisation des adénovirus en thérapie génique, la limitation majeure dans de ces vecteurs reste la réponse immunitaire intense liée à la fois à une stimulation de l'immunité innée et à l'induction d'une réponse cellulaire et

humorale spécifique de l'adénovirus dû à l'antigénicité des protéines de la capside. L'inefficacité de transfert de gène par les adénovirus est attribuée à la réponse humorale et la préexistence des anticorps chez les patients. La majorité des humains adultes possèdent des anticorps contre l'HA5 et dans 57% des cas ils sont capables de neutraliser l'Ad (Schulick et al., 1997). Plusieurs travaux montrent les régions épitopes des protéines à la surface de la capside, notamment celles de l'hexon (Crawford-Miksza and Schnurr, 1996; Toogood et al., 1992), de la fibre (Fender et al., 1995; Liebermann et al., 2001; Liebermann et al., 2002; Mei and Wadell, 1996) et de la base de penton (Hong et al., 2003). Afin d'améliorer l'efficacité de livraison du gène thérapeutique aux cellules cibles, quelques travaux visent l'altération des épitopes les plus immunogéniques et des domaines d'interaction de la capside avec les cellules du système immunitaire ainsi que l'utilisation d'adénovirus animaux

Pour minimiser la propagation incontrôlée du virus et limiter les réactions inflammatoires dues aux virus, trois générations successives de vecteurs adénoviraux non répliquatifs ont été développées (Russell, 2000). Dans la première génération de vecteur adénoviraux, les gènes E1 et/ou E3 ont été délétés, permettant l'introduction de plus de 6.5 kb d'ADN étranger dans la capside (Russell, 2000) (l'Ad sauvage ne peut incorporer que 2 kb d'ADN étranger en plus de son génome). La délétion de la région E1 du génome empêche la transcription de la région E2 (qui est activée par E1) et par conséquent inhibe la répllication indésirable du virus. Il est cependant évident que la délétion des régions E1B et/ou E3 rend les cellules infectées plus vulnérables à l'apoptose et moins favorables à la transcription du transgène.

Les vecteurs de deuxième génération, présentent une délétion partielle ou totale des gènes de la région E2. Ce type de construction évite les problèmes de répllication non contrôlée présents dans les vecteurs de la première génération (Russell, 2000). Cependant, de fortes réactions inflammatoires induites par ce type de vecteur restent un problème majeur pour ces vecteurs. Ces réactions sont particulièrement importantes lors d'infections répétées.

Dans les vecteurs de troisième génération, tous les gènes viraux ont été supprimés. Seules les régions ITRs et la séquence d'encapsulation ont été conservées (Amalfitano et al., 1998; Parks et al., 1996). Désignés aussi sous le nom de 'gutless' ou 'helper dependent', ces virus ont besoin d'un '*helper virus*' pour assurer leur production (Russell, 2000). Ils peuvent acheminer des transgènes de grande taille (en théorie environ 30 kpb) dans les cellules cibles, et contrairement aux vecteurs de première et deuxième génération, la réponse immunitaire a été largement réduite. Plusieurs problèmes sont toutefois associés à cette méthode: la

contamination de la solution de vecteurs par du '*helper virus*' infectieux et l'instabilité du vecteur.

Les vecteurs adénoviraux non réplicatifs ont été utilisés dans la thérapie anti-cancer, sans grand succès. En général, ces vecteurs ne transduisent pas de manière efficace les cellules cancéreuses et ne se propagent pas à partir de leur lieu d'injection ce qui n'entraîne qu'un effet thérapeutique très modeste. Ceci est souvent dû soit à l'absence du récepteur CAR dans certaines tumeurs, soit à des tumeurs difficiles d'accès à cause des barrières anatomiques incluant la vascularisation et la matrice extracellulaire ou encore à l'effet du vecteur limité dans le temps. Pour remédier à ce problème, un nouveau type de vecteur a été développé: les CRAAds (*conditionally replicating adenoviruses*). La majorité des approches utilisées pour générer les CRAAds vise à modifier les gènes de la région E1 qui seront complétés dans les cellules cancéreuses pour maintenir leur expression (Bischoff et al., 1996; Rein et al., 2006). La première stratégie pour arriver à ce but vise à construire des CRAAds avec des mutations de la protéine E1B55k (en temps normal cette protéine se lie à la protéine p53 et la désactive permettant ainsi la réplication du vecteur dans les cellules infectées et la propagation virale). Le vecteur mutant pour la protéine E1B55k ne peut pas réguler l'activité de p53. Après la pénétration du virus, la protéine p53 pourra donc bloquer l'infection en arrêtant le cycle viral et éviter la propagation du virus dans les cellules saines. Les cellules cancéreuses n'exprimant pas la protéine p53, ce vecteur sera capable de s'y répliquer, de provoquer la lyse cellulaire et d'infecter d'autres cellules cancéreuses (Rein et al., 2006). La deuxième stratégie employée pour construire des CRAAds vise à contrôler la réplication virale par un promoteur spécifique présent dans la tumeur (Davydova et al., 2004; Rein et al., 2006).

Parmi les nombreuses autres voies explorées en thérapie génique, nous pouvons citer le recours à des adénovirus « non humains ». Ces virus, parmi lesquels des adénovirus aviaires, ovins, bovins et canins ont un grand avantage sur leur homologue humain car dans la plupart des cas ils ne déclenchent pas de réponse immunitaire importante chez l'homme. L'Ad canin de sérotype 2 (CAV-2), qui a un tropisme pour les neurones humains, a été proposé comme vecteur pour les maladies neuronales (Kremer, 2005).

La Modification du Tropisme d'Ad:

Les HAd2 et 5 sont les sérotypes couramment utilisés pour la thérapie génique, ils reconnaissent le récepteur CAR. L'omniprésence de ce récepteur dans les tissus humains constitue un avantage lorsque l'on veut cibler plusieurs types de tissus cellulaires, mais elle présente également un inconvénient majeur lorsqu'il s'agit de cibler spécifiquement un type cellulaire précis. Il est assez aisé de détruire le tropisme naturel des adénovirus (en mutant la zone de reconnaissance du récepteur au niveau de la tête de la fibre) mais le reciblage du virus reste plus problématique et passe par une modification de la capsid. Différentes méthodes de reciblage sont utilisées (Fig. 1.21), celles-ci sont détaillées dans la revue de Campos et Barry (Campos and Barry, 2007).

- Modification génétique de la capsid induisant un nouveau tropisme

Cette méthode consiste à fusionner génétiquement une protéine de la capsid à un peptide de reciblage reconnu par un récepteur à la surface de la cellule. Plusieurs régions de la capsid telles que la boucle HI et le domaine C-terminal de la fibre, la boucle RGD de la base de penton, la région hypervariable de l'hexon et le domaine C-terminal de la protéine IX ont fait l'objet d'une fusion, (Dmitriev et al., 1998; Dmitriev et al., 2002; Vellinga et al., 2006; Wickham et al., 1996a).

- Utilisation d'un adaptateur moléculaire

Le cas le plus général consiste à conjuguer les virus à des anticorps bi-spécifiques qui se lient à la fois à la fibre et à la cellule. Cette procédure permet par ailleurs l'abolition du tropisme naturel du virus (Wickham et al., 1996b).

- Modification chimique de la capsid

La modification du virus se fait dans ce cas à l'aide d'une réaction chimique qui va lier de façon covalente une molécule à la surface de la capsid (Fisher et al., 2001; Lanciotti et al., 2003).

- Remplacement de la fibre

Un nouveau virus est construit de manière génétique en remplaçant une partie de la fibre ou la fibre entière du virus par celle d'un autre sérotype pour cibler un nouveau type de récepteur (Havenga et al., 2002; Shayakhmetov et al., 2000; Tsuruta et al., 2007).

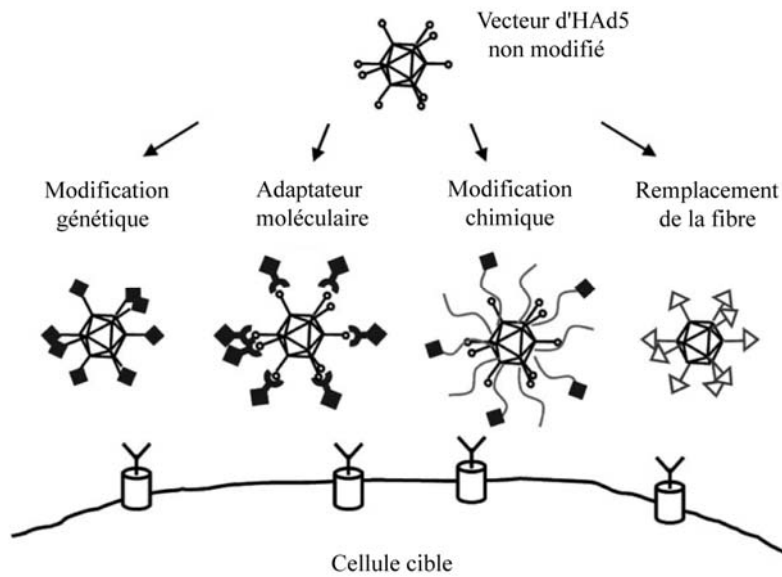


Figure 1.21: Description schématique des quatre stratégies majeures de reciblage des Ad5 vers de nouveaux types cellulaires (Campos and Barry, 2007).

Les Objectif Principaux de la Thèse

Le travail de ma thèse a été dédié à l'étude structurale des protéines de la capsid de l'adénovirus. Dans un premier temps je me suis intéressée aux protéines mineures sur lesquelles aucune donnée structurale n'existe. Dans cette optique, j'ai essayé de cristalliser plusieurs constructions des protéines IX et IIIa. Ces deux protéines se sont avérées insolubles et instables en solution et aucun cristal n'a été obtenu. Malgré l'échec des essais de cristallisation, nous avons accumulé de nombreuses données biochimiques qui aident à mieux comprendre le comportement de ces protéines.

Dans un deuxième temps, je me suis intéressée à la structure des protéines majeures de l'adénovirus. J'ai donc obtenu des cristaux de la fibre courte de l'adénovirus aviaire de sérotype 1. Ce virus possède deux fibres accrochées à la même base de penton, cette particularité a suscité notre intérêt quant au rôle de la deuxième fibre. A la fin de ma thèse le projet de l'amélioration de la diffraction des cristaux du dodécaèdre de l'HAd3 m'a été confié. Cette partie de ma thèse m'a permis d'avoir accès à un projet de cristallographie sur un très grand complexe symétrique (3.5 MDa). J'ai donc réussi à obtenir une condition de cristallisation donnant lieu à une diffraction à haute résolution, le travail de résolution de la structure sera réalisé après la soutenance de ma thèse.

2. *Matériels et méthodes*

2.1. *Sous Clonage*

Les protocoles de sous-clonage incluant PCR (*Polymerase Chain Reaction*), purification d'ADN, digestion, déphosphorylation et enfin ligation sont fournis par les différentes entreprises qui mettent au point les enzymes et les kits.

Pour tous les clonages, les clones positifs sont sélectionnés d'abord par PCR sur colonie puis vérifié par minipréparation d'ADN, digestion enzymatique et enfin par séquençage.

La protéine IX de l'HAD5:

Les différents mutants de la protéine IX (numéro d'accèsion P03281) ont été introduits dans le vecteur pET-3a-6His (résistance à l'ampicilline) par le Dr M Rosa-Calatrava sauf les domaines N et C-terminaux que nous avons introduits nous même dans le vecteur pET28a-TEV.

Les fragments N-terminal (entre le 1^{er} et le 58^{ème} acide aminée) et C-terminal (entre le 85^{ème} et le 140^{ème} aa) de la protéine IX ont été amplifié par PCR à partir du clone sauvage de la protéine IX en utilisant les amorces suivantes (les séquences soulignées présentent les sites de restriction) :

C-terminal sens 5'-CTG ATG CAT ATG CTT GCA AGC AGT GCA GC

C-terminal anti-sens 5'-CTG ATG GGA TCC TTA AAC CGC ATT GGG AGG

N-terminal sens 5'-CTG ATG CAT ATG ATG AGC ACC AAC TCG TTT GAT G

N-terminal anti-sens 5'-CTG ATG GGA TCC TTA CTC CAA CGG CGT TCC AGA C

Les fragments ont été insérés entre les sites de restriction NdeI et BamHI du vecteur pET28a-TEV. Le vecteur pET28a-TEV permet l'expression de la protéine fusionnée à une étiquette 6-

His du côté N-terminal. Cette étiquette polyhistidine qui sert à la purification de la protéine par chromatographie d'affinité, peut être éliminée avec la protéase TEV (*Tobacco Etch Virus*).

La protéine IIIa de l'HA5 :

Le clone de la protéine sauvage (numéro d'accèsion P12537) inséré dans le vecteur pET28a-His6 nous été fourni par le Dr. Saw See Hong. A partir de ce plasmide j'ai réalisé douze constructions de la protéine IIIa tronquées des côtés N et/ou C-terminaux. Six des douze constructions commencent à l'acide aminé 1 et les six autres à l'acide aminé 20, les deux amorces sens sont :

Sens-1 5'-ATG GGT CATATG ATG ATG CAA GAC GCA ACG GAC C
 Sens-20 5'-ATG GGT CAT ATG ATG CTT AAC TCC ACG GAC GAC TGG C

Les 12 constructions sont tronquées du côté C-terminal entre les acides aminés 390 et 440, les amorces anti-sens sont :

Construction	Anti-sens
1 ou 20-390 aa	5'-CTGATG <u>GAGCTCTT</u> ACCAGTGCGGGTTCAAGATGG
1 ou 20-400 aa	5'-CTGATGG <u>GAGCTCTT</u> ATCCCCGGTGTAGAAACCAGG
1 ou 20-410 aa	5'-CTGATGG <u>GAGCTCTT</u> AGAATCCATCGTTACCCTCGGG C
1 ou 20-420 aa	5'-CTGATGG <u>GAG CTCTT</u> AAAACACGCTGTCGTCTATGTCGTCC
1 ou 20-430 aa	5'-CTGATGG <u>GAGCTCTT</u> ACA ACTCTAGCAGGGTCTGCGGTTG
1 ou 20-440 aa	5'-CTGATGG <u>GAGCTCTT</u> ACAGCGCCGCCTCTGCC

Après une amplification par PCR, les inserts ont été introduits entre les sites de restriction NdeI et SacI du vecteur pET9 (résistant à la kanamycine). Le vecteur pET9 ne code pas pour l'étiquette 6-His. J'ai donc recloner les fragments 1-430 ; 20-390 ; 20-400 et 20-410 en introduisant une étiquette 6-His avec un site de coupure à la TEV du côté N-terminal de la séquence. Pour générer ce plasmide deux étapes d'amplification par PCR ont été nécessaires. La première PCR sert à rajouter le site de coupure à la TEV en utilisant les mêmes amorces anti-sens et les amorces sens suivantes:

Site TEV

Sens-1 5'-TTATGATATTCCA ACTACCGAGAATTTGTATTTTCAGGGGATGATGCAA
 GACGCAACGGACC

Site TEV

Sens-20 5'-TTATGATATTCCA ACTACCGAGAATTTGTATTTTCAGGGGCTTAACTCC
ACGGACGACTGGC

La deuxième PCR est réalisée sur le produit issu de la première PCR, elle sert à rajouter l'étiquette 6-His. Les amorces anti-sens sont toujours les mêmes et l'amorce sens est la suivante :

6-His

Sens 5'-GAGTATCATATGGGGCCACCATCATCACCACCATGATTATGATATTCCAA
CTACCGAG

Le produit issu de la PCR est introduit dans le vecteur pET9 entre les sites de restriction NdeI et SacI.

La fibre courte de l'adénovirus aviaire :

Le fragment 201-410 de la fibre courte de l'adénovirus aviaire (numéro d'accèsion Q64762) a été amplifié par PCR en utilisant les amorces suivantes :

Sens 5'-CTG ATG GGA TCC CTG TAC CAA GCG CCC ACT AG

Anti-sens 5'-CTG ATG CTC GAG TCA GAC CGT AAC GGG GGC G

L'insert a été introduit dans le vecteur d'expression pProEx HTb (*Life Technologies*) entre les sites de restriction XhoI et Bam-HI. Le plasmide permet l'expression de la protéine fusionnée à une étiquette 6 - His du côté N-terminal non clivable.

2.2. Surexpression dans le Système Bactérien

Des cellules d'*Escherichia Coli*, souche BL21star compétentes (*Invitrogen*), sont transformées par choc thermique, avec un des plasmides d'intérêts. Elles sont ensuite étalées sur un milieu LB (*Luria Broth*) solide et sélectif. Une colonie isolée d'E. Coli transformée est mise en pré-culture dans 150 ml de milieu LB en présence d'antibiotiques (35 µg/mL chloramphénicol et 50 µg/ml kanamicine ou 100 µg/ml ampicilline selon le plasmide utilisé) pendant une nuit à 37°C sous agitation. Six litres de milieu LB sont ensuiteensemencés avec

les 150 ml de préculture et mis sous agitation à 37°C jusqu'à une DO_{600nm} comprise entre 0.6 et 0.8. L'expression de la protéine recombinante est alors induite par 0.5 mM d'IPTG (*isopropyl-beta-D-thiogalactopyranoside*) à une température variant entre 14 et 30°C (selon la protéine) pendant une nuit. Les bactéries sont ensuite culottées par centrifugation à 6000 g pendant 10 min.

La protéine marquée avec de l'azote N¹⁵ (pour être utilisées en RMN) est produites dans le milieu M9 :

Pour préparer 1L de milieu M9 nous avons besoin de deux solutions préparées fraîchement: 'M9 Salt' 10X et 'Trace elements solution' 100X, leurs compositions (pour 1L) ainsi que la composition de 1L de milieu M9 est indiquée dans les tableaux ci-dessous :

'M9 Salt' 10X

60 g	NaHPO ₄
30 g	KH ₂ PO ₄
5 g	NaCl
5 g	N ¹⁵ H ₄ CL

'Trace elements solution' 100X

5 g	EDTA
0.83 g	FeCl ₃
84 mg	ZnCl ₂
13 mg	CuCl ₂
10 mg	CoCl ₂
10 mg	H ₃ BO ₃
1.6 mg	MnCl ₂

1 L de milieu M9

100 ml	'M9 Solt' 10X
10 ml	'Trace elements solution' 100X
20 ml	20% (W/V) Glucose
1 ml	1 M MgSO ₄
0.3 ml	1 M CaCl ₂
2 ml	Biotine (1 mg/ml)
2 ml	Thiamine (1 mg/ml)
	Eau et Anticorps approprié

Purification du produit d'expression

Le culot de bactéries est resolubilisé dans le tampon de lyse (20 mM Tris HCl pH 8; 300 mM NaCl; 10 mM imidazole) contenant des antiprotéases (20 µM PMSF; cocktail d'antiprotéases complet; EDTA-free de Roche). Les cellules sont cassées par sonication ou par pression (1000 Psi à l'aide de la presse de French 'Thermo Spectronic'). La protéine recombinante est récupérée dans la fraction soluble après une centrifugation à 20 000 g pendant 1 heure à 4°C. Les protéines fusionnées avec une étiquette 6-His sont d'abord purifiées par chromatographie d'affinité: l'extrait brut est chargé sur une colonne de résine agarose-cobalt ou Nickel (Sigma) pré-équilibré dans le tampon de lyse. L'étiquette polyhistidine se lie au cobalt par affinité et la protéine reste attachée dans la colonne. Celle-ci est lavée d'abord avec le tampon de lyse puis avec ce même tampon auquel on a rajouté 20 mM d'imidazole. La protéine est éluée grâce au tampon de lyse additionné de 500 mM

imidazole (un compétiteur qui décroche l'étiquette polyhistidine de la résine). Une étape de dialyse ou de dessalage sur une colonne PD10 (*desalting column Amersham*) est nécessaire pour éliminer l'imidazole. Les protéines dont les étiquettes sont clivables sont incubées avec 1/100 (pds/pds) de la protéase TEV une nuit à 4°C. La protéase TEV est elle-même également fusionnée à une étiquette 6-His qui permettra de la séparer de la protéine d'intérêt clivée sur une deuxième colonne d'affinité.

Une deuxième étape de purification peut être nécessaire, en particulier lorsque la protéine est destinée à être utilisée pour les tests de cristallisation. La chromatographie d'exclusion de taille ou 'gel filtration' permet de séparer les protéines en fonction de leur poids moléculaire. Deux colonnes ont été utilisées, la *Superdex 200* et la *Superdex 75* (*Amersham Pharmacia*) selon la taille de la protéine d'intérêt. La colonne est dans un premier temps pré-équilibrée dans un tampon (généralement 20 mM Tris-HCl pH 7.9 ; 150 mM NaCl) puis 1 mL de solution protéique concentrée est injectée en haut de cette colonne. Les fractions éluées sont suivies par absorption aux UVs puis analysées sur un gel de polyacrylamide SDS-PAGE.

2.3. Surexpression dans le Système Baculovirus

Toutes les étapes de production de la protéine recombinante sont détaillées dans le livre fourni sur le site web de Invitrogen: *Bac-to-Bac[®] Baculovirus Expression System*. Brièvement, les cellules d'insectes sont cultivées en monocouche à 27°C dans des flasques en plastique de 75 cm² à 300 cm². La production et le titrage des baculovirus recombinants sont réalisés sur des cellules Sf21 (souche *Spodoptera frugiperda*). Les cellules HF (souche *High FiveTM Invitrogen*) sont utilisées pour l'expression des protéines recombinantes. Les cellules d'insectes sont cultivées dans du milieu TC100 (*Invitrogen*) supplémenté avec 10 % de sérum de veau fœtal et des antibiotiques (pénicilline / streptomycine).

Purification du produit d'expression

Les cellules d'insecte sont resolubilisées dans le tampon de lyse (20 mM Tris-HCl pH 7.9 ; 150 mM NaCl) puis cassées par plusieurs cycles de congélation décongélation (classiquement 4 cycles entre -200 et 37°C).

Suivant ce protocole, j'ai pu exprimer la base de penton d'adénovirus aviaire mais malheureusement cette protéine était insoluble. La base de penton de l'HA3 a elle aussi été

surexprimée dans les cellules d'insecte par le Dr. Pascal Fender. Le Dr. Fender nous a fourni la protéine après une étape de purification sur gradient de sucrose suivie d'une étape de dialyse contre le tampon suivant: 20 mM Tris-HCl pH 7.9 ; 20 mM de NaCl. La protéine a été ensuite purifiée grâce à une chromatographie échangeuse d'anions. La protéine est chargée sur une Mono-Q (*Amersham*), lavée avec le tampon de lyse à 20 mM NaCl puis éluée grâce à un gradient linéaire de NaCl variant depuis 0.02 jusqu'à 1 M (la protéine est libérée de la colonne pour une concentration d'environ 120 mM). La protéine est concentrée jusqu'à 9 mg/mL, elle est suffisamment pure pour être utilisée directement dans les essais de cristallogénèse ou congelée dans de l'azote liquide et conservée à -80°C.

Une fois pure, la qualité des protéines est vérifiée par différents méthodes, notamment: gel natif ou SDS-PAGE coloré au bleu de coomassie ou au nitrate d'argent ; western blot ; spectrométrie de masse ; séquençage N-ter ; DLS (*Diffusion light scattering*) ; MALLS-SEC (*Multiple Angle Laser Light Scattering-size-exclusion chromatography*); dichroïsme circulaire ; microscopie électronique en coloration négative ... etc. Je ne développerais pas ces techniques ici.

2.4. Détermination de l'Etat d'Oligomérisation

2.4.1. MALLS-SEC

L'analyse par diffusion de lumière d'un échantillon biologique apporte des renseignements sur la taille des particules en solution indiquant ainsi son homogénéité. Pour un résultat fiable, cette technique nécessite 20 à 50 µL de protéine avec une concentration élevée (généralement plus de 5 mg/mL). L'équipement est constitué d'un détecteur MALLS (*Multiple Angle Laser Light Scattering*) couplé à une chromatographie d'exclusion de taille (*SEC size-exclusion chromatography*) et un réfractomètre (wyatt, 1993). La technique de MALLS-SEC permet d'accéder à la masse moléculaire absolue d'une molécule.

Brièvement, la colonne d'exclusion de taille est équilibrée dans un tampon Tris-HCl, pH 7.5, 150 mM NaCl, la protéine est chargée sur la colonne (branchée sur un système HPLC). Au cours de l'éluion la masse moléculaire moyenne est mesurée pour chaque point du chromatogramme. L'intensité de la lumière diffusée est directement proportionnelle à la concentration en protéine et à la masse moléculaire calculée selon la formule de Zimm $M=R_{\theta}/K*C$, où M est la masse moléculaire (g/mol) ; C la concentration en protéine (g/mL) ;

R_θ le rapport de Rayleigh mesuré pour chaque angle de diffusion θ ; K^* est une constante optique (la masse moléculaire ne dépend pas du volume d'éluion sur la colonne).

2.4.2. Pontage Chimique

L'interaction inter-moléculaire au sein d'un oligomère peut être soit stable soit transitoire. Dans ce dernier cas, son observation par MALLS en solution ne serait pas possible. Le pontage chimique couplé à l'analyse de l'échantillon sur gel SDS PAGE offre une méthode directe pour identifier les complexes issus aussi bien d'une interaction stable que d'interactions transitoires. Le principe de cette technique repose sur l'utilisation d'un agent bifonctionnel contenant deux groupements réactifs qui vont chacun créer une liaison covalente avec les groupements fonctionnels des acides aminés. Les deux réactions peuvent se faire soit au sein d'une même protéine soit entre deux protéines qui sont physiquement proches. Plusieurs agents pontants existent mais le glutaraldéhyde est probablement l'agent le plus utilisé pour le pontage chimique de protéine.

Le pontage chimique nécessite l'utilisation d'un tampon sans groupement amine. Les protéines ont été transférées dans un tampon phosphate pH 7.5. La protéine concentré à environ 0.7 mg/ml est incubée avec des quantités croissantes en glutaraldéhyde fraîchement préparé (dans un volume totale de 10 μ l). Le mélange est incubé pendant 30 min à température ambiante, la réaction est ensuite arrêtée par l'addition de 5 μ l du tampon Tris 3 M à pH 8.5. Les protéines sont ensuite analysées sur un gel d'électrophorèse dénaturant, SDS-PAGE.

La formule chimique de la glutaraldéhyde: $O=CH-(CH_2)_3-HC=O$

2.5. La Cristallographie aux Rayons X

La cristallographie aux rayons X est devenu la méthode la plus utilisée pour la détermination de la structure atomique des macromolécules biologiques. Pour ce faire plusieurs étapes sont nécessaires: 1. L'obtention de la protéine pure homogène et en grande quantité. 2. L'obtention d'un cristal ordonné qui diffracte à une résolution quasi-atomique. 3. La collecte et le traitement des donnés cristallographiques. 4. L'obtention des phases et le calcul d'une carte de densité électronique. 5. La construction et l'affinement du modèle.

2.5.1. La Cristallogenèse

La diffusion des rayons X par une molécule isolée est trop faible. Pour avoir suffisamment de signal pour résoudre une structure nous avons besoin d'un cristal. Celui-ci fournit une diffusion cohérente et simultanée par un grand nombre de molécules ordonnées permettant ainsi l'addition des contributions et donc une amplification du signal. La première difficulté pour obtenir la structure cristallographique d'une protéine est l'obtention du cristal. Comme un cristal de sel, un cristal de protéine consiste à un arrangement ordonné dans les trois dimensions de l'espace. La méthode de diffusion de vapeur dans des gouttes assises ou suspendues est la technique la plus utilisée dans ce domaine. Elle consiste à mélanger dans une goutte une solution de protéine (typiquement concentrée entre 1 et 100 mg/mL) et une solution de cristallisation contenant un agent précipitant (généralement 1 µL de chacune des solutions), la goutte est ensuite disposée dans une enceinte scellée à proximité d'un réservoir contenant la solution de cristallisation (environ 1 mL). En général, la solution de cristallisation dans le réservoir contient un agent précipitant qui est deux fois plus concentré que dans la goutte. L'équilibre s'établit entre la goutte et le réservoir par diffusion des espèces volatiles (surtout de l'eau) entraînant une augmentation progressive de la concentration de la protéine et de l'agent précipitant dans la goutte. La solubilité de la protéine diminue parfois en passant par une étape de sursaturation favorable à la nucléation et la croissance cristalline sans atteindre directement la précipitation. Le laboratoire a accès par l'intermédiaire du partenariat pour la biologie structurale PSB (*Partnership in Structural Biology*) à un robot de cristallisation. Celui-ci crible automatiquement jusqu'à 6x96 conditions commerciales avec une quantité de protéine très limitée (100 nL/goutte). La méthode utilisée pour la cristallisation est la diffusion de vapeur en goutte assise (0.1 µl de solution protéique et 0.1 µl de solution de la condition de cristallogenèse avec un réservoir contenant 100 µl de la condition de cristallogenèse).

Il n'y a pas de règle pour obtenir un cristal de protéine. La condition initiale de cristallisation est souvent obtenue de façon empirique en testant des centaines de conditions variant au niveau du précipitant, du pH et du sel. Les cristaux peuvent être ensuite améliorés afin d'obtenir la plus haute résolution possible. Plusieurs facteurs peuvent être visés tels que la concentration en précipitant et en protéine, le pH, la température, la force ionique, l'effet d'additifs, substrat, ligand, cofacteur, protéine partenaire, le cryoprotecteur ... etc.

Les noms des cribles commerciaux utilisés sont indiqués dans le tableau ci-dessous et toutes les compositions sont disponibles sur les sites des fournisseurs.

Plaquette	Cribles	Fournisseur
IBS1:	Crystal screen I & II	HR
IBS2:	Crystal screen Lite and PEG-Ion	HR
IBS3:	MembFac & Natrix	HR
IBS4:	Ammonium Sulfate, QuickScreen, Malonate, Formate	HR
IBS5:	PEG 6K, PEG LiCl, MPD, Mme	HR
IBS6:	Index Screen	HR

Hampton Research=HR

2.5.2. Principe de la Diffraction aux Rayons X

Pour étudier la structure des macromolécules à une résolution atomique, il faut disposer d'une source de rayonnement qui doit avoir une longueur d'onde du même ordre de grandeur que les distances interatomiques ($d=1.5\text{Å}$ entre deux atomes de carbone liés par une liaison covalente simple), d'où l'utilisation des rayons X qui ont une petite longueur d'onde ($\lambda=0.7$ à 2Å). Toutes les données présentées dans cette thèse ont été collectées grâce au rayonnement synchrotron sur les ligne de lumière dédiés à la cristallographie des macromolécules à l'ESRF ($\lambda\sim 1 \text{Å}$). Cette source de rayonnement est caractérisée par une faible divergence et de forte brillance et intensité réduisant considérablement le temps de collecte et améliorant significativement la qualité des données.

Un cristal peut se décrire comme un réseau aux nœuds duquel se trouve un motif: atome, ion, molécule. Il peut être décrit par la translation entière de la maille élémentaire le long des axes du réseau. Eventuellement, le contenu de la maille élémentaire est obtenu à partir de l'unité asymétrique par des rotation et translation qui définissent le group d'espace. Des plans équidistants mais d'orientation différente qui passent par les nœuds de ce réseau s'appellent plans réticulaires (l'équidistance est d_{hkl}). La diffraction des rayons X peut être assimilée à une réflexion sur ces plans réticulaires. Les ondes réfléchies sur les plans réticulaires successifs d'une même famille (caractérisées par un même indice hkl) interfèrent de façon constitutive. La diffraction ne se fait qu'à des valeurs de θ particulières qui vérifient la loi de Bragg:

$$2d_{hkl}\sin\theta = n\lambda \quad \text{Eq.1,}$$

où n est entier, λ est la longueur d'onde incident, θ est l'angle incident et d_{hkl} est la distance interréticulaire qui définit aussi la résolution de la réflexion hkl. Plus la distance d est petite, plus l'angle θ est grand, plus la résolution est grande. Pour une longueur d'onde donnée la

plus petite distance d qui respecte la loi de Bragg est $\lambda/2$, d'où l'utilisation des rayons X ($\lambda \sim 1 \text{ \AA}$) en cristallographie.

La diffraction des rayons X par un cristal se fait dans des directions privilégiées de l'espace correspondant aux noeuds du réseau réciproque. Le réseau réciproque est la transformée de Fourier du réseau direct. Les dimensions de la maille du réseau réciproque sont inversement proportionnelles à celles de la maille du réseau direct. Lorsqu'un faisceau monochromatique de rayons X éclaire un cristal, la condition de diffraction n'est vérifiée que si un noeud du réseau réciproque coïncide avec le vecteur de diffusion à la surface de la sphère d'Ewald. La sphère d'Ewald est une sphère de rayon $1/\lambda$ dont le centre coïncide avec un noeud du réseau réciproque. Pour que les noeuds du réseau réciproque et la sphère d'Ewald se croisent, on fait tourner le cristal autour d'un axe perpendiculaire au faisceau incident durant la collecte des données. La rotation du cristal entraîne la rotation du réseau réciproque associé. Il faudra donc effectuer une rotation d'un certain angle minimum pour collecter des données couvrant la plus grande partie de l'unité asymétrique de l'espace réciproque (l'angle de rotation minimum dépend du groupe d'espace, plus la symétrie est élevée plus il y a de réflexions équivalentes). Pour réduire les erreurs sur les mesures, il faut collecter plusieurs fois des réflexions qui sont symétriquement équivalentes.

2.5.3. Traitement des Données

Le traitement des données passe par trois étapes : indexation ; intégration et mise à l'échelle du jeu de données. Ce traitement est effectué de manière automatique dans les programmes tels que DENZO (Otwinowski and Minor, 1997a; Otwinowski and Minor, 1997b), MOSFLM (Leslie, 1992) ou XDS (Kabsch, 1988). La répartition des tâches sur le cliché de diffraction est liée au réseau réciproque, au groupe d'espace et l'orientation du cristal dans l'espace. L'intensité des tâches contient une partie de l'information sur la densité électronique et la position des atomes. Après la détermination des paramètres de la maille, du groupe d'espace et de l'orientation du cristal grâce aux positions des tâches sur le cliché de diffraction, tous les clichés sont indexés. On affine un certain nombre de paramètres tel que la distance cristal-détecteur, la position du centre du détecteur, l'orientation du cristal, les paramètres de maille. Pour chaque tâche un indice hkl est attribué, les profils de diffraction prédits sont ajustés aux tâches observées afin d'obtenir un meilleur accord entre les deux. Durant l'intégration l'intensité de chaque tâche est mesurée et le bruit de fond est

soustrait. Les images intégrées sont ensuite mises à l'échelle afin de constituer un jeu de données complet et homogène. Les réflexions équivalentes par symétrie sont moyennées. En effet, chaque image diffère par la dose de rayons X reçue et par la baisse du pouvoir de diffraction du cristal (dommages dus au radiation). Le rapport signal/bruit ainsi que les valeurs des R_{sym} entre les réflexions équivalentes sur l'ensemble du jeu de données permettent d'estimer la qualité du jeu de données.

$$R_{sym} = \frac{\sum_h \sum_i |I_i(h) - \langle I(h) \rangle|}{\sum_h \langle I(h) \rangle} \quad \text{Eq. 2,}$$

où $I_i(h)$ est une mesure individuelle, $\langle I(h) \rangle$ est la moyenne pondérée de toutes les mesures des réflexions équivalentes $I(h)$ et h est l'indice (h, k, l).

2.5.4. Facteur de Structure et Densité Electronique

Les rayons X sont des ondes électromagnétiques qui interagissent avec le nuage électronique des atomes. Un nombre complexe appelé facteur de structure $\mathbf{F}(hkl)$ est lié à la densité électronique. Le facteur de structure comporte l'amplitude $|F(hkl)|$ et la phase $\varphi(hkl)$ de l'onde diffusée dans une direction particulière par rapport à une onde incidente.

$$\rho(xyz) = (1/V) \sum_{hkl} \mathbf{F}(hkl) \exp[-2\pi i(hx+ky+lz)] \quad \text{Eq.3,}$$

$$\mathbf{F}(hkl) = |F(hkl)| \exp(i\varphi(hkl))$$

où V est le volume de la maille ; i est une unité imaginaire ; x, y, z sont des coordonnées fractionnelles dans l'espace direct par rapport à la maille élémentaire et h, k, l sont les indices des vecteurs du réseau réciproque.

L'amplitude est proportionnelle à la racine de l'intensité des réflexions mesurées. Cependant la phase ne peut pas être mesurée directement. Sans la phase, la densité électronique ne peut pas être calculée (Eq.3).

2.5.5. *Obtention des Phases*

Plusieurs méthodes permettent l'obtention des phases initiales. Les méthodes les plus utilisées sont le remplacement isomorphe, la diffusion anomale et le remplacement moléculaire si un modèle proche de la structure à résoudre existe. Dans cette thèse nous avons eu surtout recours au remplacement isomorphe.

Le Remplacement Isomorphe

Cette méthode nécessite l'introduction de manière isomorphe d'atome lourd dans le cristal. L'atome lourd peut être introduit par plusieurs méthodes. Le trempage du cristal natif dans une solution contenant un composé d'atome lourd permet la diffusion de ce dernier à travers des canaux de solvant dans le cristal. Le composé se fixe sur des sites précis dépendants de l'environnement chimique. Les atomes lourds les plus utilisés sont le mercure, le platine, l'or, l'uranium, le sélénium et le gadolinium. Ils interagissent de manière préférentielle avec la chaîne latérale de la méthionine, de la cystéine, de l'histidine, du glutamate, de l'acide aspartique, de la lysine et du tryptophane. Le temps de trempage dépend de la réactivité du composé contenant l'atome lourd. L'incorporation de l'atome lourd peut perturber l'empilement cristallin et entraîner un défaut d'isomorphisme entre les cristaux natifs et dérivés. D'autres méthodes consistent à introduire un atome lourd avant cristallisation de façon covalente. La protéine recombinante contenant de la séléno-méthionine (SeMet) à la place de la méthionine peut être produite dans des bactéries auxotrophes à la méthionine (Hendrickson et al., 1990).

Le cristal contenant un dérivé d'atome lourd ne diffère du cristal natif que par l'addition de l'atome lourd, on parle du remplacement isomorphe, d'où :

$$\mathbf{F}_{PH} = \mathbf{F}_P + \mathbf{F}_H \quad \text{Eq. 4,}$$

où \mathbf{F}_{PH} ; \mathbf{F}_P et \mathbf{F}_H représentent le facteur de structure d'indices hkl du cristal *dérivé* ; du cristal natif et de l'atome lourd respectivement. Les amplitudes $|\mathbf{F}_{PH}|$ et $|\mathbf{F}_H|$ sont accessibles expérimentalement lors des expériences de diffractions.

Pour résoudre la structure nous devons déterminer les positions des atomes lourds. Pour cela la fonction de différence de Patterson est utilisée. La fonction de Patterson est une série de Fourier qui contient les amplitudes du facteur de structures au carré comme coefficient, elle est donc calculée sans avoir besoin de la phase:

$$P(uvw) = (1/V) \sum_{hkl} |F(hkl)|^2 \exp(-2\pi i(hu + kv + lw)) \quad \text{Eq. 5}$$

avec u , v et w comme coordonnées relatives dans la maille de Patterson, qui a les mêmes dimensions que la maille élémentaire (réel). Un pic sur les cartes de Patterson correspond à un vecteur interatomique. Exemple: Si on a un axe d'ordre 2 selon z , l'atome en (x,y,z) aura un symétrique en $(-x,-y,z)$. La fonction de Patterson aura un pic en $(x,y,z) - (-x,-y,z) = (2x,2y,0)$. Le plan $Z=0$ est le plan dit de Harker. On y retrouve tous les atomes avec comme coordonnées $2x,2y$. Il est possible de déterminer les coordonnées des atomes lourds en calculant la fonction de différence de Patterson entre les amplitudes des facteur de structures du dérivé $|F_{PH}|$ et du native $|F_p|$. A partir des positions des pics sur les cartes de Patterson on peut déduire les coordonnées des atomes lourds dans la maille élémentaire, des programmes comme SHELXS utilise cette méthode. La phase pour chaque facteur de structure natif peut être estimée à partir de la contribution de F_H ainsi que les amplitudes $|F_H|$ et $|F_{PH}|$.

La Diffusion Anomale

Lorsque les atomes lourds sont excités à une longueur d'onde proche de leur seuil d'absorption, ils absorbent de l'énergie créant un déphasage entre l'onde incident et l'onde diffracté, ils réémettent de l'énergie sous forme de fluorescence appelé diffusion anormale. Ce phénomène se manifeste par la violation de la loi de Friedel qui dit que les intensités des réflexion hkl et $(-h-k-l)$ sont égales (paires de Friedel). La méthode de Patterson permet ensuite la localisation des diffuseurs anormaux et le calcul de phase.

Le Remplacement Moléculaire

Le remplacement moléculaire peut être utilisé lorsque l'on dispose d'une structure similaire connue. Brièvement, le modèle est d'abord orienté dans la maille par une fonction de rotation puis positionné dans le cristal par une fonction de translation à l'aide de la fonction de Patterson. Le modèle connu, une fois placé dans l'unité asymétrique, sert à calculer des facteurs de structures F_{calc} . Les phases calculées à partir du modèle φ_{calc} sont associées aux

valeurs des modules déduites des intensités diffractées $|F_{\text{obs}}|$. La densité électronique est calculée à partir des amplitudes $|F_{\text{obs}}|$ et des phases φ_{calc} selon l'équation 3. La première carte obtenue avec des phases venant du remplacement moléculaire, peut être améliorée par des modifications du modèle, par substitution des acides aminés, construction des parties manquantes ou moyennation de densités électroniques en utilisant les symétries non cristallographiques.

2.5.6. Construction et Affinement du Modèle

Lorsque la densité électronique devient interprétable, un modèle moléculaire initial est construit manuellement dans la densité électronique calculée F_{obs} via une interface graphique tel que *coot* (Emsley and Cowtan, 2004). Des nouvelles phases peuvent être calculées à partir du modèle initial. Deux cartes de densité électronique de type $2F_{\text{obs}}-F_{\text{calc}}$ et $F_{\text{obs}}-F_{\text{calc}}$ sont calculées, elles servent ensuite à l'amélioration du modèle. Les étapes d'affinement du modèle qui suivent consistent à optimiser l'accord entre les facteurs de structures expérimentaux $|F_{\text{obs}}|$ et les facteurs de structures calculés $|F_{\text{calc}}|$ en tenant compte des contraintes géométriques. L'affinement se fait via des programmes tels que *Refmac* (Murshudov et al., 1997) et *CNS* (Brunger et al., 1998). La qualité du modèle est évaluée par le facteur R:

$$R = \frac{\sum_{hkl} ||F_{\text{obs}}(hkl)| - |F_{\text{calc}}(hkl)||}{\sum_{hkl} |F_{\text{obs}}(hkl)|} \quad \text{Eq.5}$$

Plus le facteur R est petit plus le modèle est proche des données de diffraction. Pour éviter le suraffinement de la structure, le facteur R est aussi calculé pour une partie des réflexions (généralement 5%) qui n'ont pas été utilisées lors de l'affinement. Ce facteur R_{libre} (R_{free}) permet un suivi de l'affinement moins biaisé.

D'autres paramètres sont considérés durant l'affinement comme le facteur d'agitation thermique ; les longueurs de liaison et les angles de torsion. La répartition de ces angles donne donc un critère supplémentaire de qualité du modèle affiné. En effet ils doivent respecter le diagramme de Ramachandran (Ramachandran and Sasisekharan, 1968) qui donne la répartition des angles dièdres ϕ et ψ selon les structures secondaires des protéines.

RESULTATS ET DISCUSSION

Dans ce chapitre je discuterai les travaux réalisés sur les protéines structurales de l'adénovirus, notamment sur deux protéines mineures (protéines IX et IIIa), ainsi que sur deux protéines majeures (la tête de la fibre et la base de penton).

3. *La Protéine IX*

Les protéines majeures de la capsid (Hexon, Fibre, Base de penton) ont déjà été largement étudiées ; leurs structures, fonctions et positions dans la capsid sont aujourd'hui bien définies. Les protéines mineures (IX, IIIa, VI, et VIII) ont fait l'objet de moins d'attention, leurs structures et fonctions restent encore assez méconnues. Leurs positions dans la capsid restent parfois ambiguës.

Unique au sein des Mastadénovirus, la protéine IX est la plus petite protéine de la capsid virale (14.4 kDa). Elle est associée à l'hexon sous forme de trimère. Il s'agit d'une protéine impliquée dans plusieurs aspects biologiques et structuraux du virus. Nous nous sommes intéressés à son rôle architectural au sein de la capsid. La compréhension de la protéine IX est l'un des projets principaux du laboratoire ; plusieurs études sont donc menées simultanément en microscopie électronique, en cristallographie et en biochimie.

3.1. *La Solubilité de la Protéine IX*

La protéine IX sauvage a fait l'objet de plusieurs tentatives d'expression en bactéries, mais se révèle insoluble. Selon les prédictions de structures secondaires, son extrémité N-terminale serait partiellement déstructurée alors que son côté C-terminal est prédit être replié en super hélice (*coiled-coil*) (Fig. 3.1A). Il a été montré que des mutations ponctuelles introduites au niveau du *coiled-coil* ou dans la partie N-terminale de la protéine IX permettent de moduler son degré de solubilité. Nous disposons de 14 mutants de ce type (Fig. 3.1) fournis par le Dr. Manuel Rosa Calatrava (Lyon) dans le cadre d'une collaboration. Mon premier objectif fut de caractériser les différents mutants de la protéine IX afin de déterminer les degrés de solubilité et d'expression ainsi que la stabilité des différentes constructions. La liste des mutants ainsi que leur solubilité sont résumés dans le tableau de la figure 3.1.

3.2. Essai de Détermination de la Structure de la Protéine IX

Durant ces trois ans, nous avons essayé de déterminer la structure atomique de la protéine IX par plusieurs méthodes.

Tous les mutants solubles (Fig. 3.1) ont été surexprimés et ont fait l'objet d'une caractérisation biochimique et d'essais de cristallisation (voir détail pour certaines de ces protéines plus loin). Pour chacun de ces mutants des centaines de conditions de cristallisation (liste dans la partie Matériels et Méthodes) ont été criblées à différentes températures (4, 12 ou 20°C) et concentrations protéiques (entre 3 et 30 mg/ml) soit à l'aide du robot de cristallisation, soit manuellement. Seules quelques plaquettes bidimensionnelles ont été obtenues pour le mutant IX- Δ SYL-V117D (Fig. 3.2). Ces cristaux sont fragiles, instables en présence de cryoprotecteur et ne diffractent pas sur les lignes de lumière à l'ESRF.



Figure 3.2: Plaquettes bidimensionnelles obtenues pour la protéine IX- Δ SYL-V117D. 0.1 μ l de précipitant (0.1 M HEPES pH 7 ; 0.5 M formate) est rajouté à 0.1 μ l de solution protéique à 3.8 mg/mL (dans le tampon Tris 20 mM et 150 mM NaCl). Les gouttes assises ont été réalisées par le robot (PixSys4200 Cartesian) à 24° C, les cristaux apparaissent ensuite à 4° C, au bout d'une nuit et peuvent atteindre une taille de 0.6 mm dans une dimension. Ces cristaux sont reproductibles dans 24 conditions différentes, où le pH du tampon varie entre 4 et 9, et la concentration en formate varie de 0.5 à 3.2 M.

Malgré beaucoup d'acharnement, aucun cristal donnant lieu à une diffraction n'a été obtenu pour la protéine IX. Nous avons conclu que l'échec des essais de cristallisation est probablement dû à la flexibilité du domaine N-terminal (seul ou de ce domaine par rapport au domaine C-terminal) suggérée par les prédictions de structures secondaires (Fig. 3.1A). Dans le but de valider cette hypothèse, nous nous sommes intéressés à la structure des domaines N et C-terminaux isolés.

Le domaine C-terminal a été exprimé et purifié pour ensuite faire l'objet de tests de cristallogénèse qui n'ont pas abouti. Vu la petite taille du fragment (6 kDa sans les 6-His et le site TEV) nous avons projeté une étude structurale par RMN en collaboration avec le Dr Jean Pierre Simorre de l'Institut de Biologie Structurale (IBS) à Grenoble. La protéine marquée

avec l'azote ^{15}N est produite dans *Escherichia coli* en milieu minimum M9 (cf. Matériels et Méthodes) supplémenté avec du $^{15}\text{N-H}_4\text{Cl}$. La protéine a été purifiée et le marquage vérifié par spectrométrie de masse à l'IBS. La figure 3.3 montre le spectre HSQC enregistré sur un spectromètre RMN (800 MHz).

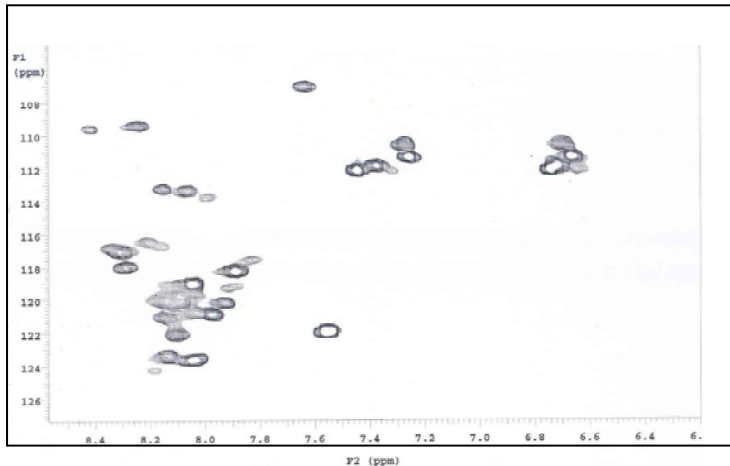


Figure 3.3: Spectre HSQC du fragment C-terminal de la protéine IX à 800 MHz: le spectre est mesuré avec 150 μM de protéine dans 20 mM NaH_2PO_4 et 200 mM NaCl pH 7.5 à 4°C. Ce spectre montre les corrélations entre l'hydrogène porté par le carbone α et le ^{15}N .

Le spectre montre les corrélations entre hydrogène ^1H et azote ^{15}N , chaque pic entre 7.6 et 8.4 ppm correspond à un N-amid et donc à un acide aminé. Sur le spectre on peut compter 23 pics représentant 23 acides aminés. Le fragment C-terminal de la protéine IX étudié est trimérique et contient au total 56 acides aminés. Nous pouvons donc conclure, qu'une grande partie de ce fragment est très flexible. Nous n'avons donc pas pu procéder à l'attribution séquentielle du squelette de la protéine, mais nous avons malgré tout pu conclure grâce à l'indice des déplacements chimiques observés, que cette région de 23 acides aminés se structure en une hélice α . Ce résultat confirme la prédiction de structure secondaire, a également été vérifié par la suite par dichroïsme circulaire.

De la même façon le fragment N-terminal de 58 acides aminés a été cloné dans le vecteur pET-TEV. Malheureusement le faible taux d'expression de cette protéine n'a pas permis d'obtenir une quantité raisonnable pour la caractérisation biochimique. Le domaine N-terminal ne semble pas structuré en solution ce qui réduit probablement sa durée de vie dans la bactérie.

3.3. Localisation de la Protéine IX

Récemment, la structure tridimensionnelle d'un adénovirus dépourvu de la protéine IX a été déterminée au laboratoire par MET (Fabry et al., 2005). Les auteurs ont noté l'absence de la densité correspondante à la protéine IX ainsi qu'une autre densité cylindrique longtemps attribuée à la protéine mineure IIIa (densité orange sur la figure 3.4A). Cette dernière densité a été plus récemment attribuée au domaine C-terminal de la protéine IX par les travaux de Saban et al. (2006).

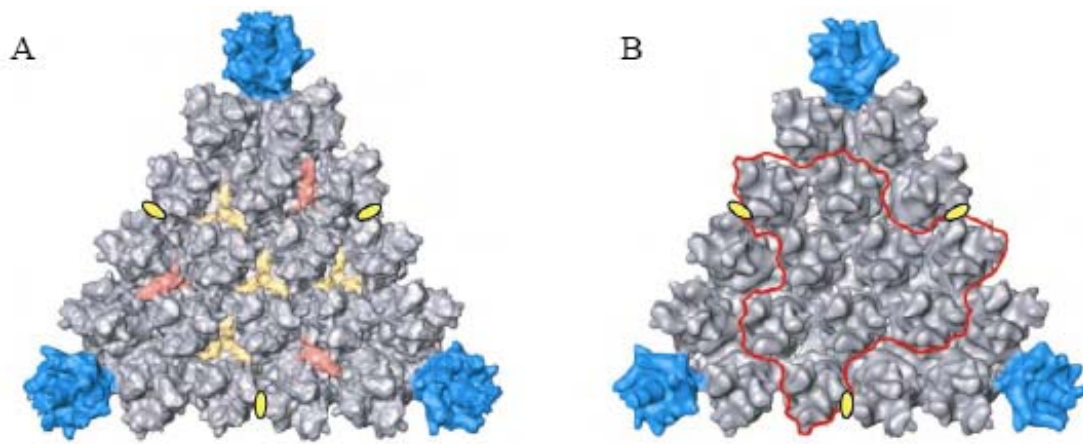


Figure 3.4: Comparaison de l'HAAd5 et de l'HAAd5 delta IX (Fabry et al., 2005).

A, Facette de l'HAAd5 sauvage (10 Å de résolution). Le penton est représenté en bleu, l'hexon en gris, la protéine IX en jaune, la densité cylindrique orange était supposée être la protéine IIIa, mais elle disparaît chez le mutant l'HAAd5 delta IX.

B, Facette de l'HAAd5 delta IX (16 Å de résolution). Les densités jaune (la protéine IX) et orange (la protéine IIIa) sont absentes dans le virus délété de la protéine IX.

Pour résoudre cette ambiguïté quant à la localisation de la protéine IX à la surface de la capside et afin de mieux comprendre comment cette protéine est incorporée dans celle-ci, plusieurs approches ont été utilisées:

- √ Etudes des états d'oligomérisation de plusieurs mutants de la protéine IX *in vitro*.
- √ Détection du domaine C-terminal de la protéine IX à l'aide d'anticorps spécifique à la surface de l'adénovirus humain par MET.
- √ Détection du domaine C-terminal de la protéine IX fusionnée à la GFP à la surface de l'adénovirus canin par MET.

3.3.1. *Etats d'Oligomérisation de la Protéine IX*

Le domaine *coiled-coil* (C-ter) de la protéine IX contient un domaine leucine zipper putatif qui pourrait être responsable d'un trimère présumé de la protéine IX (Rosa-Calatrava et al., 2001; Vellinga et al., 2005). Afin de vérifier cette proposition *in vitro* et de vérifier également si la protéine IX est capable de former des oligomères *via* son extrémité N-terminale et/ou la région centrale riche en alanine, la capacité d'oligomérisation de chaque domaine a été étudiée. Dans ce but nous avons utilisé quatre mutants:

- 1- Le domaine C-terminal de la protéine IX isolé
- 2- La protéine IX Δ SYL-V117D (un mutant C-terminal défectueux)
- 3- La protéine IX- Δ SYL (la protéine entière)
- 4- La protéine IX Δ SYL- Δ ala (un mutant délété de la région centrale riche en alanine)

a. *La Région C-terminale de la Protéine IX*

Le domaine C-terminal de la protéine IX est un domaine *coiled-coil* riche en leucine. Par analogie avec d'autres domaines *coiled-coil* qui forment des trimères ou des oligomères (Hurst, 1995; Lupas, 1996a; Lupas, 1996b; Shu et al., 1999), le domaine C-terminal de la protéine IX a été proposé comme étant responsable de la trimérisation de la protéine IX à la surface de la capsid (Rosa-Calatrava et al., 2001; Vellinga et al., 2005). Il n'est cependant pas nécessaire pour l'incorporation de la protéine dans la capsid ni pour la stabilisation de celle-ci. Ce domaine est accessible à la surface de la capsid et a fait l'objet de modifications qui ne gênent pas l'intégration de la protéine IX dans la capsid (*cf.* Introduction: protéine IX).

Surexpression et Purification

Les 56 acides aminés C-terminaux de la protéine IX (en vert Fig. 3.1A) ont été clonés dans un vecteur pET-TEV codant également pour une étiquette hexa-histidine clivable (la masse moléculaire du fragment est de 6 kDa sans les 6-His et 8.2 kDa avec la queue 6-His). Ce fragment de la protéine IX a été surexprimé chez *E. Coli* et purifié comme décrit dans la partie Matériels et Méthodes.

Le domaine C-terminal seul est produit en grande quantité. Celui-ci est soluble, pur et homogène. Le profil d'élution (Fig. 3.5) nous informe également sur l'état d'oligomérisation

du domaine C-terminal de la protéine IX. Ce domaine de 8.2 kDa est en effet élué à un volume de 11 mL correspondant à une taille standard d'environ 24 kDa soit un état trimérique consistant avec les données indiquant que le domaine C-terminal est impliqué dans la trimérisation de la protéine IX.

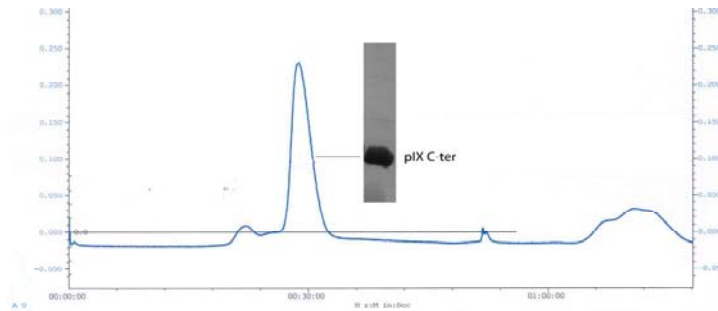


Figure 3.5: Profil d'éluion du fragment C-terminal de la protéine IX par une colonne d'exclusion de taille Superdex 75. Le gel SDS-PAGE révèle la pureté de la protéine.

Deux autres techniques ont été utilisées pour vérifier l'état d'oligomérisation du domaine C-terminal de protéine IX: le MALLS-SEC et le pontage chimique.

Etat d'Oligomérisation: MALLS-SEC et Pontage Chimique

Les expériences de MALLS et de pontage chimique sont complémentaires. Le MALLS permet la détection des populations majoritaires en solution. Le pontage chimique quant à lui stabilise les populations qui se forment *via* une interaction de faible affinité ou transitoire et qui sont peu ou pas visibles en solution.

Durant l'expérience de MALLS la masse du domaine C-terminal de la protéine IX est calculée en fonction du volume d'éluion. L'homogénéité du poids moléculaire observé tout au long du pic d'éluion indique la présence d'une seule population monodisperse correspondant au trimère (Fig. 3.6A). Ces résultats expérimentaux confirment donc la proposition selon laquelle le domaine C-terminal est responsable de la trimérisation de la protéine IX.

Le pontage covalent à l'aide de glutaraldéhyde montre que le domaine C-terminal forme également un trimère (Fig. 3.6B), en accord avec le profil d'éluion sur la S75 et le MALLS (Fig. 3.5 et 3.6A). Toutefois, d'autres populations sont présentes correspondant au monomère, au dimère et parfois au tétramère: le monomère et le dimère sont des états intermédiaires du trimère, ils sont détectés soit car le pontage est incomplet soit car le pontage intramoléculaire sature les sites potentiels de pontage intermoléculaire. Le tétramère (bande très faible) peut être un artéfact du pontage chimique mais il peut aussi être le produit d'une

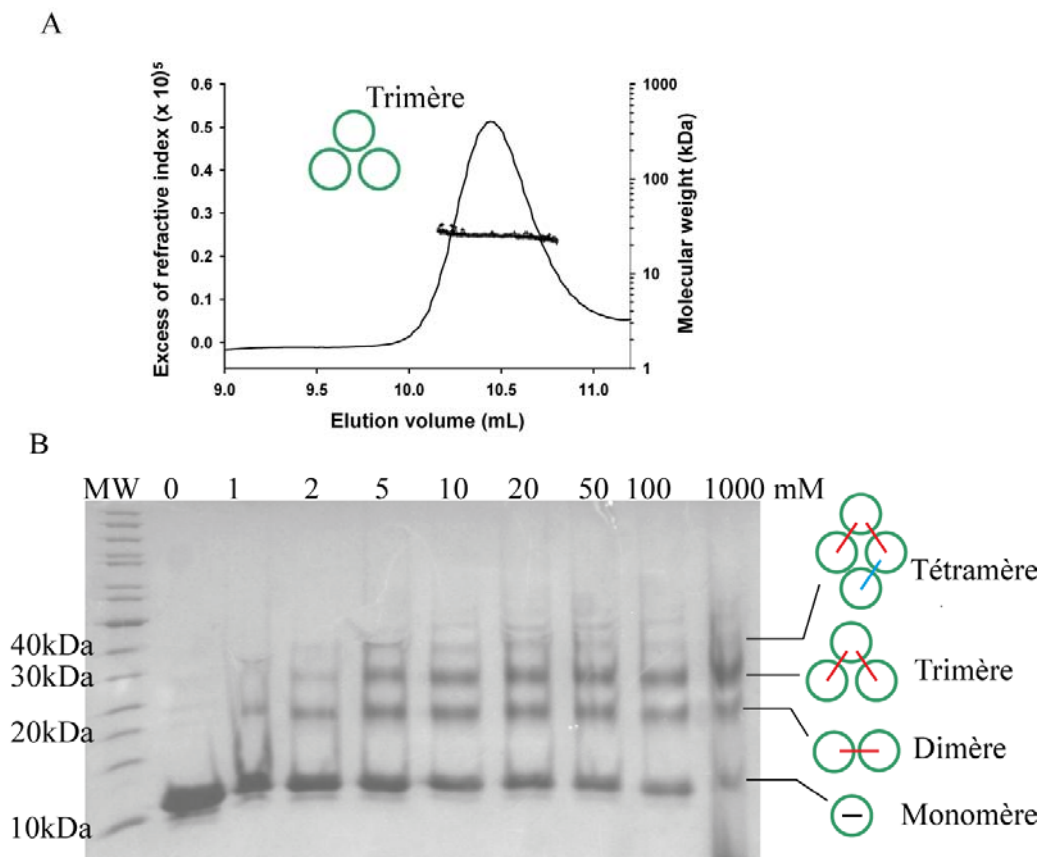


Figure 3.6: Etats d'oligomérisation du domaine C-terminal de la protéine IX

A, Expérience de MALLS-SEC: la courbe représente le profil d'éluion sur une colonne d'exclusion de taille S200. Le graphe indique la masse moléculaire déterminée à partir des mesures de diffusion de lumière et de réfractométrie en fonction du volume d'éluion. La masse moléculaire calculée est d'environ 25 kDa (la masse attendue pour le trimère est de 24.6 kDa). La masse ne varie pas au cours du chromatogramme indiquant le caractère monodisperse de l'échantillon.

B, Expérience de pontage chimique, la protéine (0.7mg/mL) est incubée avec des quantités croissantes de glutaraldéhyde pendant 30 min à température ambiante. La réaction (pontage covalent intra et inter moléculaire) est ensuite arrêtée par l'addition de tampon Tris 3 M, pH 8.5. Les protéines sont analysées sur un gel SDS-PAGE. La piste MW est la piste des marqueurs de poids moléculaires Fermentas ; Les chiffres en haut de chaque piste représentent la concentration en glutaraldéhyde. Le domaine C-terminal est schématisé avec un cercle vert, les traits indiquent le pontage intramoléculaire (noir), le pontage intermoléculaire (rouge) et un pontage intermoléculaire d'oligomère transitoire non observé en solution par le MALLS (bleu).

interaction de faible affinité au niveau du domaine C-terminal (non observée en solution par MALLS mais stabilisée par le pontage chimique). Il faut noter qu'avec les expériences de pontage chimique plus le nombre de molécules pontées est grand, moins il y a de chance que cela soit un artéfact du fait d'un grand nombre de pontage covalent nécessaire au sein du même complexe. L'étude de Saban et *al.* (2006) montre que dans la reconstruction 3D obtenue de l'HA35, la partie C-terminale de la protéine IX est présente sous forme de tétramère, en accord avec nos résultats de pontage covalent.

Pour conclure, le domaine C-terminal est donc responsable de la formation d'un trimère stable en solution, mais il est aussi capable d'induire un tétramère transitoire. Ce tétramère a été mis en évidence dans la capsid par Saban et *al.* (2006).

b. La Protéine IX Δ SYL-V117D

Dans ce mutant, en plus de la délétion de trois acides aminés SYL dans le domaine N-terminal (cette délétion est nécessaire pour obtenir la protéine sous forme soluble), une charge négative a été introduite dans le domaine *coiled-coil* en remplaçant la valine 117 par un acide aspartique. Le résidu valine 117 dans le *coiled-coil*, se révèle strictement nécessaire à la multimérisation de la protéine (Rosa-Calatrava et al., 2001). Ce mutant nous permet donc de sonder spécifiquement les propriétés d'oligomérisation des régions centrale et N-terminale de la protéine IX.

Surexpression et Purification

La méthodologie utilisée pour surexprimer la protéine IX- Δ SYL-V117D dans un système bactérien est la même que celle décrite précédemment. Elle nous permet d'obtenir la protéine en grande quantité. Le degré de pureté de l'échantillon après gel filtration, est révélé dans la figure 3.7.

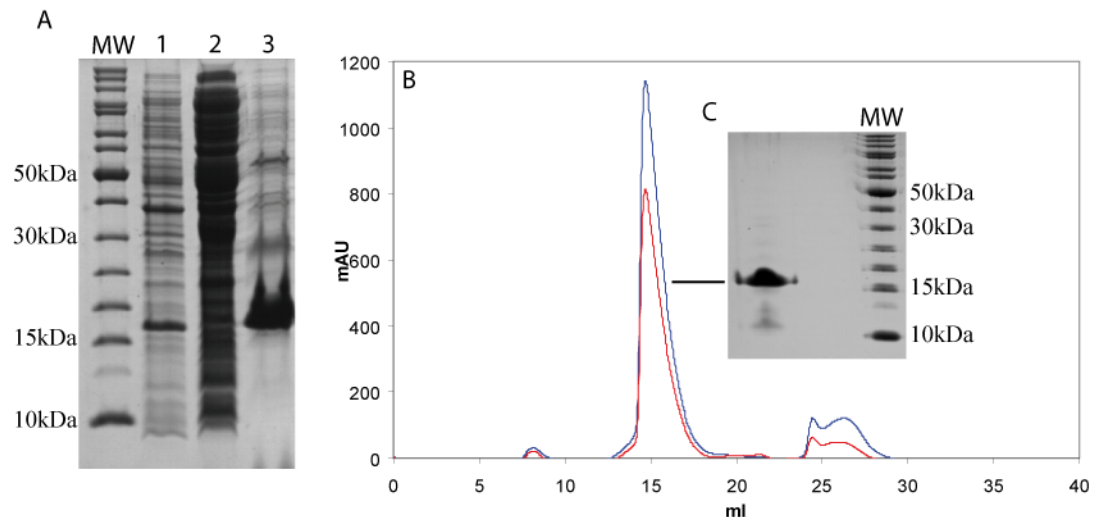


Figure 3.7: Etapes de purification de la protéine IX- Δ SYL-V117D.

A, Gel dénaturant SDS-PAGE: Piste 1: Fraction totale ; Piste 2: Lavage avec 20 mM imidazole ; Piste 3: élution avec 300 mM imidazole.

B, Profil d'élution de la protéine IX- Δ SYL V117D après gel filtration (Superdex 200). La protéine est éluee à un volume de 16 mL comme le montre le gel SDS-PAGE en C.

Selon le profil d'élution sur la colonne d'exclusion de taille (Fig. 3.7B), la protéine IX- Δ SYL-V117D est éluee à un volume de 16 mL, correspondant à la taille d'un monomère de la protéine IX. Nous pouvons donc confirmer que la mutation V117D est nécessaire et suffisante pour empêcher l'oligomérisation du domaine C-terminal. Ce mutant nous permettra donc d'étudier le rôle du domaine N-terminal et de la région centrale dans l'oligomérisation de la protéine IX.

Etat d'Oligomérisation: MALLS-SEC et Pontage Chimique

D'après le résultat du MALLS-SEC (Fig. 3.8A), le mutant IX-V117D de la protéine IX est un monomère. Ce résultat avait été envisagé grâce à la valeur du volume d'élution de cette protéine (Fig. 3.7B).

Avec les expériences de pontage chimique (Fig. 3.8B), deux espèces majoritaires sont détectées: le monomère et le dimère. La présence d'un dimère prouve que la protéine IX- Δ SYL-V117D est capable d'interagir *via* son domaine N-terminal et/ou central (la formation d'un pontage covalent entre deux monomères étant la preuve directe et évidente de leur proximité en solution). Ce dimère est probablement instable en solution et en équilibre avec le monomère.

Nous avons montré que le domaine C-terminal était responsable de la formation d'un trimère stable et d'un tétramère transitoire et que le domaine central ou le domaine N-terminal est impliqué dans une interaction de faible affinité générant un dimère de la protéine IX.

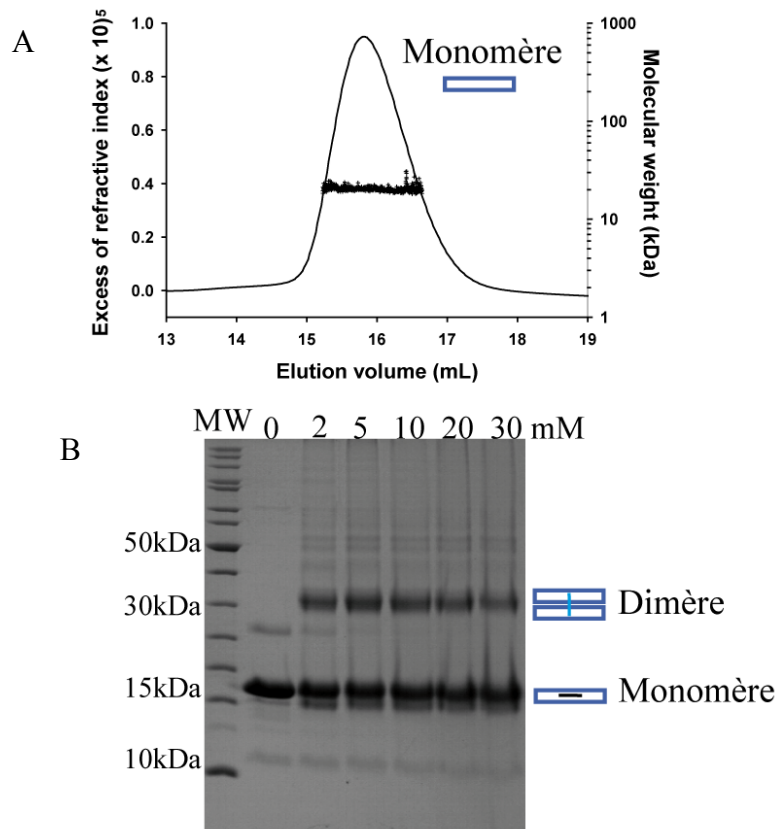


Figure 3.8: Etats d'oligomérisation du mutant IX- Δ SYL-V117D.

A, Expérience MALLS-SEC: La masse moléculaire d'environ 16 kDa ne varie pas le long du chromatogramme (la masse attendue pour le monomère est de 15.5 kDa). L'échantillon peut donc être considéré comme monodisperse et monomérique.

B, Expérience de pontage chimique: les pistes de 0 à 30 contiennent respectivement 0 ; 2 ; 5 ; 10 ; 20 et 30 mM de glutaraldéhyde. Les domaines N-terminal et central sont schématisés en rectangle bleu, les traits indiquent le pontage intramoléculaire (noir) et le pontage intermoléculaire d'oligomère transitoire non observé en solution (bleu).

c. La Protéine IX Δ SYL

La délétion Δ SYL (Ser13, Tyr14 et Leu15) du côté N-terminal dans la protéine IX permet d'obtenir la protéine sous forme soluble tout en conservant les caractéristiques et fonctions de la protéine sauvage *in vivo*. En effet, le Dr Manuel Rosa Calatrava a montré que cette protéine mutante est capable de s'incorporer dans la capsid virale, de former des corps d'inclusions dans le noyau des cellules infectées et de conserver son activité transcriptionnelle (*communication personnelle, Dr. Calatrava*). Pour la suite de nos études nous considérons ce mutant IX- Δ SYL comme un pseudo-sauvage avec au mieux un comportement similaire à celui de la protéine sauvage entière et au pire un comportement d'une protéine dont l'extrémité C-terminale et la région centrale sont fonctionnelles contrairement au domaine N-terminal.

Surexpression et Purification

La protéine IX- Δ SYL est surexprimée dans un système d'expression bactérien suivant la procédure décrite dans la section Matériels et Méthodes. La figure 3.9 montre les étapes de purification.

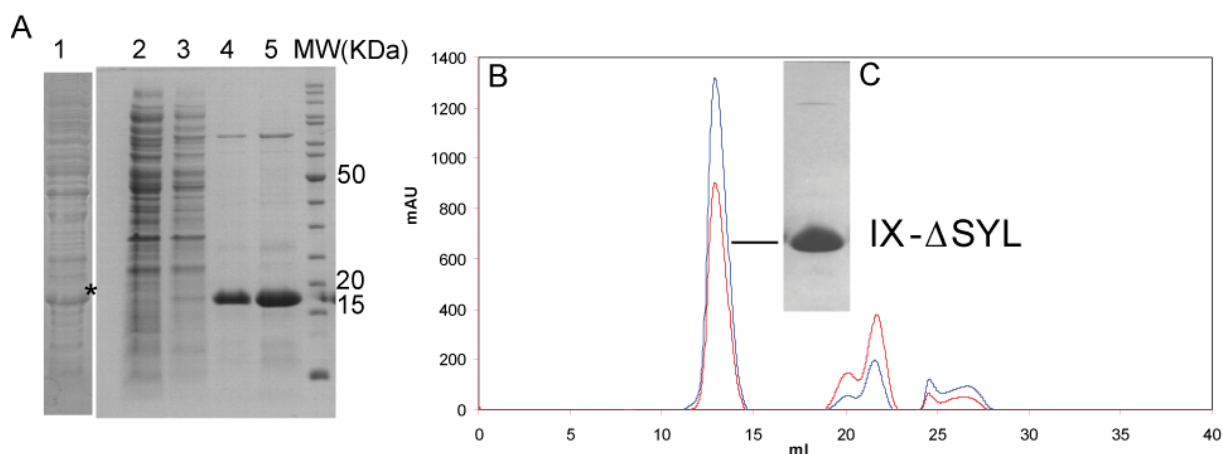


Figure 3.9: *Étapes de purification de la protéine IX- Δ SYL.*

A, Le gel SDS-PAGE dénaturant coloré au bleu de coomassie présente en piste 1: l'extrait brut; piste 2: la fraction non retenue sur la colonne d'affinité; piste 3: la fraction issue du lavage de la colonne avec 20 mM imidazole; piste 4 et 5: fraction d'élution avec 300 mM imidazole (dépôt de 2 et 5 μ l respectivement).

B, Profil d'élution de la protéine IX- Δ SYL sur une colonne d'exclusion de taille (Superdex 200). La protéine est éluee à un volume de 13 ml comme le montre le gel SDS-PAGE en C.

La protéine d'intérêt est exprimée en grande quantité, pure et homogène comme le suggère le pic symétrique obtenu lors de la gel filtration (Fig. 3.9B).

Cette protéine de 15.5 kDa (avec l'étiquette 6-His) est éluee avec un volume de 13 mL (comparé à 16 mL pour la protéine IX- Δ SYL-V117D) correspondant à une taille standard d'environ 50 kDa soit un état trimérique.

Etat d'Oligomérisation: MALLS-SEC et Pontage Chimique

Le MALLS-SEC nous indique qu'il y a au moins deux populations en solution, le trimère et l'hexamère (respectivement fin et début de l'élution) (Fig. 3.10A). A chaque état oligomérique devrait en effet correspondre un pic, or le profil d'élution montre un seul pic centré sur l'état trimérique. L'hexamère n'est donc probablement pas assez concentré pour générer un pic bien distinct. Le trimère majoritaire est naturellement formé par une interaction de forte affinité au niveau du domaine C-terminal, l'hexamère est généré par l'association de

faible affinité de deux trimères *via* le domaine central ou N-terminal, d'où sa présence en faible quantité.

Le pontage chimique de la protéine IX- Δ SYL suggère la présence de plusieurs populations en solution (Fig. 3.10B): le monomère, le dimère, deux types de trimères, plusieurs tétramères et une zone plus diffuse correspondant à des oligomères de hauts poids moléculaires. Le monomère et le dimère sont des produits de pontage intramoléculaire. Le dimère est aussi le produit d'un pontage du domaine central ou N-terminal (Fig. 3.8B). Le trimère provient majoritairement de l'interaction au niveau du domaine C-terminal mais peut aussi être le produit d'une combinaison de l'interaction des domaines C et N-terminaux d'où la présence de deux types de trimères (un léger décalage de migration sur le gel SDS-PAGE dû à deux formes différentes). Le même phénomène a lieu pour les tétramères (Fig. 3.10B). L'hexamère a été observé en faible proportion en MALLS (Fig. 3.10A) mais on ne voit pas de bande clairement attribuable à ce dernier sur le gel de pontage chimique. La stabilisation de l'hexamère par pontage covalent nécessite cinq réactions sur le même complexe ce qui diminue la probabilité de sa formation. La zone diffuse de haut poids moléculaire comporte probablement des hexamères (Fig. 3.10B).

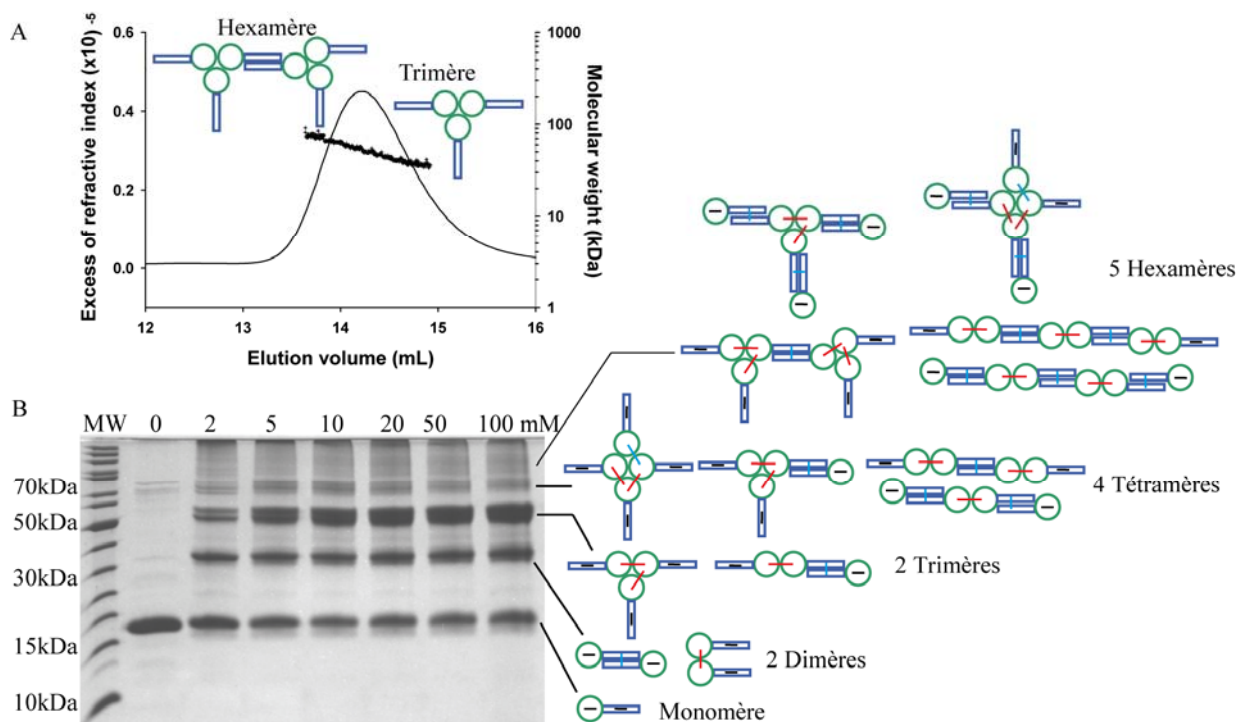


Figure 3.10: Etats d'oligomérisation de la protéine IX- Δ SYL.

A, Expériences de MALLS : La courbe représente le profil d'éluion sur une colonne d'exclusion de taille Shodex 804 KW. Les croix indiquent la masse moléculaire calculée en fonction du volume d'éluion. La masse moléculaire varie le long du pic du chromatogramme entre 40 et 80 kDa.

B, Expérience de pontage chimique : Les chiffres en haut de chaque piste présentent la concentration en glutaraldéhyde. La protéine entière est schématisée comme dans les figures 3.6 et 3.8.

Ces résultats confirment notre hypothèse selon laquelle deux domaines de la protéine IX sont capables de former des oligomères. Néanmoins, il reste encore à savoir si cette dimérisation se produit par le biais du domaine central ou N-terminal de la protéine IX.

d. La Protéine IX Δ SYL- Δ ala

Le mutant IX- Δ SYL- Δ ala de la protéine IX est délété de la région centrale riche en alanine (63-70 aa). Ce mutant a été surexprimé et purifié suivant une méthodologie identique à celles décrites précédemment, la figure 3.11 montre le profil d'élution sur une colonne d'exclusion de taille S75.

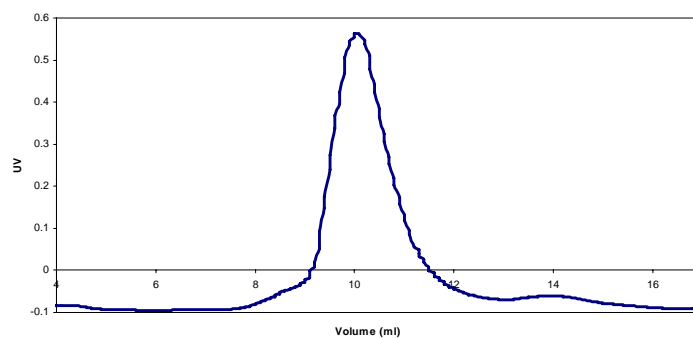


Figure 3.11: Profil d'élution de la protéine IX- Δ SYL- Δ ala sur une colonne d'exclusion de taille Superdex 75.

Selon le profil d'élution (Fig. 3.11), la protéine IX- Δ SYL- Δ ala est majoritairement trimérique. Cette protéine de 14.8 kDa est en effet éluée à un volume de 10 mL correspondant à un volume d'élution d'une protéine de 50 kDa soit un trimère.

Etat d'oligomérisation : MALLS-SEC et pontage chimique

En comparant les profils MALLS des protéines IX- Δ SYL- Δ ala (Fig. 3.12A) et IX- Δ SYL (Fig. 3.10A), nous avons constaté que la pente du graphe de la masse moléculaire est plus faible dans le cas de la protéine IX- Δ SYL- Δ ala (Fig. 3.12A), ceci est dû à l'hexamère qui est probablement présent en plus petite quantité chez ce mutant. L'hexamère est généré par dimérisation de deux trimères. Ces observations préliminaires laissent penser que le domaine central riche en alanine est impliqué dans cette dimérisation.

La comparaison des gels de pontage covalent de la protéine IX- Δ SYL- Δ ala (Fig. 3.10) et IX- Δ SYL (Fig. 3.12) indique la présence des mêmes populations oligomériques chez les

deux mutants. Elles sont cependant toutes présentes en plus petite quantité chez le mutant IX- Δ SYL- Δ ala (bandes moins intenses) (Fig. 3.10B et 3.12B). L'implication de la région centrale dans la dimérisation de la protéine IX explique parfaitement ce profil de pontage chimique: dans le cas de la protéine IX- Δ SYL, le signal est amplifié grâce à l'accumulation de plusieurs oligomères de même taille mais issus de plusieurs combinaisons d'interactions (domaine C-terminal et central) (Fig. 3.10B). Chez le mutant IX- Δ SYL- Δ ala la dimérisation est moins efficace et les oligomères sont donc issus majoritairement de la trimérisation (Fig. 3.12B) (c'est probablement pour cette même raison que les bandes sont également moins intenses lors du pontage chimique du domaine C-terminal seul ; Fig. 3.6B).

Ces résultats suggèrent que la région centrale est impliquée dans la dimérisation de la protéine IX bien que la délétion de cette région ne permette pas l'abolition totale de la formation de ces oligomères issus à la fois de la dimérisation et de la trimérisation (l'hexamère détecté par la MALLS et les deux types de trimère observés par pontage chimique ; Fig. 3.12).

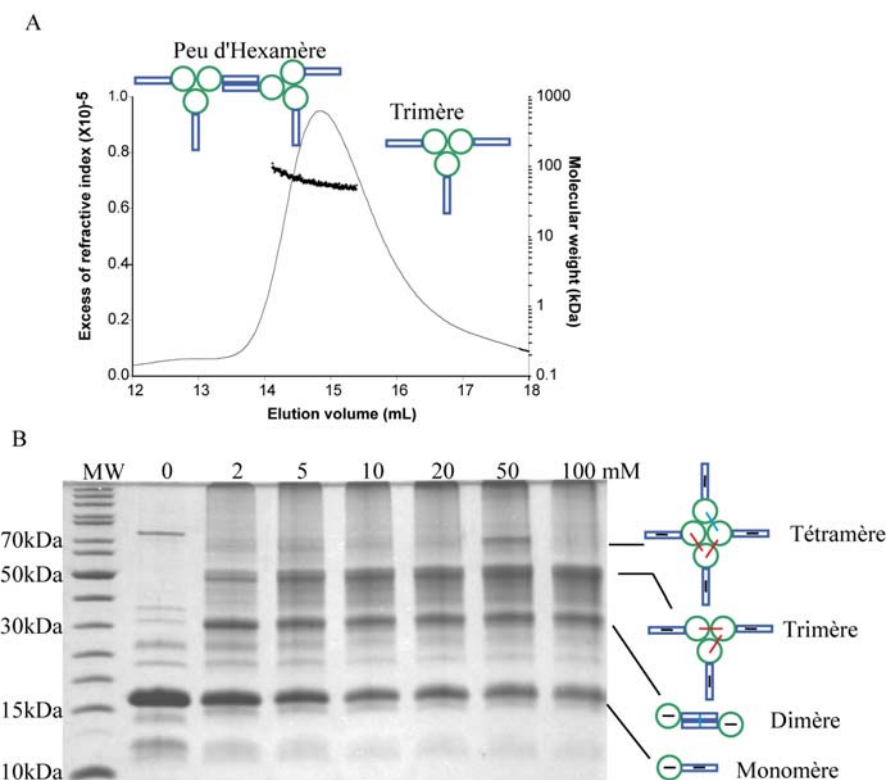


Figure 3.12: Etat d'oligomérisation du mutant IX- Δ SYL- Δ ala.

A, Expérience du MALLS-SEC: La masse moléculaire varie entre 50 et 100 kDa au cours du chromatogramme (la masse attendue pour le trimère est de 45 kDa).

B, Expérience de pontage chimique: Les concentrations en glutaraldéhyde sont indiquées en haut de chaque piste. Les différents oligomères sont schématisés comme dans la figure 3.10B. Les expériences de pontage chimique du mutant IX- Δ SYL- Δ ala et du mutant IX- Δ SYL ont été réalisées le même jour avec les mêmes concentrations en protéine et les mêmes lots de dilutions de glutaraldéhyde.

La région responsable de la dimérisation de la protéine IX est probablement située autour de la région centrale mais implique une plus grande région. Pour prouver cette hypothèse il faudra générer un mutant avec une délétion plus étendue de part et d'autre de la région centrale.

e. Conclusion

Ces expériences ont permis de prouver que la protéine IX possède deux domaines capables de s'oligomériser indépendamment:

- Le domaine C-terminal qui est responsable d'une interaction de forte affinité engendrant une trimérisation de la protéine IX et d'une interaction transitoire générant le tétramère.
- La région centrale qui forme un dimère *via* une interaction de faible affinité.

La possibilité et le degré d'oligomérisation du domaine N-terminal n'ont pas pu être déterminés. En effet, pour toutes les protéines que nous avons analysées, la séquence SYL a été éliminée du domaine N-terminal. Nous ne connaissons pas l'effet de cette délétion sur ce domaine de la protéine. Afin de compléter notre étude et en particulier différencier les deux extrémités de la protéine dans le virus, nous nous sommes intéressés à la structure de la protéine IX dans le virus.

3.3.2. Localisation du Domaine C-terminal de la Protéine IX dans la capside d'HA5

Ces travaux ont été réalisés par Guy Schoehn et Céline Fabry en collaboration avec le Dr. Manuel Rosa-Calatrava et le Pr. Pierre Boulanger (Lyon)

Des expériences ont été menées récemment au laboratoire pour déterminer la localisation du domaine C-terminal de la protéine IX à la surface de l'Ad humain de sérotype 5. Nous disposons d'un virus dans lequel la partie C-terminale de la protéine IX a été fusionnée avec un peptide SY12 (CTAYSSYMKGGKF). Le virus a été complexé à des Fabs ciblant ce peptide et une reconstruction 3D a été réalisée par microscopie électronique (Fig. 3.13).

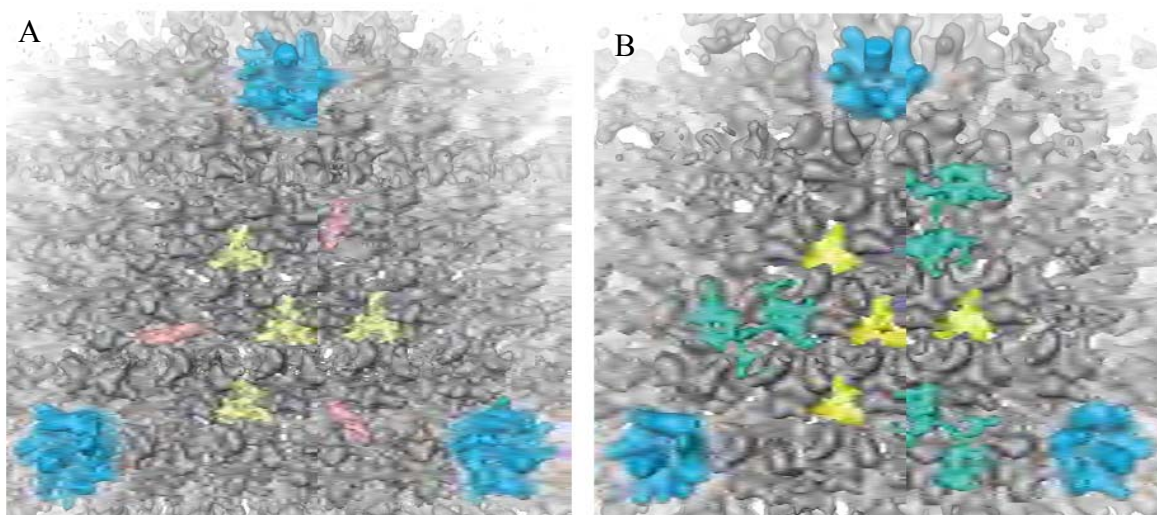


Figure 3.13: Reconstruction 3D en MET d'adénovirus humains de sérotype 5 dont les parties C-terminales des protéines IX sont fusionnées avec le peptide SY12.

A, Le virus modifié SY12, avec les parties C-terminales de la protéine IX en orange et les parties N-terminales en jaune. Les pentons sont en bleu et les hexons en gris.

B, Structure tridimensionnelle de l'adénovirus modifié neutralisé par des Fabs anti-SY12. Les parties N-terminales des protéines IX sont en jaune et les Fabs en vert.

La comparaison de la structure 3D du virus complexé aux Fabs avec celle du virus seul montre l'apparition d'une densité ressemblant très fortement à deux Fabs au dessus de la densité cylindrique récemment réattribuée au domaine C-terminal de la protéine IX par Saban et al. (2006) (*cf.* Introduction). Ces observations confirment la localisation du domaine C-terminal de la protéine IX entre deux hexons à l'interface entre deux facettes de l'icosaèdre.

3.3.3. Localisation de la Partie C-terminale de la Protéine IX dans la Capside du CAV-2

a. Introduction

L'alignement de séquence de la protéine IX d'HA5 et celle du CAV-2 indique :

- 1- La conservation du domaine N-terminal.
- 2- L'absence du domaine central riche en alanine chez la protéine IX du CAV-2.
- 3- La séquence du domaine C-terminal est moins conservée que celle du domaine N-terminal entre les deux virus.

Afin d'étudier la localisation de la protéine IX dans la capsid de l'adénovirus canin et de vérifier si la différence dans la séquence se traduit par des différences d'organisation à la surface de la capsid, la structure 3D de CAV-2 a été obtenue à 12 Å de résolution par reconstruction 3D à partir d'images obtenues par CryoMET. Dans un second temps la reconstruction 3D du CAV-2 contenant la protéine IX dont la partie C-terminale a été fusionnée à la GFP a été obtenue à une résolution d'environ 25 Å. Cette première reconstruction 3D d'adénovirus non humain a permis une comparaison avec la capsid des Ads humains. La comparaison de la structure de CAV-2 et CAV-2-IX-GFP a permis la localisation de la partie C-terminale de la protéine IX à la surface de la capsid.

b. ARTICLE I: Three-dimensional structure of canine adenovirus serotype 2 capsid

Auteurs: G. Schoehn, M. El Bakkouri, C. M. S. Fabry, O. Billet, L. F. Estrozi, L. Le, D. T. Curiel, A. V. Kajava, R. W. H. Ruigrok et E. J. Kremer.

Publié en 2008 dans *journal of Virology*, Vol. 82, pages 3329-3341.

Three-Dimensional Structure of Canine Adenovirus Serotype 2 Capsid [▽]

Guy Schoehn,^{1,2*} Majida El Bakkouri,¹ Céline M. S. Fabry,¹ Oliver Billet,^{3,4} Leandro F. Estrozi,¹ Long Le,⁵ David T. Curiel,⁵ Andrey V. Kajava,⁶ Rob W. H. Ruigrok,¹ and Eric J. Kremer^{3,4*}

Unit for Virus Host Cell Interaction, UMR 5233, Université Joseph Fourier, EMBL, CNRS, Grenoble, France¹; Institut de Biologie Structurale Jean-Pierre Ebel, UMR5075 CEA-CNRS-UJF, Grenoble, France²; Institut de Génétique Moléculaire de Montpellier, CNRS 5535, Montpellier, France³; Université Montpellier I and II, Montpellier, France⁴; Division of Human Gene Therapy, Departments of Medicine, Obstetrics and Gynecology, Pathology, and Surgery, and the Gene Therapy Center, The University of Alabama at Birmingham, Birmingham, Alabama 35294⁵; and Centre de Recherche de Biochimie Macromoléculaire, UMR 5237 CNRS-Université Montpellier I and II, Montpellier, France⁶

Received 6 November 2007/Accepted 28 December 2007

There are more than 100 known adenovirus (AdV) serotypes, including 50 human serotypes. Because AdV-induced disease is relatively species specific, vectors derived from nonhuman serotypes may have wider clinical potential based, in part, on the lack of ubiquitous memory immunity. Whereas a few of the human serotype capsids have been studied at the structural level, none of the nonhuman serotypes has been analyzed. The basis laid by the analysis of human AdV (hAdV) has allowed us to determine and compare the three-dimensional structure of the capsid of canine serotype 2 (CAV-2) to that of hAdV serotype 5 (hAdV-5). We show that CAV-2 capsid has a smoother structure than the human serotypes. Many of the external loops found in the hAdV-5 penton base and the hexon, against which the antibody response is directed, are shorter or absent in CAV-2. On the other hand, the CAV-2 fiber appears to be more complex, with two bends in the shaft. An interesting difference between the human and canine viruses is that the C-terminal part of protein IX is in a different position, making an antenna sticking out of the CAV-2 capsid. The comparison between the two viruses allows the identification of sites that should be easy to modify on the CAV-2 capsid for altering tissue tropism or other biological activities.

Adenoviridae are nonenveloped icosahedral viruses with 26- to 44-kb double-stranded DNA genomes (26). They infect a variety of vertebrates including mammals, fish, birds, and reptiles (12). The more than 50 human adenovirus (hAdV) serotypes are the most extensively studied because of their ability to induce respiratory, ocular, and enteric infections in immunocompetent individuals. hAdV infections are a significant cause of morbidity and mortality in newborns and immunocompromised individuals (28). AdVs are also used for gene therapy trials and anticancer treatments (6) although the widespread immunity against human serotypes severely compromises their use as vectors. Because AdV-induced morbidity is relatively species specific, vectors derived from nonhuman AdVs may be more clinically useful than those derived from human serotypes, based, in part, on the potential lack of ubiquitous memory immunity (cellular and humoral) (5, 11, 18, 30, 35, 36).

One of the best characterized nonhuman AdVs is the canine serotype 2 (cAdV-2, commonly referred to as CAV-2) (29). We previously showed that CAV-2 vectors could lead to preferential transduction of neurons *in vivo* (49), long-term (>1 yr) expression in the brain (50), and a lower cellular immune response (27, 49). The majority of donors in a random cohort harbored no or only low levels of anti-CAV-2 neutralizing antibodies (30, 37), and less

than half of this cohort harbored anti-CAV-2 memory T cells (38). When the donors harbored an anti-CAV-2 cellular response, it was on the average 10-fold lower than the response to hAdV serotype 5 (hAdV-5). Finally, in contrast to hAdV-5-based vectors, CAV-2 vectors poorly transduce or mature human dendritic cells or activate the complement cascade (37, 40), both of which play pivotal roles in orchestrating and bridging innate, adaptive, and memory immune responses. These findings suggest that CAV-2 vectors may have specific clinical advantages, especially in terms of safety (39).

However, there exist certain enigmas concerning various aspects of CAV-2 biology. For example, the *in vitro* paradigm for AdV tropism is predominately based on studies using hAdV-2 or hAdV-5 in epithelial cells. Initially, the fiber knob binds to the coxsackievirus AdV receptor (CAR) (4, 52). Then, a conserved Arg-Gly-Asp (RGD) motif in the penton base binds to integrins, triggering endocytosis (24). Although CAV-2 uses CAR to infect cells (48), the capsid does not contain an identifiable integrin-interacting motif, yet it traffics in epithelial cells with similar kinetics (8). Furthermore, although both hAdV-5 and CAV-2 use CAR, CAV-2 preferentially transduces neurons when injected into the brain while hAdV-5 preferentially transduces astrocytes (1, 32). CAV-2 is also capable of an impressive level of axonal transport in neurons that can be up to 100-fold greater than that of hAdV-5 (49). In addition, in contrast to previous studies with hAdV-5 (16, 17), we have been unable to complex CAV-2 virions with divalent cations or lipoplexes (unpublished data).

To resolve these and other CAV-2 enigmas, we need to better understand the capsid structure. Based on the data from different hAdVs, the AdV capsid structure appears to be rel-

* Corresponding author. Mailing address for Guy Schoehn: Unit for Virus Host Cell Interaction, 6 rue Jules Horowitz, BP181, 38042 Grenoble, France. Phone: 33 4 7620 9423. Fax: 33 4 7620 7199. E-mail: schoehn@embl.fr. Mailing address for Eric J. Kremer: Institut de Génétique Moléculaire de Montpellier, CNRS 5535, 1919 Route de Mende, 34293 Montpellier, France. Phone: 33 4 6761 3672. Fax: 33 4 6704 0231. E-mail: eric.kremer@igmm.cnrs.fr.

[▽] Published ahead of print on 23 January 2008.

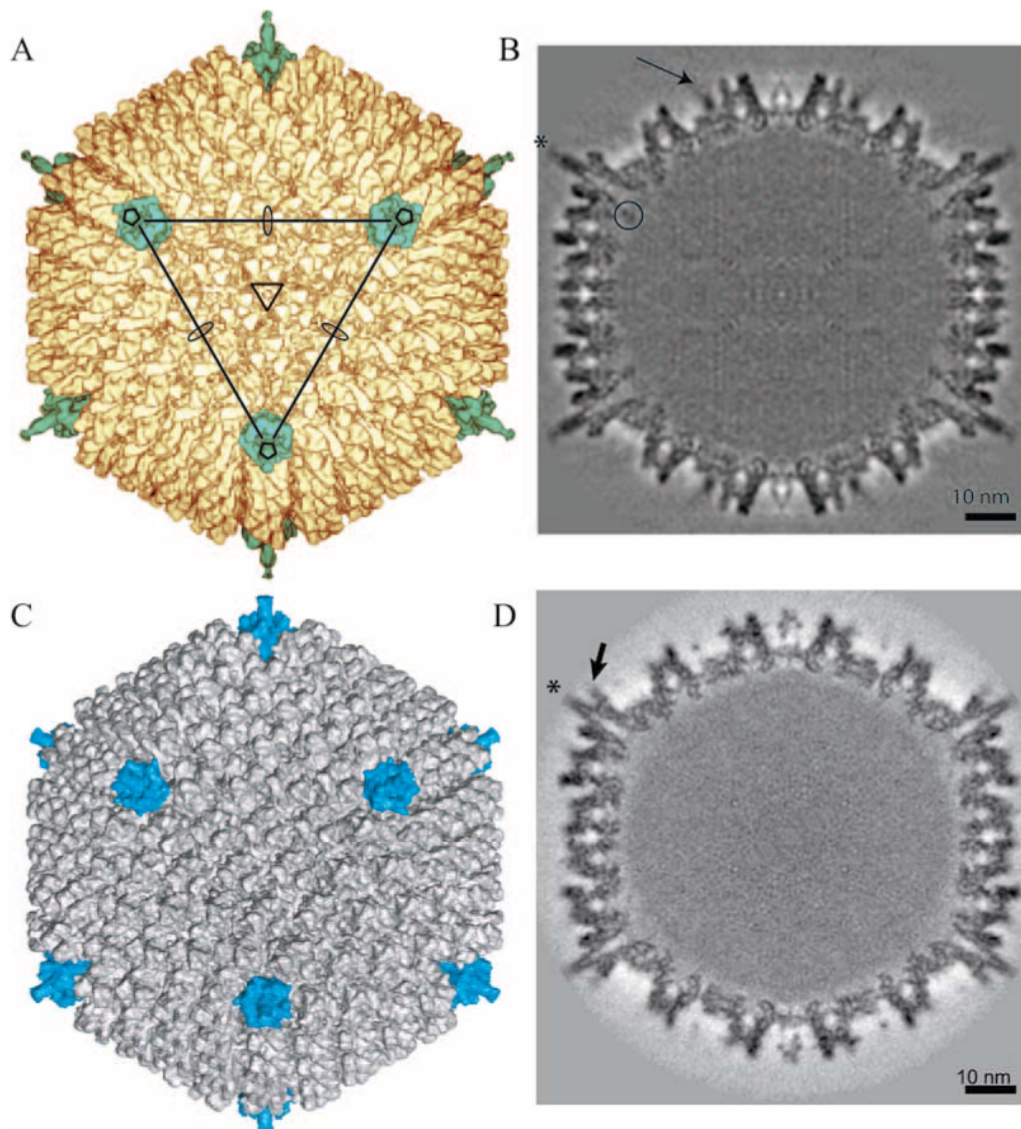


FIG. 1. 3D reconstruction of CAV-2 at 12-Å resolution and comparison with the previously determined structure of hAdV-5 (MSD accession number 1111) at 10-Å resolution. (A) Isosurface representation of CAV-2 as viewed down the threefold axis. The pentons are shown in petrol blue, and the rest of the capsid is shown in yellow. The icosahedral axes are indicated by an ellipse (twofold), a triangle (threefold), and a pentagon (fivefold). A continuous line connecting the three pentons pointing in your direction delimits a facet. (B) A central slice through the CAV-2 EM reconstruction showing that the capsid density is well defined. One of the fibers sticking out radially from the capsid is indicated by a star. A circle surrounds the nonattributed extra density present under the penton base. The arrow indicates protein IX density. Scale bar, 10 nm. (C) Isosurface representation of hAdV-5 view down the threefold axis. The pentons are shown in blue, and the rest of the capsid is shown in gray. (D) A central slice through the hAdV-5 EM reconstruction showing that the capsid density is well defined. The fibers sticking out radially from the capsid are indicated by a star. The RGD loop is indicated by an arrow. Scale bar, 10 nm.

atively well conserved. The capsid contains 7 out of the 12 polypeptides present in the virion. The major structural components are the trimeric hexons, 240 of which form the 20 facets of the icosahedron (51). At each of the 12 vertices are a pentameric penton base and an externally projecting trimeric fiber. The atomic structures of these three individual components are known for hAdV-2 or hAdV-5 (44, 54, 58).

The hexon trimer has a pseudohexagonal base with three towers extending upwards. The triangular top is rotated $\sim 10^\circ$ counterclockwise with respect to the hexagonal base. The loops extending from the irregularly shaped towers (44) have not

been solved by X-ray crystallography, although they have been visualized and modeled in recent cryoelectron microscopy (cryo-EM) structures (15, 45).

In contrast to the hexon, the penton base is wide at the top (58). It has an overall tulip shape, with each of its five petals having a right-handed twist around the fivefold axis. The top of the base is almost flat with the exception of two protuberances. One of these extensions, which are variable in size depending on the serotype, contains the above-mentioned integrin-interacting motif. The other extension is a hypervariable loop that also varies in size depending on the serotype (20).

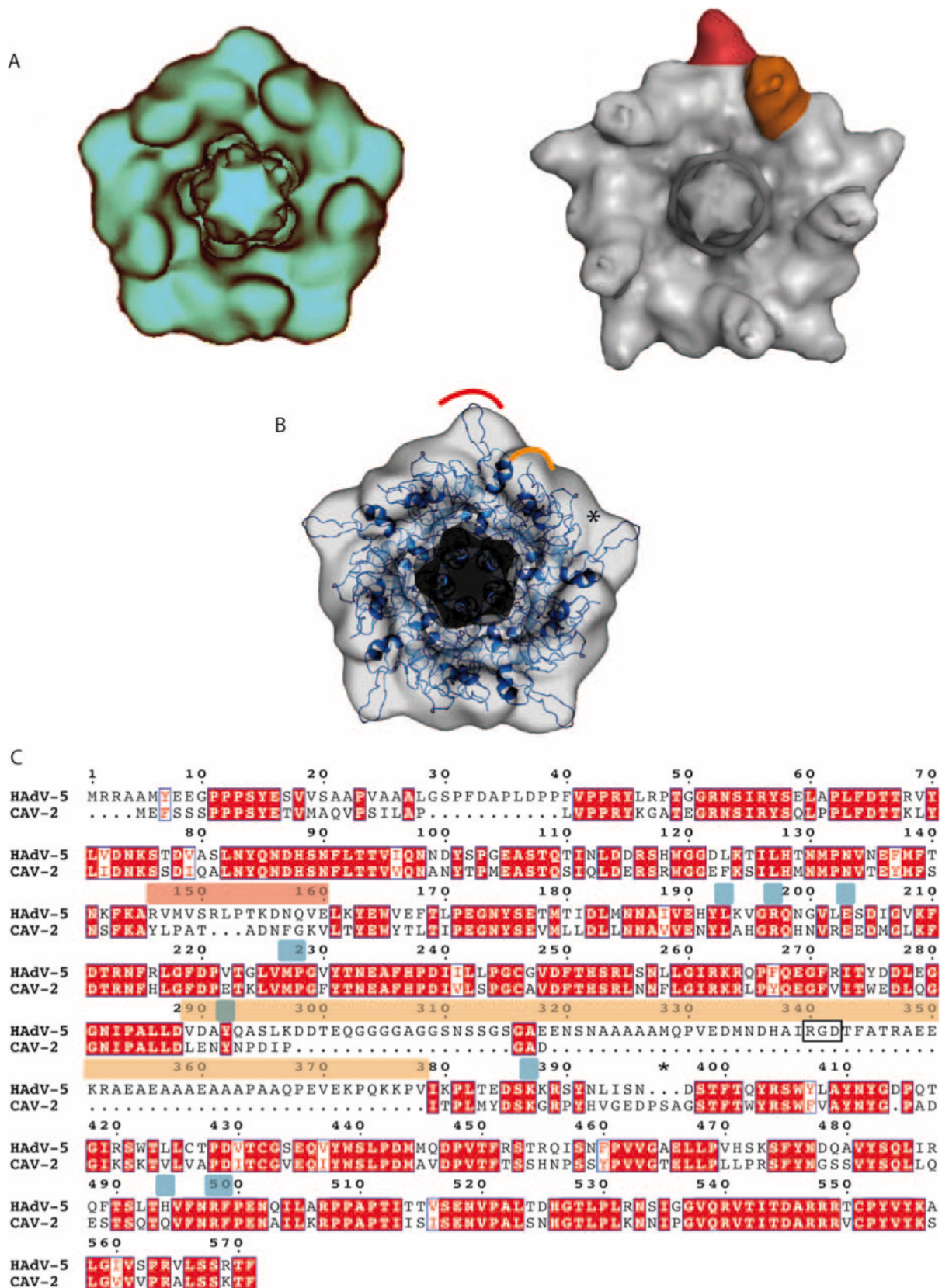


FIG. 2. The CAV-2 penton base. (A) Comparison between the computationally isolated penton from the CAV-2 capsid in blue (left) and the hAdV-5 capsid in gray (right). One of the five hypervariable regions is highlighted in red, and one protuberance harboring the RGD sequence is shown in orange in the hAdV-5 penton. (B) Fitting of the hAdV-2 penton base X-ray structure in blue (PDB accession number 1X9T) into the 12-Å CAV-2 cryo-EM envelope. The red and orange parts correspond to the regions indicated in panel A. The star indicates a region where the

The trimeric AdV fiber contains a highly conserved N-terminal sequence that binds to the base, a thin shaft of variable length, and a globular knob at the distal end. The length, the flexibility of the fiber shaft, and the sequence of the knob domain also vary between serotypes (9). Four of the 23 repeats of the hAdV-2 fiber shaft were crystallized, and the resolved structure demonstrated a novel triple β -spiral (54). Notably though, the entire fiber was not crystallized, probably due to its flexible structure. The number of repeats can vary from a minimum of five for hAdV-3 to at least 46 for bovine AdV-3 (BAV-3) (43).

Apart from these major proteins, the capsid is stabilized by hexon-associated proteins (51). Using high-resolution cryo-EM three-dimensional (3D) reconstructions combined with secondary structure predictions, a new location of the minor structural proteins has been proposed. Protein IX forms a connected complex off our trimers in the middle of each facet (45). Protein IIIa is located inside the capsid underneath the penton base, connected to its N-terminal part (45). A total of 120 copies of protein VIII are located on the inner face of the capsid, and each makes the link between four hexon trimers around the three- and fivefold axes (15).

Here, we have determined the 3D structure of the CAV-2 capsid and compared it to the known X-ray and EM structures of hAdV-5 or hAdV-2 by fitting the different atomic resolution structures into the EM density. Globally, we found that CAV-2 capsid has a smoother structure than the human serotypes. Many of the external loops found in hAdV-2 and hAdV-5 (hAdV-2/5) are either shorter or absent in CAV-2. On the other hand, the CAV-2 fiber appears to be more complex, with two bends in the shaft. The knowledge of the structure will allow the identification of sites where functional loops may be added to CAV-2 to potentially modify tissue tropism.

MATERIALS AND METHODS

CAV-2 vectors and recombinant fiber. CAVGFP and CAVpIX-GFP were prepared as previously described (30, 31). Briefly, CAVGFP is a CAV-2 vector with a deletion of the E1 region and with a green fluorescent protein (GFP) expression cassette. The capsid contains no modifications. CAVpIX-GFP is a vector containing GFP fused to the C-terminal end of protein IX. All vectors were purified by double banding on CsCl gradients, and CsCl was removed using PD-10 columns (Pharmacia). The vectors were stored in phosphate-buffered saline containing 10% glycerol. Stocks were 2×10^{12} physical particles/ml with at least 1 infectious particle/5 physical particles. The approximately 1.7-kb CAV-2 fiber coding region was cloned in pBlueBac4.5/V5-His-TOPO (Invitrogen). Recombinant baculovirus was generated, isolated, amplified, and purified following the manufacturer's protocol. The His-tagged fiber was purified using standard protocols.

Negative staining. Four microliters of the recombinant CAV-2 fiber sample (~ 0.1 mg/ml) was loaded in the mica-carbon interface as described previously (46). The sample was stained using 2% sodium silico tungstate, pH 7.5, and air dried. Images were taken under low-dose conditions in an EX1200-II JEOL electron microscope working at 100 kV and with a nominal magnification of 40,000. The images were scanned on a Z/I Imaging scanner (Photoscan TD) with a pixel size of 14 μm (3.5 \AA per pixel at the sample level). A total of 800 fibers

were selected using X3d (10), cut into 200- by 200-pixel squares, transferred into the SPIDER program (19), and subjected to classification and averaging.

Cryo-EM. For cryo-EM, the glycerol was removed, and the sample was concentrated to ~ 1 mg/ml. Four microliters of sample was loaded onto a Quantifoil R2/1 holey grid (Quantifoil Micro Tools GmbH, Germany), blotted for 1 to 2 s to remove the excess liquid, and then rapidly plunged into liquid ethane cooled using liquid nitrogen. The frozen grid was transferred into an LaB6 CM200 Philips electron microscope using a GATAN 626 cryoholder. The images were taken under low-dose conditions at 200 kV (less than $10 \text{ e}^-/\text{\AA}^2$) and with a nominal magnification of 27,500. Each area was imaged twice using two different defocus values, starting with the image closest to focus. The negatives were developed in full-strength D19 developer for 12 min.

Image analysis. Negatives were screened for astigmatism and drift by optical diffraction, and only those showing information up to 12 \AA were digitized using the Photoscan TD with a pixel size of 7 μm (2.54 \AA per pixel at the specimen level as calibrated using tobacco mosaic virus). The particles were selected interactively using X3d (10) and boxed into 447- by 447-pixel squares with a circular mask applied of radius 219 pixels plus four fade pixels. The images were corrected for the contrast transfer function effect, as described previously, using the program CTFMIX (10). A total of 5,350 particles from 21 negatives were used in the image analysis process. Determination of particle origin and orientation were performed with the model-based polar Fourier transform (PFT) programs using the previously determined 3D structure of hAdV-5 (3, 15). After PFT refinement was completed, a supplementary cycle was carried out using a new version of PFT, PFT2 (D. M. Belnap, J. B. Heymann and J. F. Conway, unpublished results). The final reconstruction was obtained using the PFT2 and EM3DR2 programs with a total of 2,780 particles (about half of the total). The resolution of the final map was estimated to 12 \AA by Fourier shell correlation (53) calculated between independent half-data set maps and by applying a correlation limit of 0.3 (not shown). The same process was essentially used for CAVpIX-GFP, but we used Quantifoil grids covered by a thin layer of continuous carbon to increase the visible virus concentration. Only 800 particles were used in this reconstruction, which limited the resolution to 25 \AA . Map visualization was performed using the programs WEB (19) and PYMOL (14), and the contour level was chosen to include an average protein density of $0.84 \text{ Da}/\text{\AA}^3$. The EM maps were uploaded into the EM database associated with the Macromolecular Structure Database (see below).

Fitting of the atomic structures of hexon and penton base into the EM density. The atomic structure of the hAdV-5 hexon and that of the hAdV-2 penton base were placed manually into the corresponding CAV-2 EM densities using the program PYMOL (14). These manual fits were then optimized using SITUS (55).

Molecular modeling of the 3D structure of CAV-2 fiber. To model the 3D structure of the CAV-2 fiber, we used the crystal structure of the homotrimeric C-terminal CAV-2 fiber head (residues 361 to 542) (47) and two known crystal structures of homologous proteins: reovirus attachment protein $\sigma 1$ (7) and the C-terminal part of the hAdV-2 fiber containing four repeats of the shaft domain followed by the head domain (54). The hAdV-2 structure allowed us to model the last four repetitive elements of the CAV-2 fiber shaft and the region where this shaft is linked to the head. The N-terminal repetitive structure of the CAV-2 fiber shaft was also modeled by using the β -spiral fragment of the known hAdV-2 structure as a template and in accordance with the repeat alignment shown on Fig. 3. The two longer repeats (see Fig. 3C, repeats 4 and 10) of hAdV-2 having 7- and 6-residue insertions in the middle of the repeats were modeled by using the crystal structure of reovirus attachment protein $\sigma 1$ (7). The $\sigma 1$ protein has one such long repeat containing a similar 5-residue insertion. The model was built by using the Insight II program (13).

Comparison with hAdV-5. All comparisons between the CAV-2 and the hAdV-5 EM structures were done using the hAdV-5 3D reconstruction previously determined in the laboratory and freely available in the EM database associated with the Macromolecular Structure Database (<http://www.ebi.ac.uk/msd-srv/emsearch/index.html>) under accession number 1111.

X-ray structure did not fill the EM density. (C) Sequence alignment of the hAdV-5 and CAV-2 penton bases visualized with ESPrpt (23). The RGD and hypervariable loops are indicated by bars above the sequence in the same colors as used in panels A and B. The RGD sequence is boxed with a black rectangle. Conserved residues are in white text and highlighted in red; partially conserved residues are in red text. The star highlights three extra amino acids present in CAV-2 compared to hAdV-5 that could fill up the density indicated by the star in panel B. The amino acids involved in fiber binding for hAdV-5 are indicated by blue bars above the sequence.

Sequence alignments and structure prediction. All sequence alignments were carried out using CLUSTAL (<http://www.ebi.ac.uk/clustalw/>). Secondary structure predictions were made using PREDATOR (<http://bioweb.pasteur.fr/seqanal/interfaces/predator-simple.html>) and coil-coiled predictions were made using <http://www.russell.embl-heidelberg.de/cgi-bin/coils-svr.pl>.

Protein structure accession numbers. EM density maps have been deposited in the Macromolecular Structure Database at the European Bioinformatics Institute under accession numbers 1462 and 1463 for CAV-2 and CAVpIX-GFP, respectively.

RESULTS AND DISCUSSION

EM model of CAV-2. We determined the 3D structure of CAV-2 capsid by cryo-EM at 12-Å resolution. The dimension and the thickness of the CAV-2 capsid shell are similar to those of hAdV-5 (Fig. 1). The icosahedral shape is well defined and exhibits the pseudo $T=25$ triangulation number imposed by the trimeric hexon protein (Fig. 1A and B). There are 12 triangular hexons per facet and one penton on each vertex. The hexons are packed in an orientation similar to that of hAdV-5, including the slight tilt at each of the 20 edges of the virion (30° versus 10° for the other hexons belonging to the same facets). The penton base also has roughly the same shape in both capsids. The CAV-2 fiber is flexible, which explains why only part of it was visible in the 3D reconstruction (Fig. 1B, star). Finally, densities are present in the middle of the facet between the hexons. This density, attributed to protein IX, exhibits a triskelion-like shape, which is similar to that of hAdV-5 (Fig. 1; see also Fig. 5).

The CAV-2 penton base is smooth. Sequence alignment between the hAdV-5 and CAV-2 penton bases shows that 310 of the 571 amino acids are identical and that 14 amino acids are similar between the two species. However, the CAV-2 penton base sequence is $\sim 20\%$ shorter than the hAdV-5 sequence (477 versus 571) (Fig. 2C). Nonetheless, the overall shape of the two pentagons is similar (Fig. 2A). A slice through the density (Fig. 1B) shows a cone-like geometry that fits perfectly into the vertices surrounded by the hexons.

The CAV-2 penton base isosurface representation is smoother than that of hAdV-2/5 (Fig. 2A). By comparing the EM structures (Fig. 2A), we identified two protuberances that are smaller in the CAV-2 base (Fig. 2A, right panel, red and orange areas). The atomic structure of the hAdV-2 penton base (99% identical to that of hAdV-5) was fitted into the CAV-2 penton base EM density in order to identify the residues that are responsible for the smooth appearance of the CAV-2 capsid (Fig. 2B). The first difference is located in the hypervariable loop (Fig. 2A, red area, and C, sequence indicated by a red bar), which is 3 amino acids shorter in the CAV-2 base. On average, the amino acids composing this loop are also smaller in size (PATADNFG) (Fig. 2C). The second difference (Fig. 2, orange area) lies on the protuberances located on the top of the base that contains the RGD sequence in most hAdV (residues 289 to 379 for hAdV-5) (Fig. 2C). Due to its flexibility, the entire RGD loop structure is not visible either in the atomic structure of the hAdV-2 penton base or in the hAdV-5 cryo-EM structure. This protuberance is also present in the CAV-2 base but is much less pronounced because it is only 12 amino acids long. The CAV-2 loop does not contain an RGD sequence. Despite the small size of the protuberance, it is clearly visible in the cryo-EM map. This rein-

forces the idea that the N-terminal part of the fiber interacts with the base of the RGD loop (20) and contributes to the size of the protuberance.

Next to the orange loop in the CAV-2 penton base there are three extra amino acids after the RGD loop, indicated by a star in the fit shown in Fig. 2B and in the sequence in panel C. This corresponds to an empty region in terms of both the fit and sequence (Fig. 2B and C, stars). This could easily explain the smoother pentagonal shape of the CAV-2 base (slightly larger on the middle of the edge and smaller on the pentagon vertex).

According to our sequence alignment (Fig. 2C), the N-terminal part of the CAV-2 penton base is also 15 amino acids shorter than that of hAdV-5. The currently unresolved N-terminal part of the penton base (amino acids 1 to 49) contains two highly conserved PPXY (where X is any residue) motifs implicated in interactions with host cell WW domain-containing ubiquitin ligases (21, 22). In the mature virion, this region is probably located under the penton base, where it interacts with other viral capsid proteins (20, 45). The PPXY sequences are also present in CAV-2, suggesting that these sequences are critical for some phase in the AdV life cycle (22). The 15-amino-acid difference is too small to be interpreted in the EM map because it is in contact with other proteins on the inside of the virus capsid.

On the CAV-2 penton base, 9 out of the 10 residues involved in fiber tail-penton base interaction are conserved compared to hAdV-5 (Fig. 2C, indicated by blue bars). The only difference is that histidine 494 of hAdV-5 is a glutamine in CAV-2. The fiber sequence that interacts with the base is also conserved: amino acids 13 to 19 follow the conserved sequence seen in hAdV (58). The binding of the fiber to the penton base should therefore be very similar to that described for hAdV-2 (58).

The CAV-2 fiber contains two bends. The CAV-2 fiber is 542 amino acids long: the first 41 amino acids make up the tail, the central 320 amino acids comprise the shaft, and the last 182 amino acids form the knob (Fig. 3C). In hAdV-5 each of the ~ 15 -amino-acid repeats is 13 Å long (42) and induces a rotation of $\sim 50^\circ$ along the fiber axis (54). Based on sequence analysis, Rasmussen et al. (41) predicted that the CAV-2 shaft had 18 repeats with 18 amino acids per repeat. Our sequence analysis suggests that there are 18.5 repeats (Fig. 3C). The number of amino acids/repeat varies: seven repeats have the quintessential 15-amino-acid motif, the two repeats with 16 amino acids have the "extra" amino acid on the outside of the repeat, the three repeats with 17 amino acids have the extra amino acids inside the repeat, one repeat has 18 amino acids (extra internal), and two repeats have 19 amino acids (extra external). Two of the three longest repeats (21 to 22 amino acids) contain six or seven extra internal amino acids. All hAdV fiber sequences contain an irregular, long repeat near the N-terminal attachment site on the penton base, which has been linked to a possible bend at this position (42). However, this irregular repeat tends to be the third repeat and not the fourth, as is the case for the CAV-2 fiber (9). The much longer fibers of BAV-3 and the avian chicken embryo lethal orphan virus have additional bends, also related to irregular repeats (25, 43). Wu et al. (57) showed that such a 6- to 7-amino-acid internal insert can introduce a bend in the hAdV-5 shaft that may allow greater access of the penton base's integrin-interacting motifs to the cell surface (56).

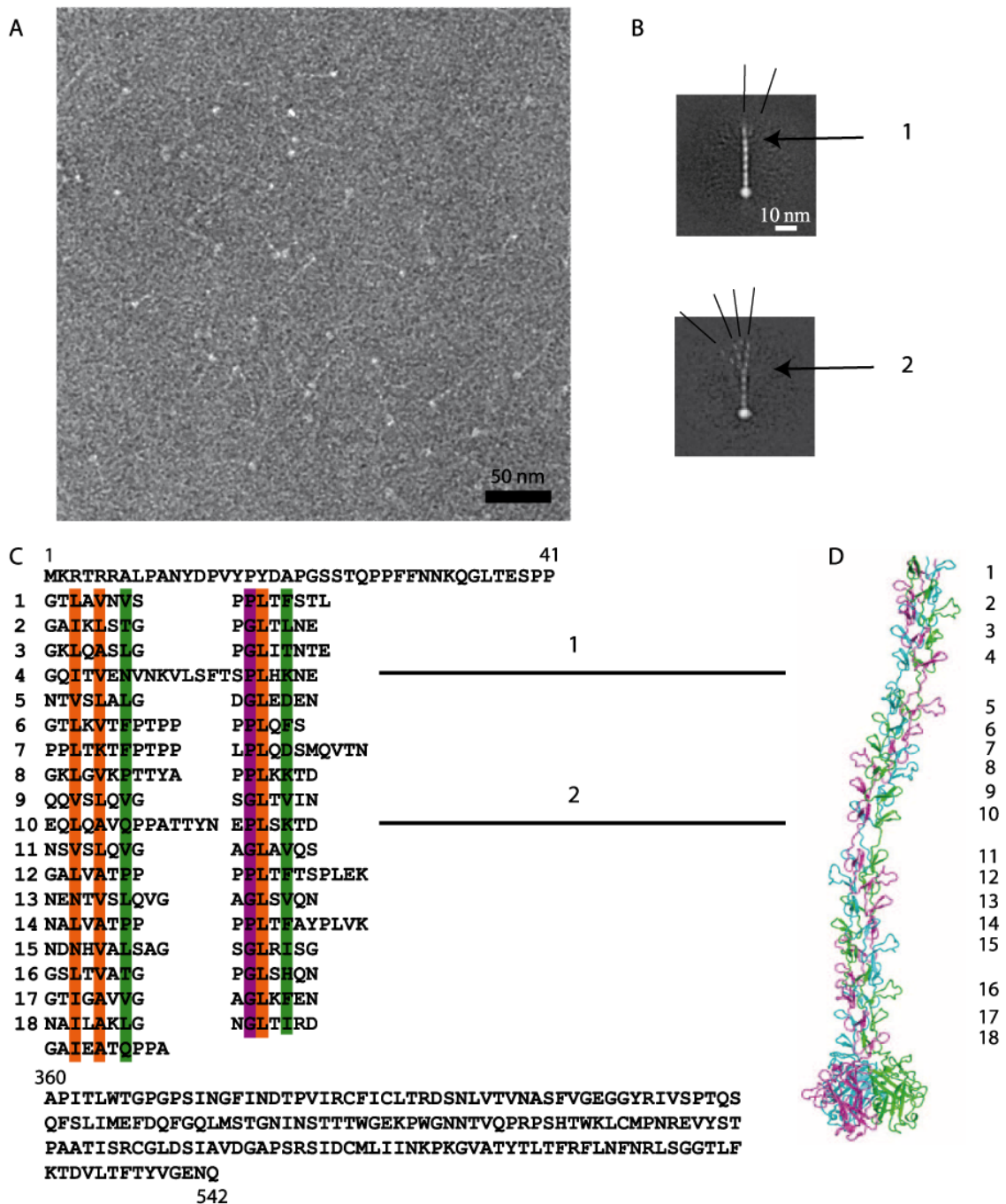


FIG. 3. The CAV-2 fiber. (A) Electron micrograph of negatively stained CAV-2 fibers. The thin shafts and their globular knobs are clearly visible. Scale bar, 50 nm. (B) Averaged image of fibers selected from micrographs like that shown in panel A. The fiber is generally straight but can exhibit two kinks as indicated by the numbers 1 and 2. The different lines show the angle that the N-terminal part makes with the rest of the shaft. (C) Sequence of the CAV-2 fiber shaft showing the repeat motifs that correspond to the location of the bends. The top part of the sequence is the N-terminal tail of the fiber, and the lower part is the C-terminal knob. The repeat residues involved in the hydrophobic core are in orange, the ones forming the peripheral hydrophobic patches are in green, and the conserved glycine or proline residues are highlighted in purple according to van Raaij et al. (54). (D) Hypothetical model of the CAV-2 fiber showing the predicted bends at repeats 4 and 10. The repeats are indicated with the same numbers as in the sequence shown in panel C.

As mentioned previously, the CAV-2 fiber was too flexible to be reconstructed by cryo-EM on the entire virion. To complete our structural analysis, we imaged recombinant CAV-2 fibers by negative staining (Fig. 3A). After image averaging (the fiber

head and the apical part of the different fibers are superposed but their N-terminal parts show the different positions they can adopt), we also found at least two bends in the shaft (Fig. 3B). The location of the hinges correlates well with the 21- and

22-amino acid-long repeats with six to seven internal amino acids (Fig. 3B, repeats 4 and 10).

To test the agreement of the EM data with the (i) suggested repeat alignment, (ii) dimensions of the predicted atomic structure, and (iii) locations and properties of the kink regions, we modeled the 3D structure of the fiber (Fig. 3D). Recently, the crystal structure of the homotrimeric C-terminal CAV-2 fiber head (residues 361 to 542) was determined (47). We also took advantage of two other known crystal structures of homologous proteins corresponding to different parts of the fiber to model its overall structure. The first structure that we used as a template was the C-terminal part of the hAdV-2 fiber containing four repeats of the shaft domain followed by the receptor-binding head domain (54). This structure allowed us to model the last four repetitive elements of the CAV-2 fiber shaft and the region where this shaft is linked to the head domains. Most of the CAV-2 repeats are similar to those in hAdV-2, and, therefore, the crystal structure of the hAdV-2 repeats was used to model the remaining part of the CAV-2 fiber. The exceptions were repeats 4 and 10 (Fig. 3C) having 7- and 6-residue insertions between two conserved β -strand motifs. These repeats were modeled by using the crystal structure of the reovirus attachment protein $\sigma 1$ (7). In the structure of $\sigma 1$, one repeat contains a similar 5-residue insertion. This insertion interrupts the progression of the regular β -spiral and introduces a bend into the $\sigma 1$ trimer. The structure of the insertion is loose, with slightly different main chain conformations in each monomer extended along the fiber axis. This structure was used to model the two longer repeats of CAV-2.

If our model is correct, the CAV-2 trimeric fiber should be about 330 Å long. The β -spiral is 280 Å long with two flexible bends and a C-terminal globular head domain contributing 50 Å (Fig. 3D). This model, with respect to the number and location of the bends and the length, is in a good agreement with our EM data.

The CAV-2 hexon. The general dimensions and twist of the homotrimeric CAV-2 hexon are very similar to those of hAdV-5 (Fig. 4A). Globally, the CAV-2 hexon sequence is shorter than that of hAdV-5 (Fig. 4C) (908 versus 952 amino acids) with all the differences located in the N-terminal half. The fit into the EM density of the CAV-2 hexon was carried out using the hAdV-5 X-ray structure (Protein Data Bank [PDB] 1P30) (Fig. 4B). Because the loops that are absent in the hAdV-5 X-ray structure (Fig. 4C, blue bars above the sequence) often correspond to a deletion in the CAV-2 sequence, the fit was good, and only one loop stays outside the density (Fig. 4B, left). This loop (amino acids 194 to 202 in

hAdV-5 versus 153 to 161 in CAV-2) was visible in the atomic structure of the hAdV-5 hexon but not in that of the hAdV-2 hexon. The sequence just before this loop is absent in the CAV-2 hexon (Fig. 4B and C, region 2), which could explain the absence of density in the EM structure. The structural rearrangement caused by this difference could fill the empty part in the CAV-2 EM hexon structure (Fig. 4B, asterisk). Further, there is a 24-amino-acid gap after amino acid 144 (Fig. 4B and C, region 1). This region is located on the top of the hAdV-5 hexon and was not ordered in the X-ray structure. This may explain the smoother shape of the CAV-2 hexon compared to the EM density of the hAdV-5 hexon (Fig. 4A, right). The last structural difference, also located on the top of the hexon, is likely caused by the significant sequence divergence around amino acid 310 (Fig. 4B and C, region 3 in the hAdV-5 sequence). Interestingly, the protruding structures in the hAdV hexons, which contain most of the epitopes recognized by anti-hexon antibodies (44), are absent in CAV-2.

As mentioned above, we have been unable to complex CAV-2 vectors with divalent cations or lipoplexes, suggesting a more neutral charge of the external CAV-2 capsid. Because the capsid charge is dominated by the hexon charges, the data presented here support this prediction. For example, region 1 in Fig. 4C is composed mainly of negatively charged residues in hAdV-5, and, more generally, all the hAdV-5 loops that are missing in the CAV-2 capsid are rich in charged residues.

We also performed a multiple fitting on the capsid using four hexons surrounding the twofold axis of the virus. The regions involved in the interhexon contacts are the same as those previously determined for hAdV-5 (15) (Fig. 4C, bold lines). Furthermore, in most cases, the amino acids are conserved between the two species. The four hexons present around the twofold axis in CAV-2 interact using the same five loops that are used in hAdV-5: loops 65 to 70, 292 to 296, 614 to 619, 674 to 676, and 683 to 690 in CAV-2 corresponding to loops 65 to 70, 340 to 344, 665 to 670, 725 to 727, and 734 to 741 in hAdV-5 (Fig. 4B and C, bold lines). Only the interaction region 2 (Fig. 4B) is absent in CAV-2.

Minor proteins. (i) Protein IX. A major difference in the 3D structure of hAdV and CAV-2 is located on the top of the triskelion structure attributed to the N-terminal domain of protein IX, where a rod-like density is present in CAV-2. This density is absent in the hAdV-5 reconstruction (Fig. 1B, arrow, and C and 5C). The sequence of protein IX is 25% shorter in CAV-2 than in hAdV-5 (105 versus 140 amino acids) (Fig. 5E). By comparing the 3D reconstructions of CAV-2 and hAdV, we should be able to locate protein IX within the capsid.

FIG. 4. The CAV-2 hexon capsomer. (A) Comparison of computer-isolated hexons from CAV-2 (yellow) and hAdV-5 (gray) imaged in isodensity surface view. (B) Fitting of hAdV-5 hexon X-ray structure (PDB 1P30) into the CAV-2 cryo-EM envelope. Either one hexon alone (left) or the four hexons surrounding a twofold axis (middle) have been fitted using SITUS and imaged at the same time in the EM density. On the right, only the bottom wall of the pseudo-atomic model of the four hexons is shown. In the hexon alone, three regions are indicated on the X-ray structure; they correspond to large differences in the sequence of CAV-2 compared to that of hAdV-5 (see below). The red loops and regions 1 to 3 in the single hexon on the left are discussed in the text and correspond to sequence elements 1 to 3 indicated in panel C. The star indicates a region in the EM density that is not filled by the density derived from the crystal structure, as discussed in the text. (C) Sequence alignment between the CAV-2 and hAdV-5 hexon bases. Conserved residues are in white text and highlighted in red; partially conserved residues are in red text. The loops not resolved in the X-ray structure are indicated by blue bars above the sequences. The amino acid sequences involved in the contacts between hexons as described by Fabry et al. (15) are underlined with a bold line. They are conserved between CAV-2 and hAdV-5. Regions 1, 2, and 3 from panel B are indicated.

In 2005, Fabry et al. (15) published the 3D structure of the wild-type capsid of hAdV-5 as well as the 3D structure of a mutant virus with a deletion of protein IX. In the mutant structure the yellow and red densities shown in Fig. 5B were not present. At the time of that study the position represented by the red density was thought to be that of protein IIIa; the authors thus assumed that the lack of the red densities in the protein IX-deleted virus was due to a diminished stability of the capsid, i.e., the absence of protein IX could have resulted in the loss of protein IIIa (15). Recently, Saban et al. (45), using a 6-Å resolution map in which the α -helices in the capsid were visible plus secondary structure predictions, reassigned the red density to a four-helical bundle built from the C-terminal parts of four copies of protein IX. The triskelion structure was considered to be made up from the N-terminal parts of three copies of protein IX. The density for protein IIIa was reassigned to a highly helical density present underneath the penton base (45). The reassignment of the position of the C-terminal domain of protein IX was confirmed by Marsh et al., who showed that a protein IX mutant virus with GFP added at the C terminus had an additional density above the reassigned position for the C terminus (33), i.e., the position of the red domains shown in Fig. 5C.

In the CAV-2 capsid (Fig. 5A and B), the triskelion N-terminal part of protein IX is present at the same position in the capsid as in hAdV-5 (Fig. 5C). The notable difference between the two reconstructions is that in CAV-2 a cylindrical density is present on the top of the triskelion, whereas the red density observed in the hAdV-5 capsid (Fig. 5C) was absent. The red density on top of the triskelion in CAV-2 (Fig. 5A and B) has more or less the same dimensions as the red rods in hAdV-5. It is therefore likely that the cylindrical density on top of the triskelion in the CAV-2 structure is also the C-terminal domain of protein IX. This interpretation is reinforced by comparing the sequences of the two proteins. Protein IX is a two-domain protein consisting of an N-terminal part that contains some β -strands and a C-terminal, α -helical part, both parts being connected by a linker, as shown by sequence alignment and secondary structure prediction (Fig. 5E). Saban et al. (45) did not detect any helices in the triskelion structure, confirming the assignment to the N-terminal domain of protein IX. This part of protein IX interacts with the hexon protein and should therefore be conserved. The N termini of the CAV-2 and the hAdV-5 sequences are the most similar between the two proteins (40% identity for the 52 first amino acids). However, the linker sequence (amino acids 40 to 81 in hAdV-5) is almost absent in CAV-2 (amino acids 39 to 53), and therefore the two domains must be close together in CAV-2. In contrast, in hAdV-5 the C and N termini could be quite far apart from each other, as suggested by Saban et al. (45). Although the CAV-2 C-terminal part of protein IX is slightly smaller and different from that of hAdV-5, this region

is still predicted to be helical and organized as a coil-coiled domain (Fig. 5E). Therefore, we suggest that three C-terminal domains of protein IX make up the rod on top of the triskelion structure in the CAV-2 capsid. To test this hypothesis we performed cryo-EM and calculated a 3D reconstruction of a CAV-2 virion harboring a C-terminal pIX-GFP fusion protein (31). While the quality of the reconstruction (25-Å resolution) is lower than that of the native CAV-2 capsid, it is good enough to recognize the triskelion, the cylindrical red density, and, at the top of that, a small extra density (Fig. 5D, green areas). This latter density is smaller than the size of GFP, probably because it is flexible. The flexible linker between the C-terminal part of protein IX and GFP (30, 31) would allow movement, and therefore only the average position is visible in the 3D reconstruction. We believe that our combined data comparing the reconstructions of CAV-2 and the capsid harboring the protein IX-GFP fusion confirm the localization of the C-terminal domain of CAV-2 protein IX. Our results also confirm the assignment of Saban et al. (45) of the red rods in the hAdV structure shown in Fig. 5C as the C-terminal domain of human virus protein IX.

(ii) The interior of the capsid. (a) Protein IIIa. As mentioned above, based on a 6-Å EM reconstruction and secondary structure predictions (45), the hAdV protein IIIa was reassigned to a helix-rich density lying under the penton base. This density is also present on the inside of the CAV-2 capsid (not shown). It exhibits the same topology as the human protein IIIa: there is an arm bound to the bottom part of the penton base, and, more toward the inside of the capsid, a globular region which resembles an empty sphere is visible and of same size as in hAdV-5. The CAV-2 protein IIIa is slightly smaller than the hAdV-5 protein (567 amino acids versus 585), but except for the 15 first amino acids, the N-terminal part predicted to be α -helical is very conserved. The C-terminal part is notably different and smaller in CAV-2. However, at the present resolution we are unable to see such differences.

(b) Protein VIII. A total of 120 copies of protein VIII have been counted on the inner part of the hAdV-5 capsid (15). The same number of proteins can be counted in CAV-2, and each interacts with four hexons (not shown). The density corresponding to protein VIII seems to be slightly smaller and less continuous in CAV-2 than in hAdV-5. The sequences of these proteins are very similar (more than 66% similarity or identity), and therefore the differences observed in the density are probably due to a lack of resolution.

Finally, there is also an extra density just under the CAV-2 penton base (Fig. 1B, circle) that is absent in the hAdV-5 structure, but at this stage it remains unassigned.

Conclusion. This work shows the first 3D structure of a nonhuman AdV. There are well over 100 AdV serotypes currently identified, with less than half of them human. We predict that this number of nonhuman AdVs is only the tip of the

penton and the hexons are as described in panel A. Protein IX is in yellow (N-terminal part) and red (C-terminal part), and the extra density assigned to GFP is in green. (E) Sequence alignment of hAdV-5 and CAV-2 protein IX. Conserved residues are in white text and highlighted in red; partially conserved residues are in red text. The arrows and open bars represent predicted β -strands and α -helices, respectively, as given by the secondary structure prediction software (<http://bioweb.pasteur.fr/seqanal/interfaces/predator-simple.html>). The predicted coil-coiled is boxed into a rectangle.

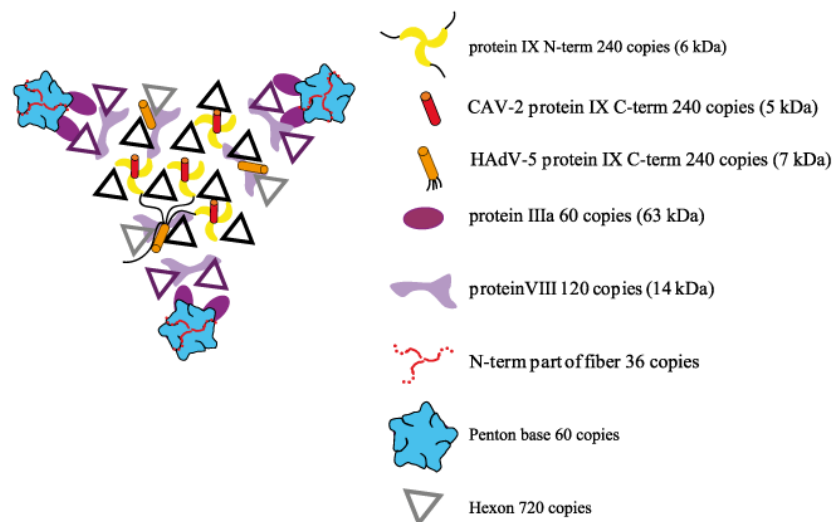


FIG. 6. Schematic view of one facet of the CAV-2 capsid seen from the outside of the capsid. The figure shows a summary of all structural data on the protein localization in the AdV capsid. The hexons highlighted in black belong to the group of nine whereas the violet ones are the peripentonal hexons (26). The different proteins occupy the same positions in the capsid (hAdV and the CAV-2) except for the C-terminal part of protein IX that can either be on the top of the N terminus of protein IX in the CAV-2 capsid or be near the twofold axis in the hAdV capsid. For the hAdV protein IX, a thin black line indicates the position of the linker between the N- and the C-terminal parts of protein IX as suggested by Saban et al. (45).

iceberg. We are just beginning to understand and uncover the potential of many of these viruses, in particular, of the nonhuman serotypes. Our study highlights the structural differences and similarities between hAdV-2/5 and CAV-2 and definitively localizes the N- and C-terminal parts of protein IX in this CAV-2 capsid. The molecular organization of the AdV capsid as it is known today is shown schematically in Fig. 6. The CAV-2 penton base and hexon have fewer and/or smaller loops around the conserved bodies of these capsid proteins. In the hAdV-2/5 penton base these loops are involved in internalization via the integrin-interacting domain (34), possibly binding coagulation factors (2), and are targeted by the humoral immune response following infection (37). The biological significance of many of the structural differences is likely to be complex. Yet we can now begin to rationally understand and resolve some of the CAV-2 enigmas. The capsid structure will allow us to rescan the external CAV-2 capsid for integrin-interacting motifs that could possibly induce internalization. Because of its C-terminal domain extending outside of the capsid wall, the CAV-2 protein IX may have an additional role during capsid disassembly at the nuclear pore by binding the anterograde motor kinesin (U. Greber, unpublished data). The smoother shape of the CAV-2 hexons and penton bases due to the lack of highly antigenic loops present in human serotypes explains the lack of cross-reacting human neutralizing antibodies (37).

Biologically active loops can be added onto the CAV-2 penton base and hexon at the sites where such loops are present in the human viruses. In particular, it will be interesting to add an RGD loop to the CAV-2 penton base at the same site where this is found in the human virus. Apart from adding loops to pentons and hexons, the specific position of the C-terminal domain of CAV-2 protein IX is an ideal site to position additional protein domains for specific interactions with host cells.

Finally, we may also be able to better understand and assay CAV-2 tropism by manipulating the rigidity of the flexible shaft of the fiber.

ACKNOWLEDGMENTS

We thank Pascal Fender and Harry Wodrich for discussions and Emmanuelle Neumann for help with the electron microscope. We thank Vasiliki Kalatzis for critical reading of the manuscript. We thank the other members of our laboratories for constructive comments, suggestions, and help during the course of this study. We thank the IFR 122 recombinant protein platform for the production of the CAV-2 fiber.

E.J.K. is an INSERM fellow. G.S. was supported in part by a Jeunes Chercheurs grant from the Agence Nationale pour la Recherche. D.T.C. was funded by NIH award 5R01CA111569. This work was supported, in part, by the Association Française contre les Myopathies (AFM), Vaincre les Maladies Lyosomales, and the Fondation pour la Recherche Médicales. O.B. was an AFM postdoctoral researcher, M.E.B. is a Ph.D. student financed by the Rhone Alpes region, and C.M.S.F. is a Ph.D. student financed through a MENRT fellowship from the French government.

The authors have no conflicting financial interests.

REFERENCES

1. Akli, S., C. Caillaud, E. Vigne, L. D. Stratford-Perricaudet, L. Poenaru, M. Perricaudet, A. Kahn, and M. R. Peschanski. 1993. Transfer of a foreign gene into the brain using adenovirus vectors. *Nat. Genet.* 3:224-228.
2. Baker, A. H., J. H. McVey, S. N. Waddington, N. C. Di Paolo, and D. M. Shayakhmetov. 2007. The influence of blood on in vivo adenovirus biodistribution and transduction. *Mol. Ther.* 15:1410-1416.
3. Baker, T. S., and R. H. Cheng. 1996. A model-based approach for determining orientations of biological macromolecules imaged by cryoelectron microscopy. *J. Struct. Biol.* 116:120-130.
4. Bergelson, J. M., J. A. Cunningham, G. Droguett, E. A. Kurt-Jones, A. Krithivas, J. S. Hong, M. S. Horwitz, R. L. Crowell, and R. W. Finberg. 1997. Isolation of a common receptor for coxsackie B viruses and adenoviruses 2 and 5. *Science* 275:1320-1323.
5. Both, G. W. 2002. Xenogenic adenoviruses, p. 447-479. In D. Curiel and J. Douglas (ed.), *Adenoviral vectors for gene therapy*. Academic Press, San Diego, CA.
6. Campos, S. K., and M. A. Barry. 2007. Current advances and future chal-

- lences in adenoviral vector biology and targeting. *Curr. Gene Ther.* 7:189–204.
7. Chappell, J. D., A. E. Prota, T. S. Dermody, and T. Stehle. 2002. Crystal structure of reovirus attachment protein signal¹ reveals evolutionary relationship to adenovirus fibre. *EMBO J.* 21:1–11.
 8. Chillon, M., and E. Kremer. 2001. Trafficking and propagation of canine adenovirus vectors lacking a known integrin-interacting motif. *Hum. Gene Ther.* 12:1815–1823.
 9. Chroboczek, J., R. W. Ruigrok, and S. Cusack. 1995. Adenovirus fiber. *Curr. Top. Microbiol. Immunol.* 199:163–200.
 10. Conway, J. F., and A. C. Steven. 1999. Methods for reconstructing density maps of “single” particles from cryoelectron micrographs to subnanometer resolution. *J. Struct. Biol.* 128:106–118.
 11. D’Ambrosio, E., N. Del Grosso, A. Chicca, and M. Midulla. 1982. Neutralizing antibodies against 33 human adenoviruses in normal children in Rome. *J. Hyg.* 89:155–161.
 12. Davison, A. J., M. Benko, and B. Harrach. 2003. Genetic content and evolution of adenoviruses. *J. Gen. Virol.* 84:2895–2908.
 13. Dayringer, H. E., A. Tramontano, S. R. Sprang, and R. J. Fletterick. 1986. Interactive program for visualization and modeling of protein, nucleic acids and small molecules. *J. Mol. Graph.* 4:82–87.
 14. DeLano, M. C., and Y. Cao. 2002. High b-value diffusion imaging. *Neuroimaging Clin. N. Am.* 12:21–34.
 15. Fabry, C. M., M. Rosa-Calatrava, J. F. Conway, C. Zubieta, S. Cusack, R. W. Ruigrok, and G. Schoehn. 2005. A quasi-atomic model of human adenovirus type 5 capsid. *EMBO J.* 24:1645–1654.
 16. Fasbender, A., J. H. Lee, R. W. Walters, T. O. Moninger, J. Zabner, and M. J. Welsh. 1998. Incorporation of adenovirus in calcium phosphate precipitates enhances gene transfer to airway epithelia in vitro and in vivo. *J. Clin. Investig.* 102:184–193.
 17. Fasbender, A., J. Zabner, M. Chillon, T. O. Moninger, A. P. Puga, B. L. Davidson, and M. J. Welsh. 1997. Complexes of adenovirus with polycationic polymers and cationic lipids increase the efficiency of gene transfer in vitro and in vivo. *J. Biol. Chem.* 272:6479–6489.
 18. Flomenberg, P., V. Piaskowski, R. L. Truitt, and J. T. Casper. 1995. Characterization of human proliferative T cell responses to adenovirus. *J. Infect. Dis.* 171:1090–1096.
 19. Frank, J., M. Radermacher, P. Penczek, J. Zhu, Y. Li, M. Ladjadj, and A. Leith. 1996. SPIDER and WEB: processing and visualization of images in 3D electron microscopy and related fields. *J. Struct. Biol.* 116:190–199.
 20. Fuschiotti, P., G. Schoehn, P. Fender, C. M. Fabry, E. A. Hewat, J. Chroboczek, R. W. Ruigrok, and J. F. Conway. 2006. Structure of the dodecahedral penton particle from human adenovirus type 3. *J. Mol. Biol.* 356:510–520.
 21. Galinier, R., E. Gout, H. Lortat-Jacob, J. Wood, and J. Chroboczek. 2002. Adenovirus protein involved in virus internalization recruits ubiquitin-protein ligases. *Biochemistry* 41:14299–14305.
 22. Garcel, A., E. Gout, J. Timmins, J. Chroboczek, and P. Fender. 2006. Protein transduction into human cells by adenovirus dodecahedron using WW domains as universal adaptors. *J. Gene. Med.* 8:524–531.
 23. Gouet, P., X. Robert, and E. Courcelle. 2003. ESPript/ENDscript: Extracting and rendering sequence and 3D information from atomic structures of proteins. *Nucleic Acids Res.* 31:3320–3323.
 24. Greber, U. F. 2002. Signalling in viral entry. *Cell Mol. Life Sci.* 59:608–626.
 25. Hess, M., A. Cuzange, R. W. Ruigrok, J. Chroboczek, and B. Jacrot. 1995. The avian adenovirus penton: two fibres and one base. *J. Mol. Biol.* 252:379–385.
 26. Horwitz, M. 2001. Adenoviruses, p. 2149–2171. In D. M. Knipe, P. M. Howley, D. E. Griffin, R. A. Lamb, M. A. Martin, B. Roizman, and S. E. Straus (ed.), *Fields virology*, 4th ed. Lippincott Williams & Wilkins, Philadelphia, PA.
 27. Keriel, A., C. Rene, C. Galer, J. Zabner, and E. J. Kremer. 2006. Canine adenovirus vectors for lung-directed gene transfer: efficacy, immune response, and duration of transgene expression using helper-dependent vectors. *J. Virol.* 80:1487–1496.
 28. Kojaoghlianian, T., P. Flomenberg, and M. S. Horwitz. 2003. The impact of adenovirus infection on the immunocompromised host. *Rev. Med. Virol.* 13:155–171.
 29. Kremer, E. J. 2004. CAR chasing: canine adenovirus vectors-all bite and no bark? *J. Gene. Med.* 6(Suppl. 1):S139–S151.
 30. Kremer, E. J., S. Boutin, M. Chillon, and O. Danos. 2000. Canine adenovirus vectors: an alternative for adenovirus-mediated gene transfer. *J. Virol.* 74:505–512.
 31. Le, L. P., J. Li, V. V. Ternovoi, G. P. Siegal, and D. T. Curiel. 2005. Fluorescently tagged canine adenovirus via modification with protein IX-enhanced green fluorescent protein. *J. Gen. Virol.* 86:3201–3208.
 32. Le Gal La Salle, G., J. J. Robert, S. Berrard, V. Ridoux, L. D. Stratford-Perricaudet, M. Perricaudet, and J. Mallet. 1993. An adenovirus vector for gene transfer into neurons and glia in the brain. *Science* 259:988–990.
 33. Marsh, M. P., S. K. Campos, M. L. Baker, C. Y. Chen, W. Chiu, and M. A. Barry. 2006. Cryoelectron microscopy of protein IX-modified adenoviruses suggests a new position for the C terminus of protein IX. *J. Virol.* 80:11881–11886.
 34. Medina-Kauwe, L. K. 2003. Endocytosis of adenovirus and adenovirus capsid proteins. *Adv. Drug Deliv. Rev.* 55:1485–1496.
 35. Olive, M., L. Eisenlohr, N. Flomenberg, S. Hsu, and P. Flomenberg. 2002. The adenovirus capsid protein hexon contains a highly conserved human CD4⁺ T-cell epitope. *Hum. Gene Ther.* 13:1167–1178.
 36. Paillard, F. 1997. Advantages of nonhuman adenoviruses versus human adenoviruses. *Hum. Gene Ther.* 8:2007–2009.
 37. Perreau, M., M. C. Guerin, C. Drouet, and E. J. Kremer. 2007. Interactions between human plasma components and a xenogenic adenovirus vector: reduced immunogenicity during gene transfer. *Mol. Ther.* 15:1998–2007.
 38. Perreau, M., and E. J. Kremer. 2005. Frequency, proliferation, and activation of human memory T cells induced by a nonhuman adenovirus. *J. Virol.* 79:14595–14605.
 39. Perreau, M., and E. J. Kremer. 2006. The conundrum between immunological memory to adenovirus and their use as vectors in clinical gene therapy. *Mol. Biotechnol.* 34:247–256.
 40. Perreau, M., F. Mennechet, N. Serratrice, J. N. Glasgow, D. T. Curiel, H. Wodrich, and E. J. Kremer. 2007. Contrasting effects of human, canine, and hybrid adenovirus vectors on the phenotypical and functional maturation of human dendritic cells: implications for clinical efficacy. *J. Virol.* 81:3272–3284.
 41. Rasmussen, U. B., Y. Schlesinger, A. Pavirani, and M. Mehtali. 1995. Sequence analysis of the canine adenovirus 2 fiber-encoding gene. *Gene* 159:279–280.
 42. Ruigrok, R. W., A. Barge, C. Albiges-Rizo, and S. Dayan. 1990. Structure of adenovirus fibre. II. Morphology of single fibres. *J. Mol. Biol.* 215:589–596.
 43. Ruigrok, R. W. H., A. Barge, S. K. Mittal, and B. Jacrot. 1994. The fibre of bovine adenovirus is very long, but bent. *J. Gen. Virol.* 75:2069–2073.
 44. Rux, J. J., P. R. Kuser, and R. M. Burnett. 2003. Structural and phylogenetic analysis of adenovirus hexons by use of high-resolution X-ray crystallographic, molecular modeling, and sequence-based methods. *J. Virol.* 77:9553–9566.
 45. Saban, S. D., M. Silvestry, G. R. Nemerow, and P. L. Stewart. 2006. Visualization of α -helices in a 6-angstrom resolution cryoelectron microscopy structure of adenovirus allows refinement of capsid protein assignments. *J. Virol.* 80:12049–12059.
 46. Schoehn, G., F. M. Vellieux, M. Asuncion Dura, V. Receveur-Brechot, C. M. Fabry, R. W. Ruigrok, C. Ebel, A. Roussel, and B. Franzetti. 2006. An archaeal peptidase assembles into two different quaternary structures: a tetrahedron and a giant octahedron. *J. Biol. Chem.* 281:36327–36337.
 47. Seiradake, E., H. Lortat-Jacob, O. Billet, E. J. Kremer, and S. Cusack. 2006. Structural and mutational analysis of human Ad37 and canine adenovirus 2 fiber heads in complex with the D1 domain of coxsackie and adenovirus receptor. *J. Biol. Chem.* 281:33704–33716.
 48. Soudais, C., S. Boutin, S. S. Hong, M. Chillon, O. Danos, J. M. Bergelson, P. Boulanger, and E. J. Kremer. 2000. Canine adenovirus type 2 attachment and internalization: coxsackievirus-adenovirus receptor, alternative receptors, and an RGD-independent pathway. *J. Virol.* 74:10639–10649.
 49. Soudais, C., C. Laplace-Builhe, K. Kissa, and E. J. Kremer. 2001. Preferential transduction of neurons by canine adenovirus vectors and their efficient retrograde transport in vivo. *FASEB J.* 15:2283–2285.
 50. Soudais, C., N. Skander, and E. J. Kremer. 2004. Long-term in vivo transduction of neurons throughout the rat central nervous system using novel helper-dependent CAV-2 vectors. *FASEB J.* 18:391–393.
 51. Stewart, P. L., R. M. Burnett, M. Cyrklaff, and S. D. Fuller. 1991. Image reconstruction reveals the complex molecular organization of adenovirus. *Cell* 67:145–154.
 52. Tomko, R., R. Xu, and L. Philipson. 1997. HCAR and MCAR: the human and mouse cellular receptors for subgroup C adenoviruses and group B coxsackieviruses. *Proc. Natl. Acad. Sci. USA* 94:3352–3356.
 53. Van Heel, M. 1987. Angular reconstitution: a posteriori assignment of projection directions for 3D reconstruction. *Ultramicroscopy* 21:111–123.
 54. van Raaij, M. J., A. Mitraki, G. Lavigne, and S. Cusack. 1999. A triple beta-spiral in the adenovirus fibre shaft reveals a new structural motif or a fibrous protein. *Nature* 401:935–938.
 55. Wriggers, W., and S. Birmanns. 2001. Using situs for flexible and rigid-body fitting of multiresolution single-molecule data. *J. Struct. Biol.* 133:193–202.
 56. Wu, E., and G. R. Nemerow. 2004. Virus yoga: the role of flexibility in virus host cell recognition. *Trends Microbiol.* 12:162–169.
 57. Wu, E., L. Pache, D. J. Von Seggern, T. M. Mullen, Y. Mikyas, P. L. Stewart, and G. R. Nemerow. 2003. Flexibility of the adenovirus fiber is required for efficient receptor interaction. *J. Virol.* 77:7225–7235.
 58. Zubieta, C., G. Schoehn, J. Chroboczek, and S. Cusack. 2005. The structure of the human adenovirus 2 penton. *Mol. Cell* 17:121–135.

c. Conclusion

Plusieurs structures tridimensionnelles d'Ads ont été obtenues par microscopie électronique (Fabry et al., 2005; Saban et al., 2006; Stewart et al., 1991). Cependant des ambiguïtés persistaient quant à la localisation de la protéine IX à la surface de la capsid. La reconstruction 3D du CAV-2 a été réalisée au laboratoire (Schoehn et al., 2008). Cette reconstruction montre que le virus canin a une composition structurale très similaire à celle des adénovirus humains. Les protéines structurales de l'adénovirus canins sont légèrement plus petites en nombre d'acides aminés que leurs équivalents humains. En particulier, la protéine IX canine ne possède pas la partie centrale riche en alanine présente dans la protéine humaine (la séquence est 25% plus courte). La comparaison de la structure 3D de l'HA5 et du CAV-2 a permis la localisation de la protéine IX dans la capsid. La région N-terminale de la protéine (conservée entre les deux virus) en forme de triskélion est présente au même endroit chez les deux virus. En revanche la région cylindrique située à l'interface des facettes icosaédrique du HA5 (en orange Fig.5C de la publication) n'apparaît pas chez le CAV-2. Pour le CAV-2 une densité cylindrique similaire à celle située à l'interface des facettes chez l'adénovirus humain est positionnée au dessus du triskélion. La structure 3D du CAV-2 dont le domaine C-terminal de la IX a été fusionné à la GFP a permis d'attribuer la densité cylindrique au dessus du triskélion au domaine C-terminal de la protéine IX. De même pour l'HA5, une structure 3D du virus complexé avec des Fabs ciblant la protéine IX du côté C-terminal a permis d'attribuer la densité cylindrique située à l'interface des facettes icosaédrique au domaine C-terminal de la IX (voir plus haut ; Fig. 3.13).

Si l'organisation de l'extrémité N-terminale est conservée chez le virus canin et humain, la localisation du domaine C-terminal de la protéine IX n'est pas conservée.

3.3.4. Discussion

Les résultats concernant la localisation de la protéine IX dans la capsidie par MET, en plus des résultats obtenus sur l'état d'oligomérisation de la protéine, nous ont permis de proposer un modèle d'organisation de cette protéine dans la capsidie (Fig. 3.14D). Ce modèle qui confirme celui de Saban et *al.*, 2006 (Fig. 3.14B) est basé sur des données biochimiques et structurales.

Dans le modèle que nous proposons pour l'organisation de la protéine IX, celle-ci crée un réseau à la surface de la capsidie entre les capsomères (Fig. 3.14D). En accord avec les résultats d'immunodétection (Fig. 3.13), nous avons positionné le domaine C-terminal entre deux hexons à l'interface entre deux facettes de l'icosaèdre (en vert Fig. 3.14D). La structure d'HAd35 à 6 Å de résolution publiée par Saban et al. (2006) révèle un cluster de 4 hélices α à l'emplacement du domaine C-terminal (Fig. 3.14B). Cette observation est cohérente avec la présence de la forme tétramérique de ce domaine lors des expériences de pontage chimique (Fig. 3.6B). Dans le modèle de Saban et *al.*, 3 hélices proviennent de la même facette et la quatrième de la facette adjacente. La densité de la quatrième hélice est très faible dans la structure et seule une petite proportion des adénovirus possède cette quatrième hélice (*P Stewart, communication personnelle*). Ceci explique la présence d'un trimère majoritaire et d'un tétramère minoritaire dans les expériences de pontage chimique (Fig. 3.6B). Le domaine C-terminal apparaît plus accessible que le domaine N-terminal de la protéine IX en accord avec les résultats d'immunodétection déjà publiés (Akalu et al., 1999).

Le domaine N-terminal est localisé au milieu de la facette entre trois hexons sous forme de triskélion (en bleu dans le modèle, en jaune sur la facette d'HAd5 et en rouge indiqué par une flèche dans la facette d'HAd35 à 6 Å de résolution Fig. 3.14A, B et D). La région N-terminale est responsable de l'incorporation de la protéine IX dans la capsidie ainsi que de sa stabilisation (Rosa-Calatrava et al., 2001; Vellinga et al., 2005), sa séquence et sa position dans la capsidie sont conservées par rapport à la région C-terminal. Le domaine N-terminal représente une grande surface de contact avec les hexons indiquant que c'est sans doute cette région de la protéine IX qui contribue le plus à la stabilisation de la capsidie.

Les structures 3D des virus montrent que les surfaces d'interaction entre les domaines C-terminaux et l'hexon sont réduites par rapport à celles existantes entre les hexons et les régions N-terminales. Ceci est en accord avec le fait que le domaine C-terminal n'est ni nécessaire pour l'incorporation de la protéine IX dans la capsidie ni pour sa stabilisation.

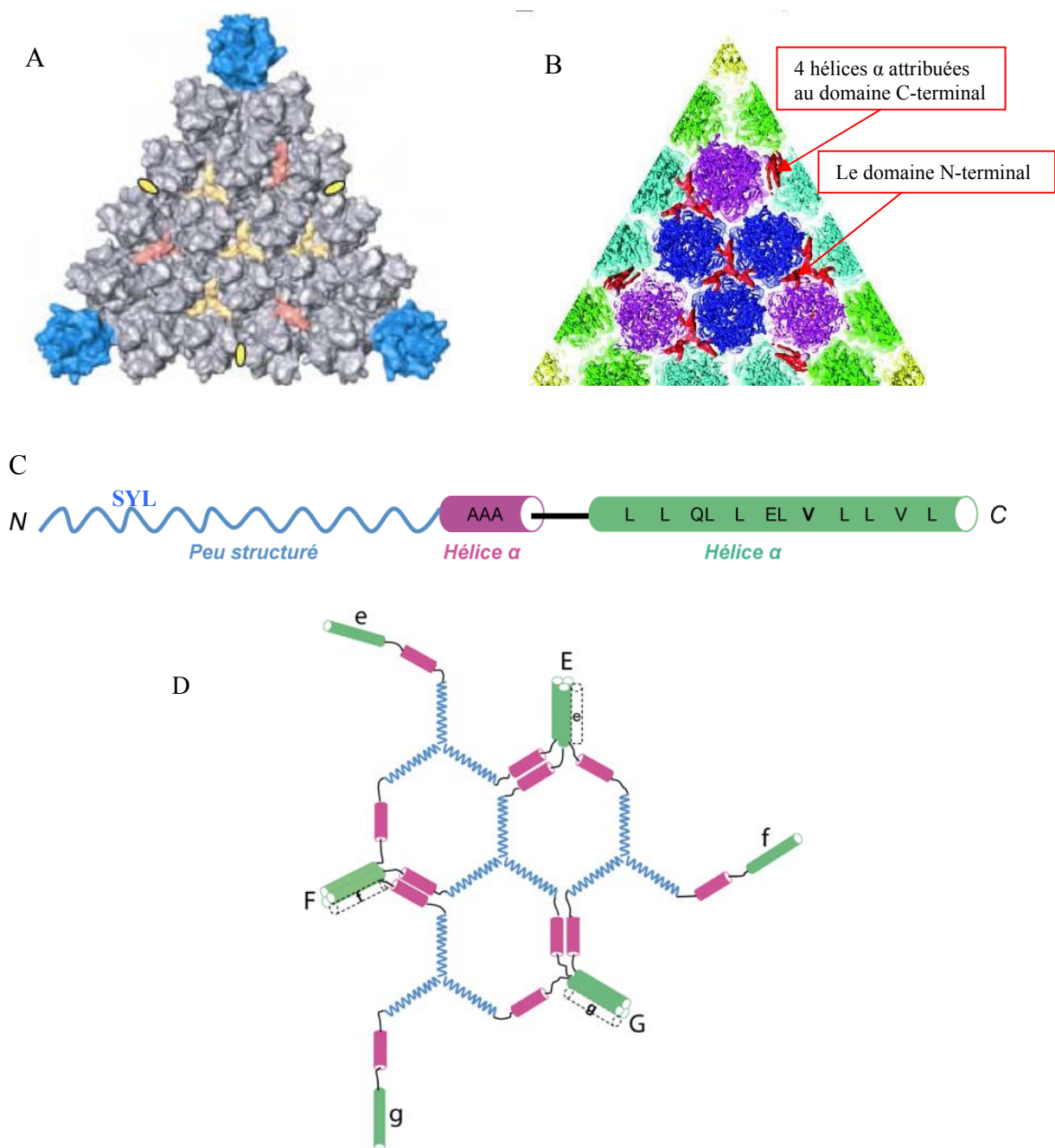


Figure 3.14: Modèle d'organisation de la protéine IX à la surface de la capside.

A, Facette de l'HAd-5 sauvage à 10 Å de résolution avec le domaine N-ter de la protéine IX en jaune et le domaine C-ter en orange (d'après Fabry et al., 2005).

B, Facette de la capside virale d'HAd35 à 6 Å, en rouge la densité attribuée à la protéine IX (Saban et al., 2006). Le domaine C-terminal est indiqué par une flèche, il est composé de 4 hélices α .

C, Schéma d'un monomère de la protéine IX basé sur les prédictions de structure secondaire. La région N-ter est peu structurée, la région centrale ainsi que le domaine C-ter sont prédites en hélice α .

D, Schéma d'organisation de la protéine IX à la surface de la capside virale, les couleurs utilisées pour les différents domaines sont les mêmes que pour la partie C. En se basant sur le modèle de Saban et al. le domaine C-terminal est constitué de 4 hélices α , parmi eux trois proviennent de la même facette et le quatrième de la facette adjacente.

Le domaine N-terminal de la protéine IX apparaît sous forme trimérique à la surface de la capsid, nous n'avons pas pu mettre en évidence ce trimère avec les expériences de pontage covalent ou de MALLS. Ceci est expliqué soit par le fait qu'il n'y a pas d'interaction spécifique en solution au niveau du domaine N-terminal, soit par la désorganisation de ce domaine provoquée par la délétion des trois acides aminés SYL du côté N-terminal (cette délétion était nécessaire pour l'obtention de la protéine sous forme soluble).

Le domaine central (en rose Fig. 3.14D) est positionné entre deux hexons faisant le lien entre les deux extrémités de la protéine IX. Prédite pour être organisée en hélice α , la région centrale de 24 acides aminés d'une longueur attendue de 36 Å correspond approximativement à la distance entre les domaines N et C-terminaux qui ont été attribués à leur densité dans la capsid. Cette région est sous forme de dimère dans le modèle (Fig. 3.14D) en accord avec les résultats du pontage chimique indiquant l'implication de cette région dans la dimérisation de la protéine IX.

En résumé, la protéine IX adopte une organisation en un maillage structuré à la surface de la capsid rappelant le cas d'autres protéines virales connues pour établir de tel réseau (la protéine P32 du bactériophage PRD1 ; (Abrescia et al., 2004)).

4. La Protéine IIIa

4.1. Introduction

La protéine IIIa est présente en 60 copies sous forme de monomère associé à la base de penton du côté interne de la capsid (Fig. 1.12). La protéine IIIa est l'une des protéines mineures sur laquelle aucune donnée structurale n'existe. Dans le cadre d'un projet d'étude structurale générale des protéines mineures de l'adénovirus nous nous intéressons également à cette protéine.

La protéine IIIa sauvage a fait l'objet de plusieurs tentatives d'expression en bactéries, elle est exprimée en très grande quantité (50 mg/L de culture) mais elle est totalement insoluble. Les prédictions de structures secondaires de la protéine IIIa ont permis de constater que cette protéine était constituée majoritairement d'hélices α (Fig. 4.1). Afin d'obtenir cette protéine sous forme soluble nous avons suivi deux stratégies :

- Purification de la protéine à partir des corps d'inclusion suivie d'étapes de dénaturation / renaturation.
- Désignation de fragments de la protéine susceptibles d'être solubles en se basant sur les données bioinformatiques.

dénaturante (les mêmes étapes qu'une purification en condition native comme décrit dans le chapitre Matériels et Méthodes avec 6 M de chlorure de guanidine dans les tampons).

-Renaturation par dilution rapide :

La protéine dénaturée purifiée est concentrée jusqu'à 10 mg/mL à l'aide d'un concentrateur (*Amicon*). La renaturation se fait par dilution rapide en ajoutant goutte à goutte la solution concentrée dans le tampon de renaturation (50 mM Tris pH 7.9 ; 500 mM NaCl) sous agitation. Le volume de ce tampon est relativement élevé, il faut faire en sorte que la concentration finale de la protéine soit de 0.1 mg/ml. La protéine IIIa que l'on espère avoir renaturée est soumise à une nouvelle colonne d'affinité dans des conditions natives.

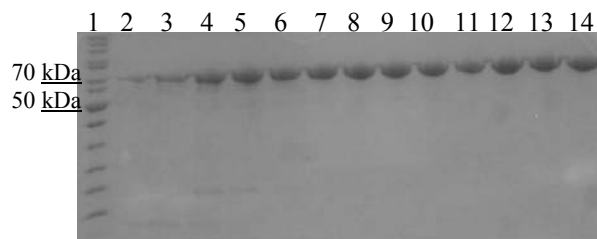


Figure 4.2: Purification de la protéine IIIa. Gel SDS-PAGE dénaturant coloré au bleu de coomassie: Piste 1 : marqueur de poids moléculaire (Fermentas). Pistes 2 à 14 : fractions d'élué de la protéine IIIa purifiée sur la colonne d'affinité en condition native. La protéine est éluee avec 500 mM d'imidazole. Malgré cette forte concentration en imidazole la protéine est présente tout au long de l'élué. La protéine est éluee sur une vingtaine de millilitre (au lieu de 6-8 ml pour une protéine naturellement soluble).

Comme le montre la figure 4.2, la protéine IIIa purifiée après dénaturation et renaturation est bien soluble. Un échantillon a été observé par microscopie électronique révélant une protéine globulaire apparemment bien repliée. Malheureusement lors de sa concentration, cette protéine précipite. Ce comportement, connu pour des protéines renaturées après dénaturation, indique que la protéine n'est pas correctement. Nous avons testé plusieurs additifs connus pour aider à obtenir une bonne renaturation de cette protéine (arginine, glycérol et sucrose) mais aucun n'a donné de résultats concluants.

4.3. Recherche de fragments solubles

Plusieurs programmes, dont les adresses internet sont listées ci-dessous, ont été utilisés pour prédire la nature et la présence de structures secondaires dans la protéine IIIa:

<http://dis.embl.de/>

<http://globplot.embl.de/>

<http://www.strubi.ox.ac.uk/RONN>

<http://bip.weizmann.ac.il/fldbin/findex>

<http://www.ist.temple.edu/disprot/predictorVSL2.php>

<http://iupred.enzim.hu/>

<http://www.pondr.com/>

<http://bioinf.cs.ucl.ac.uk/disopred/disopred.html>

Les résultats obtenus par ces différents algorithmes sont cohérents entre eux: la protéine IIIa est visiblement désordonnée du côté N-terminal (les 20 premiers acides aminés) et également du côté C-terminal (les 180 derniers acides aminés) comme le montre l'exemple de la figure 4.3.

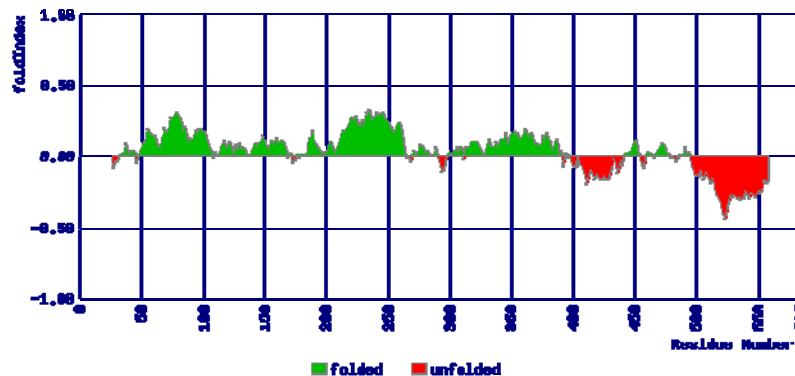


Figure 4.3: Prédiction de repliement de la protéine IIIa à partir de sa séquence primaire. Les résultats du programme FoldIndex sont représentés. La région centrale (20 -400) semble être structurée, alors que les régions N (1-20 aa) et C-terminales (400-585) sont prédites non repliées.

D'après la figure ci-dessus, les régions N et C-terminales de la protéine IIIa sont désordonnées. Elles sont suspectées d'être à l'origine de l'insolubilité de la protéine IIIa. En vue de générer un fragment soluble de cette protéine, différentes constructions tronquées soit du côté C-terminal soit des côtés N et C-terminaux ont été réalisées. Un total de douze constructions différentes de la protéine IIIa a été élaboré (Fig. 4.4) basé sur les résultats de prédiction obtenus précédemment.

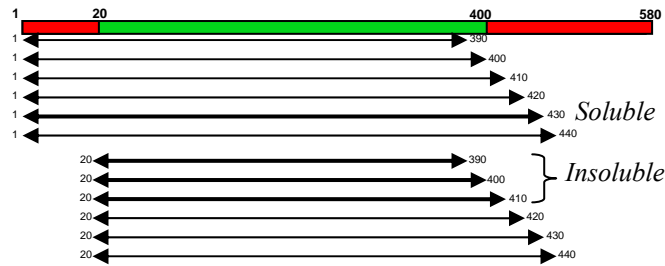


Figure 4.4: Vue schématique des 12 constructions de la protéine IIIa clonées pour tester leur solubilité. Dans les six premières constructions le domaine N-terminal a été conservé mais le gène codant pour cette protéine a été tronqué de manière incrémentale pour éliminer les acides aminés dix par dix du côté C-terminal. Les six dernières constructions sont équivalentes aux six premières mais avec une délétion de 20 acides aminés du côté N-terminal.

La construction 1-430

Nous nous sommes intéressés dans un premier temps à une des plus grandes constructions (1-430). La protéine est produite sous forme soluble chez la bactérie suivant la méthodologie décrite dans le chapitre Matériels et Méthodes. Nous concluons que la région C-terminale désordonnée était probablement à l'origine de la formation des corps d'inclusion lors de l'expression de la protéine sauvage.

La protéine recombinante a été purifiée d'abord grâce à une chromatographie d'affinité à l'aide de l'étiquette poly-histidine et ensuite sur une colonne d'exclusion de taille. Le résultat de cette dernière étape de purification est présenté sur la figure 4.5.

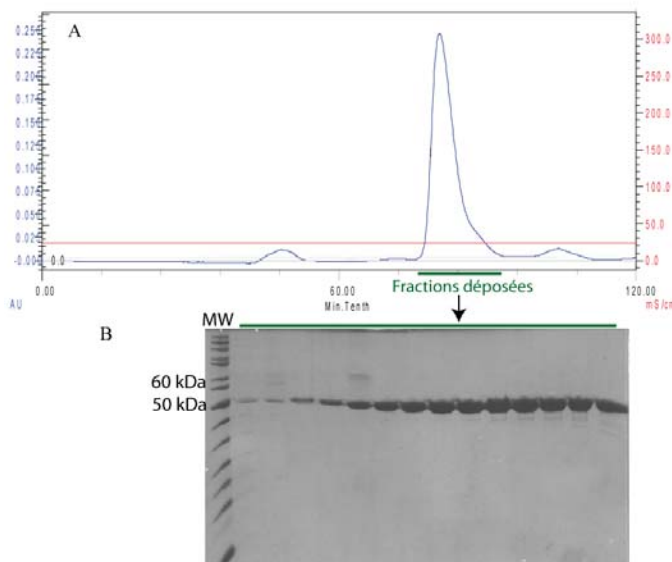


Figure 4.5: Purification de la construction 1-430 de la protéine IIIa.

A, Profil d'éluion de la protéine sur une colonne d'exclusion de taille Superdex 200 (Amersham).

B, Gel dénaturant SDS-PAGE analysant les différentes fractions du pic à la sortie de colonne.

La pureté de l'échantillon, révélée par le gel de la figure 4.5 permet son utilisation directe pour les essais de cristallogénèse.

Essais de cristallogenèse

Les essais de cristallisation de la protéine IIIa sont réalisés dans des nano-gouttes à l'aide de la plate-forme robotique du PSB. Le robot a testé 6x96 conditions (*cf.* Matériels et Méthodes) différentes pour plusieurs concentrations protéiques (entre 5 et 20 mg/ml) et différentes températures (4, 12, et 20°C).

Malgré la pureté et l'homogénéité de l'échantillon, seuls quelques cristaux uniques de très petite taille et non reproductibles ont été obtenus.

Après les résultats non concluants de la cristallisation de la construction 1-430 de la protéine IIIa, nous nous sommes intéressés à des constructions de plus petites tailles de cette protéine et particulièrement aux constructions 20-390, 20-400 et 20-410. Ces constructions se sont avérées insolubles. Nous en avons conclu que la région N-terminale est probablement indispensable à un bon repliement de la protéine.

4.4. Conclusion et perspectives

Durant ces études, nous nous sommes rendus compte que la protéine IIIa sauvage était totalement insoluble. La caractérisation des formes tronquées de la protéine a permis de constater que la région C-terminale est probablement responsable de l'insolubilité de cette protéine et que la région N-terminale semble nécessaire à un repliement correct. Ce travail sur la protéine IIIa semble être sur la bonne voie. A notre connaissance, nous avons été les premiers à obtenir des constructions solubles de cette protéine. Par manque de temps, je n'ai pas pu tester toutes les constructions de cette protéine (1-390 ; 1-400 ; 1-410 ; 1-420 et 1-440) ni explorer toutes les méthodes et conditions de cristallisation.

Dans le cadre du partenariat pour la biologie structurale (PSB) nous collaborons avec l'équipe du Dr. Darren Hart, qui a mis au point une technique permettant de trouver des fragments solubles à partir de n'importe quelle protéine: ESPRIT (*Expression of Soluble Proteins by Random Incremental Truncation*). Le principe de cette technique repose sur la création d'une librairie complète de cDNA aléatoires codant pour des fragments d'une longueur variable de la protéine d'intérêt. Certaines de ces constructions peuvent aboutir à des fragments solubles. (http://www-db.embl.de/jss/EmblGroupsOrg/g_225.html). Il est envisagé d'utiliser cette méthode pour rechercher d'autres domaines solubles de la protéine IIIa.

Dans la structure 3D de d'adénovirus obtenue par cryoMET, la densité assimilée à la protéine IIIa est constituée par deux régions globulaires (Fig. 1.12). Cette protéine semble donc organisée en deux domaines structurellement indépendants que nous devrions arriver à isoler par la méthode ESPRIT.

La détermination de la structure atomique de la protéine IIIa sera un grand pas vers l'élucidation de son rôle dans la capsid. Son placement dans la structure 3D de la capsid obtenue par MET révélera en particulier ses zones d'interaction avec les autres protéines de la capsid et permettra une attribution définitive de sa densité dans la capsid du virus.

5. *La Base de Penton de l'Adénovirus Humain de Sérotype 3*

5.1. *Introduction*

Pour certains sérotypes d'adénovirus, il arrive que lors de l'infection, les bases de penton s'assemblent pour former des sous-particules virales hautement symétriques appelées dodécaèdres. Ces particules sont constituées de 12 bases pentamérique (60 monomères) et ont été mises en évidence la première fois dans des cellules infectées par l'adénovirus humain de sérotype 3 en 1966 (Norrby, 1966). Depuis leur découverte, les dodécaèdres ont été mis en évidence dans des cellules infectées par différents sérotypes d'adénovirus humains dont le sérotype 4 (Norrby and Wadell, 1967), le sérotype 7 (Neurath et al., 1968), le sérotype 11 (Norrby, 1968), les sérotypes 9 et 15 (Norrby et al., 1967). L'observation de ces particules *in vivo* dans des cellules infectées par des adénovirus humains de sérotype 1, 2, 5, 6, 12, 16, 40 et 41 n'a pas abouti (Norrby and Ankerst, 1969; Norrby and Skaaret, 1968; Wadell and Norrby, 1969). Il faut néanmoins noter que la base de penton de l'adénovirus de sérotype 2 a été cristallisé sous forme dodécaédrique (Zubieta et al., 2005) en présence de dioxane.

Les particules dodécaédriques issues de l'HAd3 sont les plus étudiées d'un point de vue biologique et structural. Grâce à leur cavité de 80Å et à leur capacité à entrer dans les cellules, ces sous-particules virales ont été proposées comme vecteur de transfert en thérapie génique (Fender et al., 1997; Fender et al., 2003; Fuschiotti et al., 2006a).

La séquence de la base de penton est très conservée pour les sérotypes humains, il y a typiquement 70 % d'identité de séquences entre celles des HAd2 et 3. Comme le montre l'alignement de séquences entre les bases de ces deux virus (Fig. 5.1A), les différences majeures sont localisées au niveau de trois régions: le domaine N-terminal, la boucle variable et la boucle RGD. La base de penton de l'HAd3 possède 34 résidus de moins dans la boucle RGD et 9 résidus de plus dans la boucle variable par rapport à celle de l'HAd2. La structure

atomique de la base de penton de l'HAD2 a été résolue par Zubieta et al. en 2005 (Fig. 5.1B).

Dans cette structure, la boucle RGD qui est flexible est totalement absente (74 aa). La partie

A

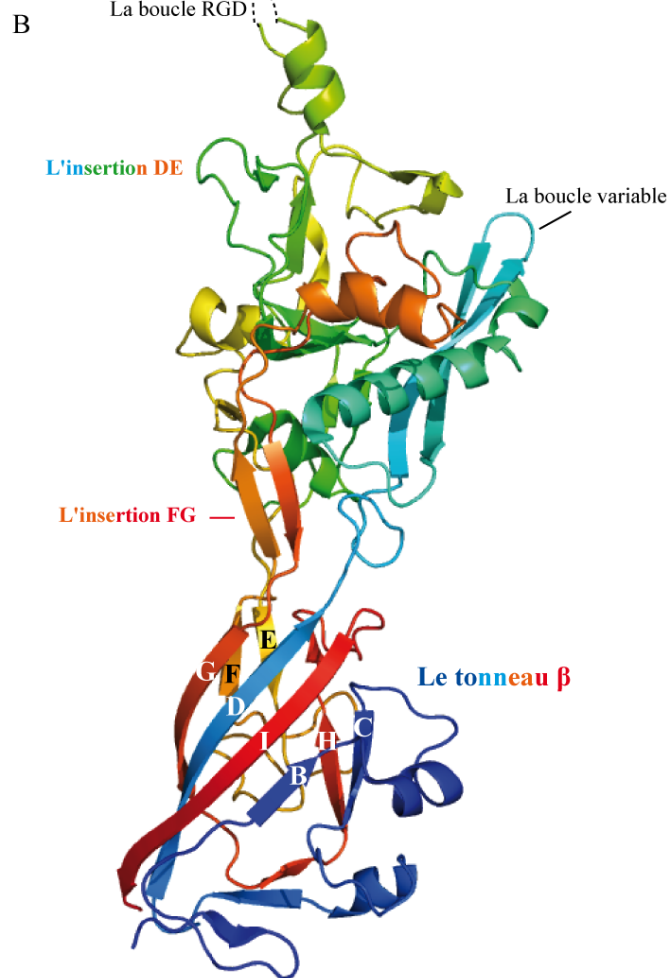


Figure 5.1: La base de penton des HAdV3 et 2.

A, Alignement des séquences primaires des bases de penton des HAd3 et 2. Les structures secondaires sont indiquées en bleu (brin β) et rose (hélice α) pour HAd2 dont la structure atomique est connue. Les boucles variable et RGD sont colorées respectivement en bleu et vert. Les flèches grises indiquent le début de la séquence des protéines cristallisées.

B, La structure cristallographique de la base de penton de l'HAd2 (Zubieta et al., 2005), avec les noms des différents domaines.

N-terminale a été également éliminée (49 premiers acides aminés) de la construction utilisée en cristallographie car cette partie présentait une forte tendance à l'auto protéolyse. Nous nous intéressons à la structure du dodécaèdre issu de l'HAd3 car nous espérons découvrir le repliement des régions manquantes dans la structure actuelle de la base de penton de l'HAd2. Nous espérons en particulier résoudre les structures de la boucle RGD et d'une partie des acides aminés composant l'extrémité N-terminale de la base. Ces domaines sont très importants pour le virus car ils permettent la fixation des intégrines (*via* le motif RGD de la boucle RGD) et des domaines WW des ligases de l'ubiquitine (*via* les deux motifs xPPxY de la partie N-terminale) (Galinier et al., 2002).

Des essais de cristallisation avaient déjà été entrepris ces dernières années par plusieurs personnes de l'EMBL sur des dodécaèdres de l'HAdV3. Des cristaux de ces particules avaient été obtenus mais leur pouvoir de diffraction n'a jamais dépassé 5 Å. Nous avons choisi de reprendre ce projet en collaboration avec le Dr. Pascal Fender de l'IBS (Institut de Biologie Structurale).

5.2. Expression et Purification

Comme pour la protéine entière décrite précédemment par Fender et al. (1997), le fragment 38-544 aa de la base de penton de l'HAd3 a été cloné dans un vecteur d'expression du système de baculovirus. La protéine recombinante est tronquée du côté N-terminal pour éviter un phénomène d'auto-protéolyse constaté au niveau des acides aminés E37-A38. Il semble de plus que cette protéolyse soit indispensable à la formation correcte des dodécaèdres (Fuschiotti et al., 2006b). La protéine est produite dans les cellules d'insecte et les particules dodécaédriques sont purifiées sur un gradient de sucrose. Après une dialyse pour éliminer le sucrose, la protéine nous a été fournie par le Dr. Fender. J'ai alors réalisé une étape de purification supplémentaire par chromatographie échangeuse d'anions (Mono Q ; *Amersham*). La figure 5.2 représente le chromatogramme généré lors de la MonoQ. Les différentes fractions correspondant aux différents pics ont été déposées sur gel SDS-PAGE dénaturant, et observées par microscopie électronique en coloration négative (Fig. 5.3).

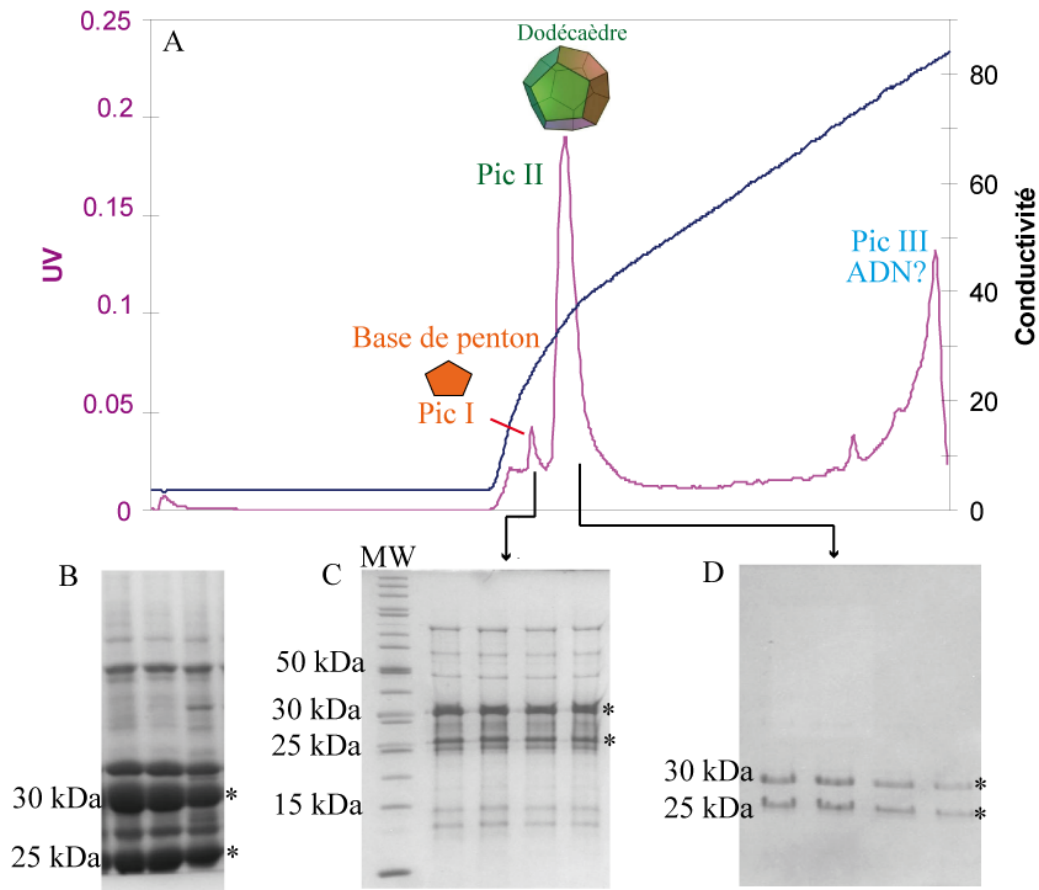


Figure 5.2: Purification de la base de penton de l'HAd3.

A, Chromatogramme à la sortie de la colonne échangeuse d'anions MonoQ, la courbe en violet correspond à l'absorption de la fraction vis-à-vis des ultra violet, la courbe noire correspond à la conductivité.

B, Gel SDS-PAGE (réalisé par le Dr. Fender) avec différentes fractions issues du gradient de sucrose, avant la MonoQ.

C, Gel SDS-PAGE (12 %) des fractions correspondant au pic I.

D, Gel SDS-PAGE (8 %) des fractions correspondant au pic II.

Sur les trois gels B, C et D les deux bandes à 25 et 30 kDa indiquées par une étoile correspondent à deux fragments de la base de penton dégradée.

Après purification, sur le gel SDS-PAGE, nous observons deux bandes de tailles différentes, à 25 et 30 kDa (Fig. 5.2D), correspondant à deux fragments de la base de penton. Pour déterminer la composition de ces bandes nous avons eu recours au séquençage N-terminal. Le plus grand des deux fragments correspond au côté N-terminal de la protéine (commençant par l'aa 38-APFVP), l'autre fragment débute par l'aa 334-ITEKQKR. La protéine est donc clivée au niveau l'aa 334 juste après le motif RGD (indiqué par * dans la figure 5.1A). Ce clivage ne se produit pas habituellement lors de la production de la base de penton, le Dr. Fender nous a informé qu'il avait ajouté à son échantillon de la DNase qui été probablement contaminé par des protéases. En attendant d'avoir un échantillon non clivé nous avons donc choisi de travailler sur cet échantillon.

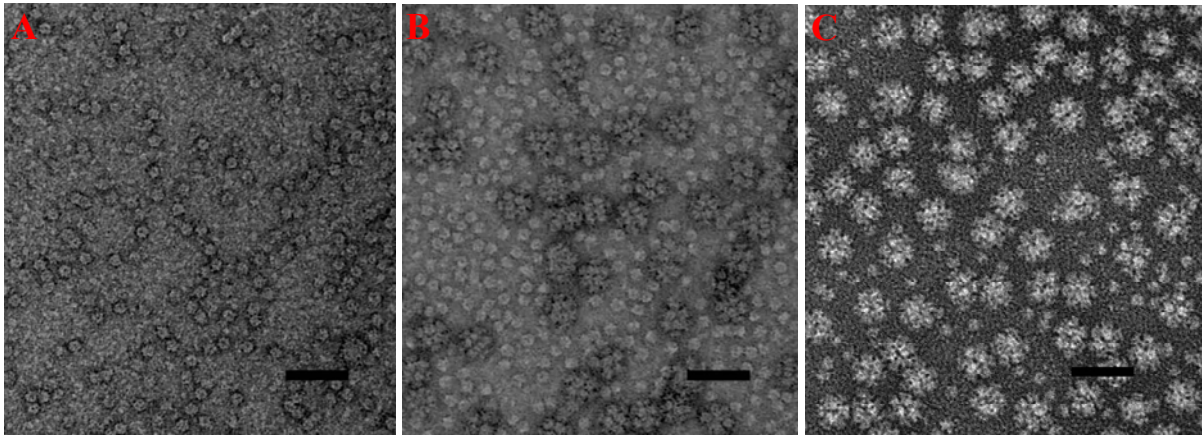


Figure 5.3: Images de MET en coloration négative des pics I et II du chromatogramme.

A, Echantillon provenant du pic I et contenant uniquement les bases de penton libres. La barre d'échelle, 50 nm.

B, Echantillon provenant du pic II et contenant un mélange de deux populations homogènes, la base de penton et le dodécaèdre.

C, Echantillon provenant du pic II et qui a été incubé une heure en présence de 10 % de dioxane et 1.6 M de sulfate d'ammonium. La composition de cette solution est identique à celle utilisée par Zubieta et al., (2005) pour cristalliser la base de penton de l'HAd2 sous forme de dodécaèdre. Il y a une nette diminution de la quantité de base libre au profit des dodécaèdres.

Les images de MET de la figure 5.3 nous indiquent que le pic I est constitué majoritairement de la base de penton et le pic II de la base de penton et du dodécaèdre. Le clivage observé au niveau de l'aa 334 ne gêne visiblement pas la formation des particules dodécaédriques. Malgré la présence de la coupure au niveau de la boucle RGD, les différentes fractions du pic II ont été rassemblées, concentrées jusqu'à 9 mg/mL puis utilisées pour tester différentes conditions de cristallisation.

5.3. Cristallisation du Dodécaèdre de l'HAdV3

La solution de dodécaèdres purifiés a été soumise à des tests de cristallisation dans des nano-gouttes avec un robot PixSys4200 (cf. le chapitre Matériels et Méthodes). Environ 600 conditions de cristallisation ont été testées à l'aide de kit de cristallisation commerciaux (listes dans le Matériels et Méthodes). Plusieurs conditions de cristallisation ont abouti à la formation de petits cristaux (Fig. 5.4). Ces cristaux apparaissent après une nuit d'équilibrage à 20°C dans des conditions de cristallisation contenant préférentiellement du PEG 4000, 6000 ou 8000 comme précipitant. Quelques exemples sont présentés dans la figure ci-dessous.

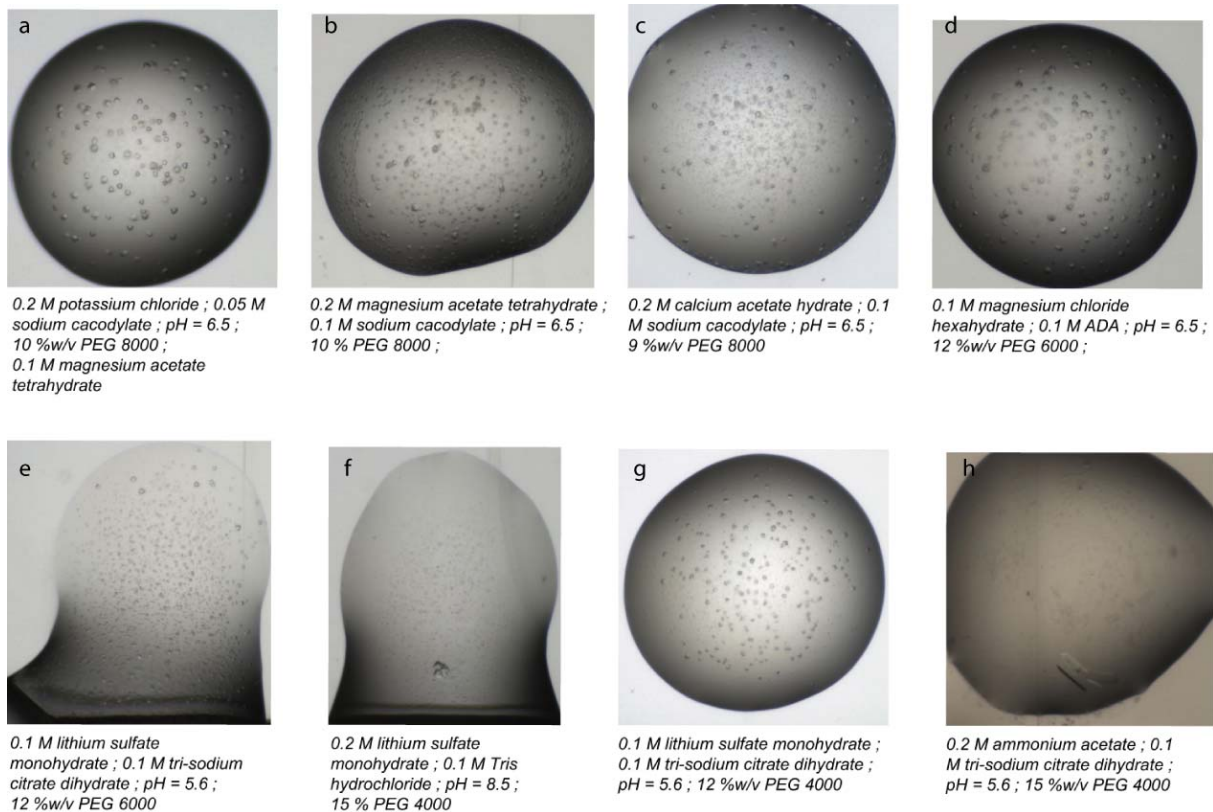


Figure 5.4: Cristaux du dodécaèdre de l’HAdV3.

Les premiers cristaux ont été obtenus dans des nano-gouttes assises par diffusion de vapeur. La goutte est composée de 0.1 μL de solution de cristallisation et de 0.1 μL de solution de protéine à 9 mg/mL, le réservoir contient 100 μL de la solution de cristallisation. Les cristaux sont obtenus après une nuit d’équilibrage à 20°C.

Les cristaux obtenus par le robot sont trop petits pour être testés directement à l’ESRF. Nous avons donc reproduits manuellement les conditions de cristallisation en utilisant des volumes plus importants (1 μL de la solution de cristallisation est mélangé à 1 μL de la solution de protéine à 9 mg/mL ; le réservoir contient dans ce cas 800 μL de la solution de cristallisation). Le temps d’apparition des cristaux ne change pas. Ces cristaux sont congelés au sein d’une boucle dans de l’azote liquide (100 K) en présence d’une solution de cryoprotectant (25% glycérol en concentration finale ajouté à la solution de cristallisation). La meilleure diffraction (7Å) a été obtenue pour les cristaux de la condition **a** (Fig. 5.4) (ID23-EH2, ESRF). La solution initiale de cristallisation utilisée est composée de 0.05 M sodium cacodylate pH = 6.5 ; 10 % pds/v PEG 8000 ; 0.2 M chlorure de potassium et 0.1 M magnésium acétate tétrahydraté. Cette condition de cristallisation a été choisie comme point de départ pour améliorer la diffraction des cristaux du dodécaèdre.

Les cristaux obtenus sont très fragiles, ils se cassent facilement lors de l'ajout de la solution de cryoprotectant et de la manipulation mécanique avec la boucle. Les stratégies adoptées pour améliorer leur diffraction visent donc à obtenir des cristaux plus solides tout en réduisant le nombre d'étapes de manipulations mécaniques:

1 – Production des cristaux directement en présence du cryoprotectant (glycérol).

L'obtention de cristaux directement en présence de glycérol permet d'éviter le transfert des cristaux dans une solution qui peut les fragiliser. Le glycérol ralentit également la vitesse d'empilement cristallin qui peut affecter le degré d'organisation dans le cristal.

2 – Additifs

En plus de la composition de la solution de cristallisation, nous avons utilisé différents additifs susceptibles d'améliorer la diffraction en stabilisant la protéine du point de vue structural. Ces additifs peuvent être des atomes lourds, des sels, des acides aminés, des agents chaotropiques, des co-facteurs, des détergents, des molécules organiques volatiles ... etc. Le dioxane fait partie de cette dernière catégorie d'additifs. Comme le montrent les images de microscopie électronique (Fig. 5.3), le dioxane (en présence de sulfate d'ammonium 1.6 M) stabilise la forme dodécaédrique et donne un échantillon homogène du point de vue oligomérique.

En pratique, un criblage de 96 additifs différents (*Additive Screen hampton research*) a été effectué à partir de la condition de cristallisation initiale à laquelle avait déjà été ajoutée 15% de glycérol. Des cristaux ont été obtenus dans quasiment toutes les conditions. Nous avons testé le pouvoir de diffraction de la plupart d'entre eux. La figure 5.5 montre 4 cristaux représentatifs des différentes formes cristallines obtenues.

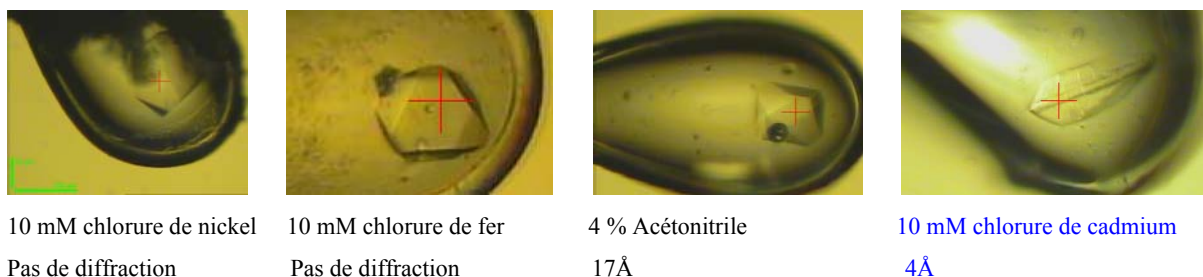


Figure 5.5: Images de 4 types de cristaux différents obtenus avec différents additifs

La condition de cristallisation commune est: 0.05 M sodium cacodylate pH = 6.5 ; 10 % pds/v PEG 8000 ; 0.2 M chlorure de potassium et 0.1 M magnésium acétate tétrahydraté plus 15 % de glycérol. Pour chacun des quatre cristaux, le type d'additif utilisé ainsi que la résolution de la diffraction sont indiqués. Les cristaux apparaissent en 1-2 jours (selon l'additif) et sont ensuite congelés dans de l'azote liquide en présence de 25% de glycérol.

Nous constatons que l'ajout de 15 % de glycérol et 10 mM de chlorure de cadmium à la condition initiale de cristallisation a permis l'augmentation des limites de résolution de 7 à 4 Å. Les cristaux obtenus dans ces conditions ressemblent plus à des plaquettes de faible épaisseur et sont très sensibles aux rayons X. En augmentant la concentration en glycérol jusqu'à 35% nous avons réussi à obtenir des cristaux tridimensionnels (Fig. 5.6A). Ces derniers diffractent jusqu'à 3.5 Å de résolution et sont beaucoup moins sensibles aux dommages d'irradiation causés par les rayons X, probablement grâce à la forte concentration en glycérol utilisée.

5.4. Collecte et traitement des données

Les données cristallographiques ont été collectées à l'ESRF (*European Synchrotron Radiation Facility*) sur la ligne de lumière ID14-EH4. Le meilleur cristal a diffracté les rayons X jusqu'à une résolution de 3.5 Å. Ce cristal du groupe d'espace $P2_1$ possède les paramètres de maille suivants: $a=264.55\text{Å}$ $b=382.46\text{Å}$ $c=268.23\text{Å}$ et $\alpha=\gamma=90^\circ$ $\beta=90.4^\circ$. L'unité asymétrique contient une particule dodécaédrique (60 monomères de base de penton de 507 acides aminés et d'une masse moléculaire de 57808 Da) avec un coefficient de Matthews (V_M) de $3.9 \text{Å}^3/\text{Da}$ correspondant à 68 % de solvant dans le cristal (Matthews, 1968). Les données ont été intégrées et mises à l'échelle avec XDS (Kabsch, 1988). Les statistiques obtenues après ces traitements sont indiquées dans le tableau 5.1.

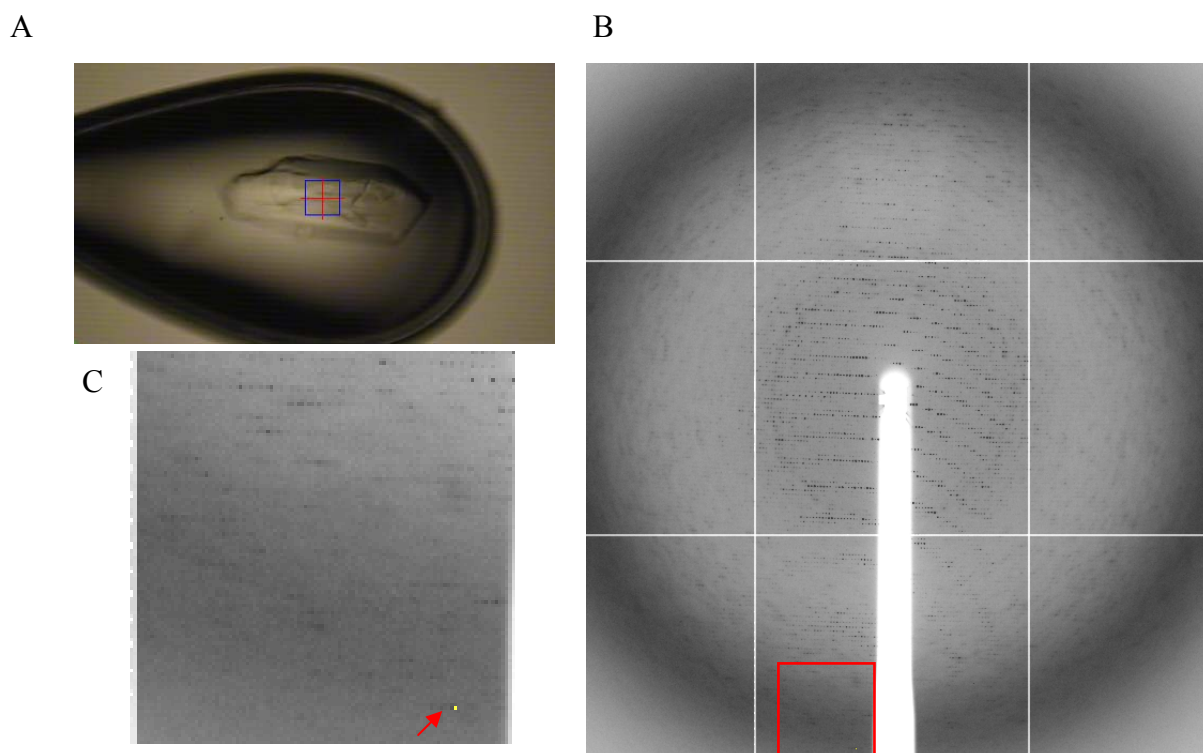


Figure 5.6: Diffraction des rayons X par les cristaux du dodécaèdre de l'HAd3.

A, Photo du cristal de dodécaèdres obtenu dans la condition de cristallisation initiale plus 35% de glycérol et 10 mM CdCl₂. La photo a été prise alors que le cristal était déjà congelé et en place pour une collecte de données à l'ESRF.

B, Cliché de diffraction du cristal obtenu sur la ligne de lumière ID14-EH4 (ESRF).

C, Zoom de la région indiquée par un cadre rouge dans la partie B. Des taches de diffraction sont observées jusqu'à 3.5 Å de résolution.

Collecte des données

source	ESRF ID14-EH4
Détecteur	ADSC
Longueur d'onde (Å)	0.976
Groupe d'espace	P2 ₁
Paramètre de maille (Å)	a=264.55 b=382.46 c=268.23
	$\alpha=\gamma=90^\circ$ $\beta=90.4^\circ$
Contenu en solvant	68 %
Résolution Å	49-4.01
N° de réflexions observées	633670
N° de réflexions uniques	352673
Complétude (%)	78.8 (62.6)
Redondance	1.79 (1.77)
Le facteur R (%)	15.4 (45.1)
I/σ (I)	5.79 (1.95)

Tableau 5.1: Statistiques cristallographiques de collecte des données sur les cristaux du dodécaèdre. Les données à 3.5 Å de résolution (statistiques non montrées) sont médiocres mais peuvent être exploités à cause de la symétrie non cristallographique avec les 60 copies dans l'unité asymétrique.

L'affinement de la condition de cristallisation initiale donnant des cristaux du dodécaèdre a donc permis l'obtention de cristaux tridimensionnels qui diffractent jusqu'à 3.5 Å de résolution. Le travail de résolution de la structure sera réalisé après la soutenance de ma thèse durant un stage post-doctoral de 3 mois. Nous envisageons la résolution de la structure par remplacement moléculaire en utilisant la structure de la base de penton de l'HA2 comme modèle (Zubieta et al., 2005). Les nombreux opérateurs de symétries non cristallographiques (NCS) du dodécaèdre seront utilisés pour moyennner les cartes de densité électronique et améliorer ainsi la qualité des cartes.

Essais de cristallisation de la base de penton non clivée

Le Dr Fender nous a fourni un échantillon de dodécaèdre non clivé au niveau de la boucle RGD (Fig. 5.7). La protéine a été purifiée comme décrit précédemment et les essais de cristallisation ont été entrepris dans des conditions proches de la condition initiale **a**. Cette protéine cristallise après une nuit d'équilibration à 20°C. Les cristaux sont de grande taille (jusqu'à 300 µm) mais malheureusement ils ne diffractent qu'à basse résolution (environ 30 Å). Nos résultats préliminaires nous permettent de conclure que la coupure au niveau de la boucle RGD permet probablement un empilement plus compact et plus ordonné des dodécaèdres. La structure de la base de penton de l'HA2 a également été obtenue avec des protéines présentant une coupure dans la boucle RGD.

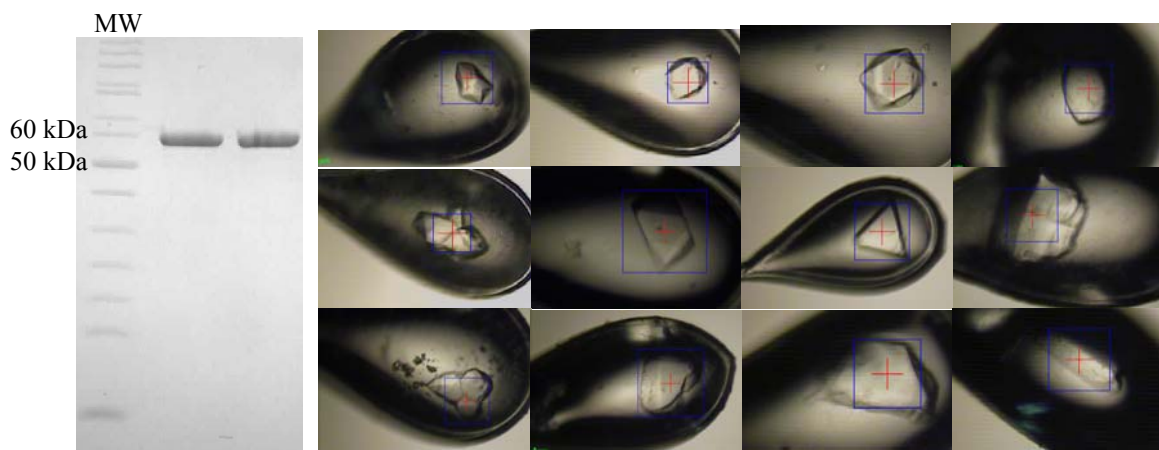


Figure 5.7: La base de penton d'HA3 non clivée au niveau de la boucle RGD.

A, Gel SDS-PAGE de la base de penton non clivée après purification sur la colonne échangeuse d'anions MonoQ.

B, Photos de quelques cristaux du dodécaèdre. Tous ces cristaux ont été obtenus dans la condition initiale (0.2 M chlorure de potassium ; 0.05 M sodium cacodylate pH = 6.5 ; 10 % w/v PEG 8000 ; 0.1 M magnésium acétate tétrahydraté) additionné de 25-35% glycérol et 10 mM CdCl₂ ou 20 mM CaCl₂.

5.5. Conclusion et Perspective

Nous avons constaté, lors des collectes des données sur les cristaux du dodécaèdre issus de l'HAAd3, qu'un clivage au niveau de la boucle RGD permet une augmentation significative des limites de diffraction. Cette boucle est probablement flexible et perturbe l'empilement cristallin.

Il a été montré que le dodécaèdre issu de l'HAAd3 était apte à être internalisé par des cellules exprimant des héparanes sulfates à leur surface (Fender et al., 2008) alors que les particules virales entières n'en sont pas capables (Dehecchi et al., 2000). L'analyse de la séquence primaire de la base de penton a mis en évidence une région fortement basique (KQKR) proche du motif RGD qui pourrait être un site de fixation potentiel aux héparanes sulfates. Une reconstruction 3D préliminaire obtenue à partir d'images de cryoMET des dodécaèdres en complexe avec un oligosaccharide d'HS composé de 8 saccharides a effectivement montré une densité supplémentaire proche de la boucle RGD et attribuée au HS (Fender et al., 2008). Dans un futur proche nous envisageons des essais de cristallisation de ces complexes (dodécaèdre non clivé + HS). Ceci permettra non seulement l'identification de ou des sites de fixation des HSPGs mais aussi l'analyse de leur mécanisme d'interaction à l'échelle atomique. Nous espérons également que l'interaction avec les HS permettra la stabilisation d'une conformation de la boucle RGD et donc l'obtention des cristaux de la protéine non clivée qui diffracteraient à haute résolution.

6. *La Fibre Courte de l'Adénovirus Aviaire de Sérotype 1*

6.1. *Introduction*

Deux étapes permettent l'internalisation de l'adénovirus: dans un premier temps la fibre s'attache au récepteur primaire à la surface de la cellule hôte. La base de penton se lie ensuite aux intégrines *via* la boucle RGD entraînant l'internalisation du virus. L'adénovirus aviaire de type 1 (CELO) comporte deux fibres liées à la même base de penton. La base de penton ne contenant pas de motif RGD, ce virus utilise probablement une voie alternative à celle décrite précédemment. Le CELO sauvage infecte des cellules CHO exprimant le CAR. Afin de connaître le rôle respectif de chacune des fibres dans cette interaction, des mutations ont été introduites dans le génome pour désorganiser soit la fibre longue soit la fibre courte (Tan et al., 2001). Le virus CELO dont le génome est dépourvu de la fibre courte est incapable de générer des particules virales, la fibre courte est donc essentielle pour la propagation du virus. Le virus dépourvu de la fibre longue est quant à lui produit et capable d'infecter les cellules de poules, mais il perd sa capacité à infecter des cellules humaines exprimant le CAR (Tan et al., 2001). La fibre longue est donc responsable de l'interaction avec le CAR à la surface des cellules humaines (Tan et al., 2001). Le récepteur de la fibre courte reste encore inconnu. Le mécanisme d'internalisation de ce virus dans les cellules humaines reste inconnu, il est 'intégrine indépendant' et fait probablement intervenir les deux fibres. Il est possible que lors de l'infection la fibre longue s'attache à la surface de la cellule *via* le CAR et que la fibre courte interagisse avec le récepteur secondaire et induise l'endocytose (cette deuxième interaction remplacerait l'interaction entre le motif RGD et les intégrines). Cette hypothèse peut s'appliquer aux HAd40 et 41 qui comportent également deux fibres: une courte et une longue, la fibre longue peut interagir avec le CAR alors que le récepteur de la fibre courte reste inconnu, leur base de penton ne contient pas de motif RGD (Favier et al., 2004).

Dans l'optique de révéler les propriétés de la fibre courte d'adénovirus aviaire nous avons résolu sa structure atomique par cristallographie.

6.2. ARTICLE II: Structure of the C-terminal head domain of the fowl adenovirus type 1 short fibre

Majida El Bakkouri¹, Elena Seiradake^{2§}, Stephen Cusack², Rob W.H. Ruigrok¹ and Guy Schoehn^{1,3*}.

¹Unit for Virus Host Cell Interactions, UMR5233 UJF-EMBL-CNRS, CIBB, 6 rue Jules Horowitz, BP181, 38042 Grenoble Cedex 9, France.

²EMBL Grenoble Outstation, 6 rue Jules Horowitz, BP181, 38042 Grenoble Cedex 9, France.

§ Present address: The Division of Structural Biology, Henry Wellcome Building for Genomic Medicine, Roosevelt Drive, Oxford, OX3 7BN, UK.

³ Institut de Biologie Structurale Jean –Pierre Ebel, 41 rue Jules Horowitz, 38027 Grenoble Cedex 1, France.

* Corresponding author. UVHCI, UMR 5233 UJF-EMBL-CNRS, CIBB, BP 181, 38042 Grenoble cedex 9, France. Tel: (33) 4 76 20 94 23 Fax: (33) 4 76 20 94 00

Email: schoehn@embl.fr

Keywords: structure, X-ray crystallography, CELO short fibre head domain, adenovirus



Structure of the C-terminal head domain of the fowl adenovirus type 1 short fibre

Majida El Bakkouri^a, Elena Seiradake^{b,1}, Stephen Cusack^b, Rob W.H. Ruigrok^{a,?}, Guy Schoehn^{a,c,?}^a Unit for Virus Host Cell Interactions, UMR5233 UJF-EMBL-CNRS, CIBB, 6 rue Jules Horowitz, BP181, 38042 Grenoble Cedex 9, France^b EMBL Grenoble Outstation, 6 rue Jules Horowitz, BP181, 38042 Grenoble Cedex 9, France^c Institut de Biologie Structurale Jean -Pierre Ebel, 41 rue Jules Horowitz, 38027 Grenoble Cedex 1, France

ARTICLE INFO

Article history:

Received 4 April 2008

Returned to author for revision 1 May 2008

Accepted 14 May 2008

Available online xxxx

Keywords:

Structure

X-ray crystallography

CELO short fibre head domain

adenovirus

ABSTRACT

There are more than 100 known adenovirus serotypes, including 50 human serotypes. They can infect all 5 major vertebrate classes but only Aviadenovirus infecting birds and Mastadenovirus infecting mammals have been well studied. CELO (chicken embryo lethal orphan) adenovirus is responsible for mild respiratory pathologies in birds. Most studies on CELO virus have focussed on its genome sequence and organisation whereas the structural work on CELO proteins has only recently started. Contrary to most adenoviruses, the vertices of CELO virus reveal pentons with two fibres of different lengths. The distal parts (or head) of those fibres are involved in cellular receptor binding. Here we have determined the atomic structure of the short-fibre head of CELO (amino acids 201–410) at 2.0 Å resolution. Despite low sequence identity, this structure is conserved compared to the other adenovirus fibre heads. We have used the existing CELO long-fibre head structure and the one we show here for a structure-based alignment of 11 known adenovirus fibre heads which was subsequently used for the construction of an evolutionary tree. Both the fibre head sequence and structural alignments suggest that enteric human group F adenovirus 41 (short fibre) is closer to the CELO fibre heads than the canine CADV-2 fibre head, that lies closer to the human virus fibre heads.

© 2008 Elsevier Inc. All rights reserved.

Introduction

Adenoviruses are non-enveloped, icosahedral viruses, with a linear, 26–45 kbp, double-stranded DNA molecule (Horowitz, 1996; Rekosh et al., 1977; Stewart et al., 1993). Whereas adenoviruses can be found infecting all 5 major vertebrate classes (Davison et al., 2003) only Aviadenovirus that infects birds and Mastadenovirus that infect mammals have been well studied. Avian (fowl) adenovirus 1 or CELO (chicken embryo lethal orphan) virus was isolated in 1957 (Yates and Fry, 1957) and is responsible for mild respiratory pathologies in birds (Dubose and Grumbles, 1959). CELO virus can be isolated from healthy chickens, it does not cause visible disease when it is experimentally introduced into chickens (Cowen et al., 1978) and has not been associated with major economic losses or pathologies in chickens. Most studies on CELO virus have focussed on its genome sequence and organisation whereas virus–cell and virus–host interactions are still poorly understood and structural work on CELO virus proteins has only recently started (Guardado-Calvo et al., 2007; Xu et al., 2007).

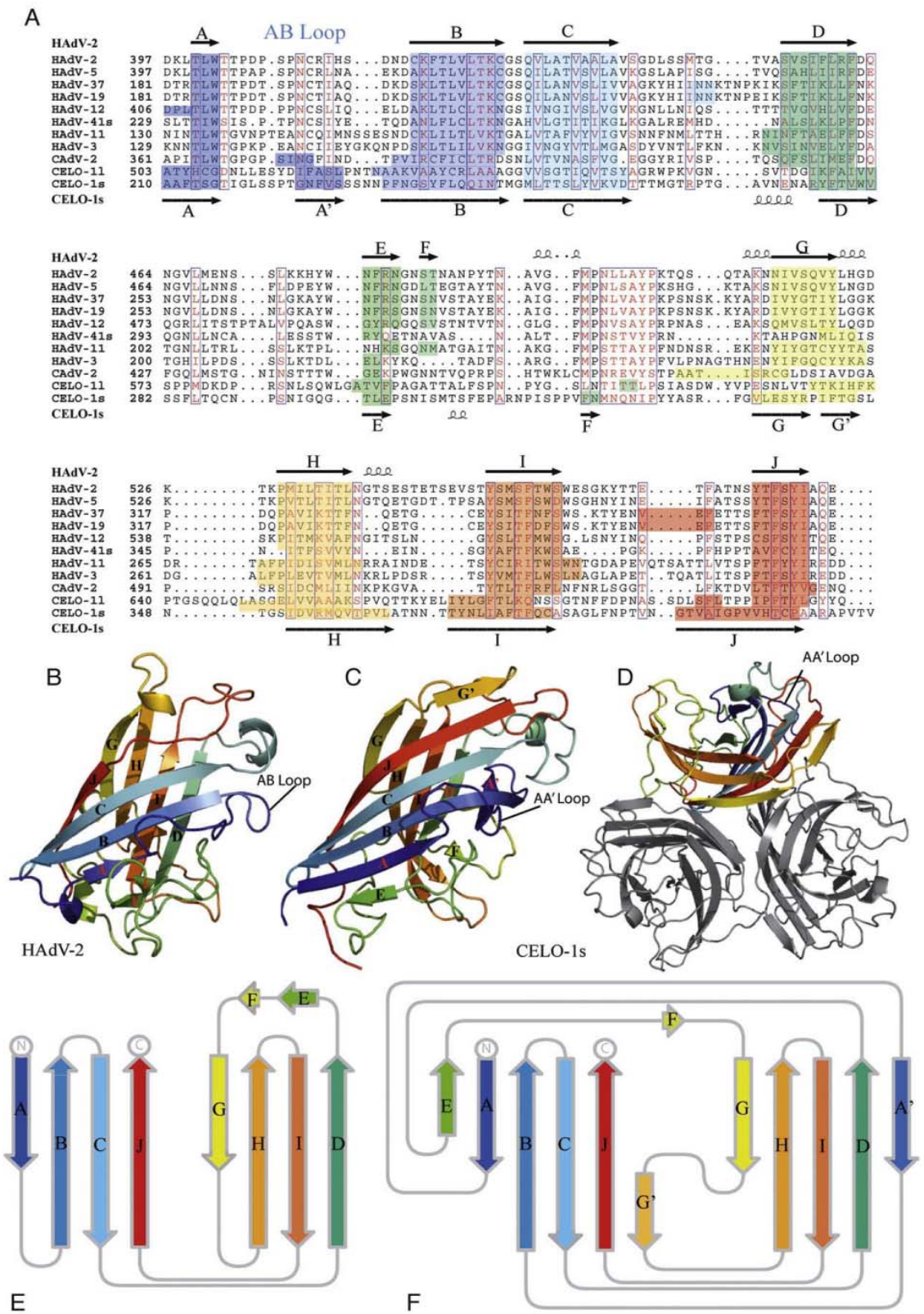
The CELO virus genome, 43,804 bp in length, was sequenced (Chiocca et al., 1996) and its transcriptional organization established (Payet et al., 1998). Like the human adenoviruses, CELO virus possesses an icosahedral capsid with a pseudo T number of 25 and a diameter of about 90 nm. Adenovirus capsids are built with 240 copies of the hexon, a homo-trimeric protein that forms the 20 facets of the icosahedron and from 12 pentons that form the 12 vertices (Fabry et al., 2005; Stewart et al., 1991, 1993). The X-ray structure of the CELO virus hexon has been solved at 3.9 Å resolution and has a fold that is similar to human hexons (Xu et al., 2007). For most human and animal adenoviruses, the penton is a non-covalent complex between the pentameric penton base and a single trimeric fibre protein (Fuschiotti et al., 2006; Zubieta et al., 2005). The fibre is a homo-trimer with three distinct regions: the head, the shaft and the tail. The fibre binds to the penton by its N-terminal tail and to the cellular receptor with the head domain. In contrast to other adenoviruses, early observations of CELO virus by electron microscopy revealed pentons with two fibres (Gelderblom and Maichle-Lauppe, 1982; Hess et al., 1995; Laver et al., 1971). The two fibres are of different lengths: the short fibre (fibre 2) is about 8.5 nm long for 410 amino acids and extends straight from the top of the penton base whereas the bottom part of the long fibre (42.5 nm and 710 amino acids in length) lies almost parallel to the top of the penton base (Hess et al., 1995).

The structure of the long-fibre head (amino acids 496–710) was recently solved by X-ray crystallography (Guardado-Calvo et al., 2007). It shows the same over-all fold as the other known adenovirus fibre

? Corresponding authors. R. Ruigrok is to be contacted at UVHCI, UMR 5233 UJF-EMBL-CNRS, CIBB, BP 181, 38042 Grenoble cedex 9, France. Fax: +33 4 76 20 94 00. G. Schoehn, Fax: +33 4 76 20 94 00.

E-mail addresses: ruigrok@embl.fr (R.W.H. Ruigrok), schoehn@embl.fr (G. Schoehn).

¹ Present address: The Division of Structural Biology, Henry Wellcome Building for Genomic Medicine, Roosevelt Drive, Oxford, OX3 7BN, UK.



Please cite this article as: El Bakkouri, M., et al., Structure of the C-terminal head domain of the fowl adenovirus type 1 short fibre, Virology (2008), doi: 10.1016/j.virol.2008.05.011

fibre, Virology

heads such as human adenovirus type 5 fibre head (Xia et al., 1994), type 2 (van Raaij et al., 1999), type 12 (Bewley et al., 1999), type 3 (Durmort et al., 2001), types 37 and 19p (Burmeister et al., 2004) and the type 41 short fibre head (Seiradake and Cusack, 2005) and the canine adenovirus type 2 fibre head (Seiradake et al., 2006). Each monomer of the trimeric head is formed by an 8–9 stranded antiparallel β -sandwich connected by loops. The main differences between the various fibre head structures are found in the loop regions that connect the β -strands and in the definition of one or two of the shorter strands. Even though the structures of all fibre heads are very similar, there is very little sequence similarity between the heads of the human adenovirus fibres and even less between fibres of human and animal viruses. The long and short CELO fibre heads show only 15% sequence identity.

Here we have determined the atomic structure of the short-fibre head of CELO (amino acids 201–410) at 2.0 Å resolution. We have used this structure and that of the long-fibre head for a structure based alignment of 11 adenovirus fibre heads that was subsequently used for the construction of an evolutionary tree. Interestingly, both the fibre head sequence and structural alignments suggest that human group F adenovirus 41 (short fibre) is closer to the CELO fibre heads than the canine CAdV-2 fibre head that lies closer to the other human virus fibre heads.

Results and discussion

The head of the short fibre of CELO virus was cloned and expressed in *E. coli*. The protein was crystallised in 0.1 M lithium sulphate, 0.1 M tri-sodium citrate pH 5.6 and 12% w/v polyethylene glycol (PEG) 4000 and its structure solved by isomorphous replacement using a Pt compound. The structure resembles that of all other known structures of adenovirus fibre heads (Fig. 1). Each monomer of the trimer folds into an anti parallel beta sandwich made up by two β -sheets which pack together at an angle of about 45°.

Comparison of the fibre head structures of CELO virus (short fibre) and HAdV2

The nomenclature proposed by Xia et al. (1994) for the fibre head of HAdV5 can also be applied to the short CELO virus fibre head: the beta sandwich is composed of sheets ABCJ and GHID (Fig. 1). The major differences between the CELO and HAdV2/5 heads are: 1; the CELO virus A- and J-strands are longer than the corresponding HAdV2 strands, allowing the A-strand to pack to the E-strand and the J-strand to the G'-strand, see below. 2; the HAdV2 AB loop is interrupted by β -strand A' in the CELO virus fibre head. 3; strand G is interrupted by a Proline residue in the CELO virus head resulting in two strands, G and G' where strand G occupies the same position as in the HAdV2 head but where strand G' contributes to the opposite ABCJ sheet. Point 1 plus point 3 result in a β -sheet that does not have 4 strands but 6: EABCJG'. The extra strand A' packs onto strand D in the opposite sheet, which, therefore, is composed of 5 strands and not of 4 as in the HAdV2 fibre head. The overall effect of these differences is that the CELO fibre head is taller with wider sheets but with shorter loops. The fact that the monomer is taller does not lead to a more compact trimer. The percentage of interface buried upon trimerization is 27 for HAdV2 and

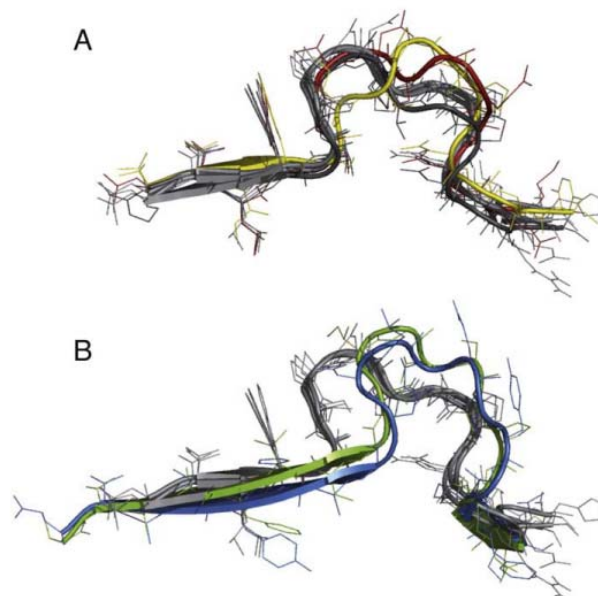


Fig. 2. Comparison of AB loop-structures of CAR-binding and non-CAR-binding fibre heads. Ribbon and line representations of the AB loops after structural superposition of different adenovirus fibre head models with Coot (Emsley and Cowtan, 2004). A, AB loops of HAdV-2, 5, 12 (light grey) ((1QHV) (van Raaij et al., 1999); (1KNB) (Xia et al., 1994); (1KAC) (Bewley et al., 1999)), HAdV-37 (dark grey) ((2J12) (Seiradake et al., 2006)), HAdV-41s (yellow) ((2BZV) (Seiradake and Cusack, 2005)) and HAdV-11 (red) ((2O39) (Persson et al., 2007)). Note that the conformation of HAdV-37 AB loop corresponds to that found in complex with CAR D1. B, AB loops of HAdV-2, 5, 12 (light grey). AA' loops of the CELO short (green) and CELO long fibres (blue) ((1IUM) (Guardado-Calvo et al., 2007)). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

only 21 for the CELO short fibre head, due to the fact that the bottom of the CELO fibre head spreads out and sheets A and E hardly contribute to the trimer interface. The main difference between the structures of the heads of the long and short CELO fibres is that in the long fibre head strand A' does not cross over to the GHIJ sheet but continues in the same direction next to strand B (Guardado-Calvo et al., 2007). This is due to the fact that the B-strand is more strongly curved. In contrast to the long fibre J-strand is broken up in two parts that are separated by a TPP sequence that continues in the same direction as the second part of the J strand and is also in a rather extended conformation, the short fibre J-strand is uninterrupted (Guardado-Calvo et al., 2007).

Receptor binding

The various human adenoviruses bind to different primary receptors. The best known receptor is CAR (coxsackie and adenovirus receptor) (Bergelson et al., 1997) that is used by subtypes 2, 5, 12 and 41 long fibre (list non-exhaustive) (Bergelson et al., 1997; Bewley et al., 1999; Roelvink et al., 1998; Tomko et al., 1997). Other human viruses use CD46 (subtypes 3, 7, 11 and 35) (Fleischli et al., 2007; Gaggar et al.,

Fig. 1. Adenovirus fibre head sequences and structures. A, Sequence alignment of all adenovirus fibre heads with known crystal structures: HAdV-2 (van Raaij et al., 1999); HAdV-5 (Xia et al., 1994); HAdV-12 (Bewley et al., 1999); HAdV-11 (Persson et al., 2007); HAdV-3 (Durmort et al., 2001); HAdV-37, HAdV-19p (Burmeister et al., 2004); HAdV-41s short (Seiradake and Cusack, 2005); CAdV-2 (Seiradake et al., 2006); CELO-11 long (Guardado-Calvo et al., 2007) and CELO-1s short fibre head. Sequences were aligned with ClustalW (Thompson et al., 1994), edited manually with SEAVIEW (Galtier et al., 1996) and visualized with ESPript (Gouet et al., 1999). B, Ribbon diagram of a monomer of the HAdV-2 fibre head, rainbow-coloured from blue (N terminus) to red (C terminus) made with PyMOL (DeLano, 2002). The β -strands are labelled according to the nomenclature introduced by Xia et al. (Xia et al., 1994). C, Ribbon diagram of a monomer of the CELO short fibre head, colours as in B. D, Ribbon diagram of CELO short fibre head trimer looking along the 3-fold axis toward the virus. Two monomers are coloured in grey, the third monomer is rainbow-coloured, like in C. E, Strand diagram for the HAdV-2 fibre head monomer (TopDaw, (Bond, 2003)). The colour codes are the same as the one used in part B. F, Strand diagram for CELO short fibre head (TopDaw (Bond, 2003)). The colour codes are the same as the one used in part C. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

Please cite this article as: El Bakkouri, M., et al., Structure of the C-terminal head domain of the fowl adenovirus type 1 short fibre, Virology (2008), doi: 10.1016/j.virol.2008.05.011

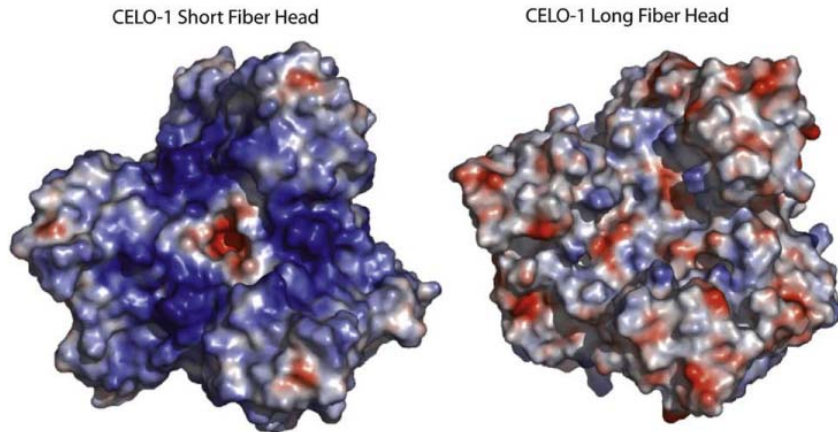


Fig. 3. Comparison of the surface potentials of the CELO short (left) and long (right) fibre heads. Colours are according to the calculated electrostatic surface potential and range from red (potential of -5 kT) to blue ($+5$ kT). The molecular surface was generated using PyMOL (DeLano, 2002) and electrostatic surface potential was calculated using APBS (Holst et al., 1994). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

2003; Persson et al., 2007; Segerman et al., 2003; Sirena et al., 2004) or sialic acid (subtypes 19 and 37) (Burmeister et al., 2004) as receptors. In order to be precise, the recombinant HAd37 fiber knob but also the intact virus bind to CAR *in vitro* but *in vivo* this virus does not use CAR to enter host cells (Wu et al., 2001). Concerning CELO virus, a mutagenesis study showed that the short fibre is essential for infecting chicken cells although the avian receptor responsible for binding of this short fibre is still unknown (Tan et al., 2001). CELO virus could infect CAR expressing mammalian cells but for this the presence of the long-fibre was essential (Tan et al., 2001). However, Guardado-Calvo and coworkers were unable to show *in vitro* binding between the long-fibre head and domain 1 of CAR by surface resonance experiments (Guardado-Calvo et al., 2007). In the HAdV types 2 and 5 the loop most important for CAR binding is the AB loop. Seiradake and Cusack (2005) have suggested that the fibre heads that bind to CAR have a specific conformation of the AB loop. Fig. 2A shows the structures of the AB loops of HAdV2, 5, 12 (in light grey) and 37 (in dark grey) that all bind CAR. The conformation of the HAdV37 loop is that of the CAR-bound structure (Seiradake et al., 2006). Also shown in Fig. 2A are the loops of HAdV41 short fibre (yellow) and that of HAdV11 (red) that do not bind CAR (Fleischli et al., 2007; Roelvink et al., 1998) and have distinctly different conformations.

Fig. 2B again shows the HAdV2, 5 and 12 AB loops (in grey) plus the AA' loops of the CELO short (green) and long fibres (blue). Due to the fact that the CELO A strand is longer and to the presence of the extra A' strand, the AA' loop is squeezed between the two strands and adopts a totally different conformation from that of the CAR-binding loops. This

difference can also be seen in the images of the monomers of HAdV2 and CELO in Figs. 1B and C. Therefore, it seems unlikely that the long or short fibres would bind to CAR in the same manner as human adenoviruses types 2, 5 and 12.

Fig. 3 shows the surface charges of the short and long fibres of CELO. The long-fibre head has a slightly acidic to neutral top surface, corresponding to an isoelectric point (pI) of 6.1. The top surface of the short-fibre head is more varied in charge with strongly positively charged patches at the monomer interfaces, the pI of this head is 8.5. These differences in surface charges suggest that the two fibres bind to different receptors. Because of the positive charges on the short fibre head we have tested the binding of sialic acid to both fibre heads. Sialic acid is used as receptor by subgroup D adenoviruses (serotypes 8, 15, and 19 and 37) (Arnberg et al., 2000; Burmeister et al., 2004) whose fibre heads all have high pI values of around 9. However, we were not able to show binding of either of the CELO fibre heads to sialic acid-containing glycoprotein by surface plasma resonance (data not shown).

Evolution

We compared the avian fibre head structure with that of the fibre head of bacteriophage PRD1 (Merckel et al., 2005). It has been suggested that adenovirus and PRD1 fibre heads may be evolutionary related but that a characteristic difference between the adenovirus and bacteriophage fibre heads is that both the adenovirus sheets contain 4 strands whereas the PRD1 sheets contain 5 strands (Bamford et al., 2005; Merckel et al., 2005). Because the CELO heads contain sheets

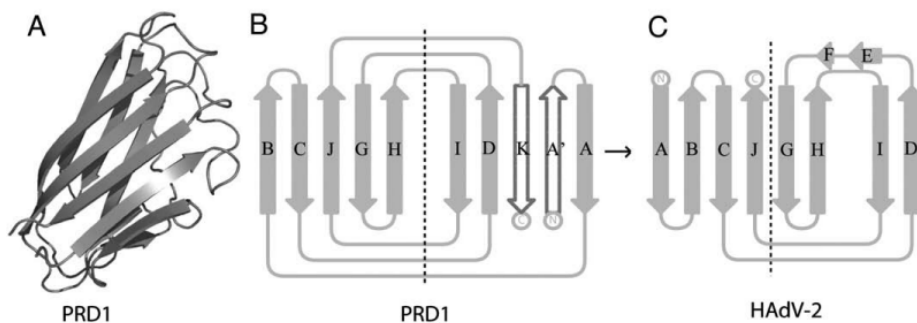


Fig. 4. Structure comparison of the HAdV-2 fibre head and PRD1 vertex protein P5. A, Ribbon diagram of the P5 monomer ((1YQ8) (Merckel et al., 2005)). B, Strand diagram for the PRD1 vertex protein P5. C, Strand diagram for the HAdV-2 fibre head ((1QHV) (van Raaij et al., 1999)) (Figure A was made with PyMOL (DeLano, 2002) and figures B and C were made with TopDaw, (Bond, 2003)).

with 5 and 6 strands, such difference is not characteristic. Fig. 4 shows the structure of a monomer of the PRD1 head plus the strand diagram for the PRD1 and HAdV2 heads (adapted from Fig. 5 in Merckel et al. (2005)). These authors have suggested that deletion of the first (A) and the last strand (K) in the PRD1 head and shift-over of the A strand to the B-strand could explain the difference between the two heads.

Comparison of Figs. 4B and C with Fig. 1F clearly shows that the adenovirus heads are much closer in structure to each other than to the bacteriophage head and that the extra strands per sheet have evolved differently for the PRD1 and the CELO heads. Another characteristic difference between the human adenovirus and the bacteriophage head is that the PRD1 sheets cross at an angle between 25 and 30° whereas

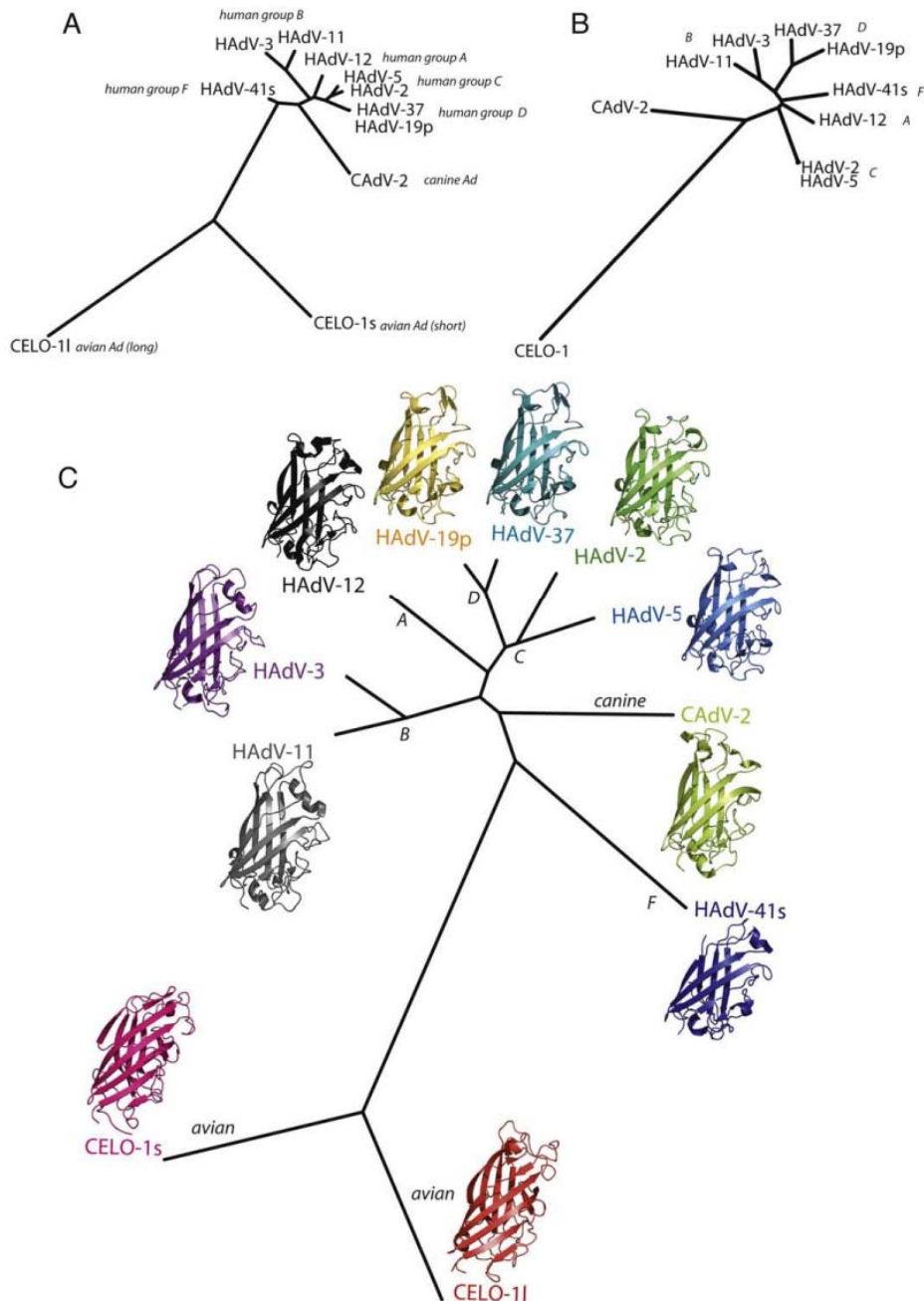


Fig. 5. Phylogenetic trees of adenovirus proteins. A, Distance tree based on Ad fiber head sequences. Protein sequences were aligned as in Fig. 1. The tree was generated with the PHYLIP package (Felsenstein, 1989; Felsenstein, 1997). B, Distance tree based on penton base sequences (PHYLIP). C, Phylogenetic tree of adenovirus fiber head domains based solely on structural comparison using MAPS (Lu, 2000). The PHYLIP package (Felsenstein, 1989, 1997) was used to draw the tree. We compared 11 fiber head structures available from the protein data bank: HAdV-2 (1QHV) (van Raaij et al., 1999); HAdV-5 (1KNB) (Xia et al., 1994); HAdV-12 (1KAC) (Bewley et al., 1999); HAdV-11 (2O39) (Persson et al., 2007); HAdV-3 (1H7Z) (Durmort et al., 2001); HAdV-37 (1UXE); HAdV-19p (1UXB) (Burmeister et al., 2004); HAdV-41s short fiber head (2BZV) (Seiradake and Cusack, 2005); CAdV-2 (2J2J) (Seiradake et al., 2006); CELO-11 long fiber head (1IUM) (Guardado-Calvo et al., 2007) and CELO-1s short fiber head.

Please cite this article as: El Bakkouri, M., et al., Structure of the C-terminal head domain of the fowl adenovirus type 1 short fiber, Virology (2008), doi: 10.1016/j.virol.2008.05.011

Table 1
Sequence identity and root mean square deviations (RMSD) between different Ad fibre heads as calculated by Secondary Structure Matching (Krissinel and Henrick, 2004)

Sequence identity (%)											
	HAdV2	HAdV5	HAdV37	HAdV19	HAdV12	HAdV41s	HAdV11	HAdV3	CAdV2	CELO-11	CELO-1s
HAdV2		66.7	55.8	55.8	40.8	32.5	37.5	38.3	23.3	16.7	8.3
HAdV5	1.43		55.8	55.8	46.7	33.3	33.3	34.2	22.5	15	10
HAdV37	1.558	1.158		99.2	46.7	29.2	37.5	36.7	26.7	14.2	12.5
HAdV19	1.498	1.138	0.321		46.7	29.2	37.5	36.7	26.7	14.2	12.5
HAdV12	1.579	1.551	1.468	1.427		43.3	30.0	36.7	25.8	14.2	10
HAdV41s	1.931	1.777	1.889	1.847	1.733		31.7	30	25.0	14.2	10.8
HAdV11	2.252	1.886	1.728	1.714	1.965	2.292		54.2	21.7	13.3	15.8
HAdV3	2.034	1.772	1.707	1.692	1.886	2.184	1.152		20.8	14.2	10.8
CAdV2	2.248	2.127	2.002	1.929	2.119	2.450	2.297	2.322		15	14.2
CELO-11	2.907	2.863	2.566	2.503	2.653	2.891	2.834	2.648	2.672		15
CELO-1s	2.835	2.825	2.726	2.638	2.865	3.003	3.132	3.011	3.149	2.607	

RMSD (Å).

the adenovirus sheets cross at about 45°. Here as well the CELO head is characteristic of the adenoviruses crossing at about 50°.

Adenovirus phylogenesis shows co-evolution of the viral subtypes with their host species (Davison et al., 2003). The evolutionary trees are easily visualised using conserved proteins like hexon (Davison et al., 2003) penton base (Madisch et al., 2007) or non-capsid proteins. Comparison of the trees derived from hexons and from pentons suggests recombination events between adenovirus genomes (Madisch et al., 2007). However, the loops on the hexons, on the pentons and virtually the entire fibre are subject to antigenic pressure and show a strong accumulation of mutations. This is the reason why these parts of the capsid are considered to be less suitable for the visualisation of the evolutionary relationships between the viruses. The variability of the fibre sequences could lead to the lining up of β -strand regions in one head with loop regions in other heads.

Table 2
Crystallographic data collection and refinement statistics

Data collection		
	Native	PIP
Detector	ADSC Q315r	ADSC Q315r
Space group	P4 ₁ 2 ₁ 2	P4 ₁ 2 ₁ 2
Unit cell (a, b, c, Å)	235.75, 235.75, 61.75	233.81, 233.81, 61.85
($\alpha, \beta, \gamma, ^\circ$)	90.00, 90.00, 90.00	90.00, 90.00, 90.00
Wavelength (Å)	0.933	0.933
Resolution	47.6–2.0 (2.10–2.00)	47.6–2.5 (2.65–2.50)
Unique reflections	117179 (15721)	112908 (18068)
Multiplicity	3.63 (3.61)	3.8 (3.8)
Completeness (%)	99.8 (100.0)	99.6 (99.5)
R-merge	12.1 (45.6)	0.118 (0.506)
I/sigma	8.99 (3.19)	8.95 (3.30)
SAD Phasing (SHARP)		
Resolution range	47.4–2.8	
Number heavy atoms	12 Pt	
Phasing power (acentric)	1.14	
R _{culis} (acentric)	0.841	
FOM (centric/acentric)	0.156/0.298	
Refinement (REFMAC)		
Resolution	47.6–2.0 (2.05–2.00)	
Reflections used	111098 (7938)	
Reflections R _{free}	5865 (402)	
R-factor	0.177 (0.228)	
R-free	0.214 (0.281)	
No. of protein heavy atoms	9245 (2 trimers)	
No. water molecules	1178	
sulphate, glycerol	10, 2	
Average B value	15.8 Å ²	
RMS deviation bond/angles	0.015/1.487	
Ramachandran plot (Molprobit)		
Favored regions	97.8%	
Allowed regions	100%	

Therefore it has been suggested that the alignment of protein structures should tell us more about phylogenetic relationships between viruses than just amino acid sequences (Bamford et al., 2005). We made phylogenetic trees of adenovirus subtypes based both on sequence alignment of penton base and fibre head sequences and a tree based on the structures of the fibre heads. For these alignments we have only used the subtypes for which the head structures were known.

Table 1 shows the amount of sequence identity and the root mean square deviations (RMSD) of the structures for all known head structures. As could be expected, the higher the sequence identity between a pair of heads, the lower the RMSD. The most deviating heads are the two CELO virus heads, the short fibre head being most different from the human virus fibre heads. The canine virus fibre head is much closer to the human than to the avian virus heads.

Fig. 5 shows phylogenetic trees based on alignments of the penton base sequences (Fig. 5B) and on the fibre heads (Fig. 5A) of all viruses for which the head structures are known. The tree based on the hexon sequences is very similar to the penton base tree (data not shown). The main difference between the "head" and the "penton base" trees are that group C viruses are closest to the animal viruses in the penton base tree whereas HAdV41 (group F virus) is closest to the animal viruses in the head tree. In fact, the HAdV41 short fibre head sequence locates closer to the CELO head sequence than the canine virus sequence. Apart from the length of the branches, the tree based on the structures of the fibre heads (Fig. 5C) is the same as that based on the head sequences. The two CELO fibre heads have clearly evolved from a common ancestor. If the evolution of mammalian and avian virus sequences have proceeded at similar rates, the duplication of the avian virus fibre sequence must have happened before the human and dog virus lineages separated. In the absence of more mammalian virus head structures it is not possible to say whether all mammalian virus head structure cluster close together or that a transmission occurred from humans to dogs, as is the case for the exclusive human virus that is present in a high percentage of pet dogs (Chiou et al., 2005). It is also possible that a recombination occurred in which a canine virus obtained a human virus fibre or fibre head.

Materials and methods

Protein production

A DNA-fragment encoding residues 201 to 410 of the avian adenovirus short fibre (accession number Q64762) was produced by PCR using primers 5'-CTG ATG GGA TCC CTG TAC CAA GCG CCC ACT AG -3' (forward primer) and 5'-CTG ATG CTC GAG TCA GAC CGT AAC GGG GGC G-3' (reverse primer). The PCR product was then cloned between the XhoI and BamHI restriction sites of the expression vector pProEx HTB (Life Technologies). The resultant plasmid encodes the short fibre fragment fused to an N-terminal purification six-histidine tag.

Protein was expressed over-night at 20 °C in transformed *E. coli* B121 star cells (Invitrogen) in Luria Broth medium supplemented with 100 µg/ml ampicilline and 34 µg/ml chloramphenicol. Expression was induced at OD₆₀₀ 0.6 by adding 0.5 mM isopropyl β-D-1-thiogalactopyranoside (IPTG). The cells were harvested by centrifugation (7 min at 7000 × g), resuspended in lysis buffer (20 mM Tris –HCl pH 7.9, 500 mM NaCl, 10 mM imidazole, complete EDTA-free protease inhibitor cocktail (Roche) and 0.1 mM PMSF and lysed by sonication. The cell lysate was centrifuged for 45 min at 35,000 × g and the supernatant containing soluble protein was incubated with cobalt-agarose resin (Sigma) in lysis buffer at 4 °C for 1 h. The resin was washed with the same lysis buffer and then with wash buffer (20 mM Tris –HCl pH 7.9, 500 mM NaCl and 20 mM imidazole) to remove non-specifically bound proteins. The His-tagged short-fibre head protein was then eluted with elution buffer (20 mM Tris –HCl pH 7.9, 200 mM NaCl and 250 mM imidazole). The eluted protein was concentrated and purified by gel filtration on a Superdex 200 column (Amersham Pharmacia Biotech), changing the elution buffer to crystallisation buffer (20 mM Tris –HCl pH 7.9, 200 mM NaCl). Protein homogeneity was checked by SDS-PAGE, negative stain electron microscopy and dynamic light scattering (DLS) (not shown). The protein was concentrated up to 30 mg/ml using a Centricon concentrator (Milipore) and used with the His-tag for crystallization trials.

Crystallogenesis and structure determination

Initial conditions for crystallization of the short-fibre head were identified using a sitting-drop vapour-diffusion screen of commercial sparse-matrix crystallization conditions. The crystallization trials were set up with a Cartesian PixSys 4200 crystallization robot (Genomic Solutions, U.K.) using Greiner Crystal Quick plates (flat bottom, untreated). After optimisation, the best crystals were found to grow at 4 °C in 0.1 M lithium sulphate, 0.1 M tri-sodium citrate at pH 5.6 and 12% w/v polyethylene glycol (PEG) 4000. The crystals were soaked for 5–10 s in a cryoprotectant solution of crystallization buffer plus 25% glycerol. The soaked crystals were mounted on cryo-loops (Hampton), flash frozen at 100 K and placed in a stream of nitrogen at 100 K for all diffraction experiments.

Data were collected on European Synchrotron Radiation Facility (ESRF), beam line ID14-EH4 with an ADSC Q315r CCD detector. The crystals contained two trimers in the asymmetric unit (space group P4₁2₁2), with unit-cell parameters a=235.75 Å, b=235.75 Å, c=61.75 Å and α=β=γ=90°. The best native dataset extended to a resolution of 2.0 Å. All data were integrated and scaled with XDS (Kabsch, 1988).

As described by Guardado-Calvo et al. (Guardado-Calvo et al., 2006) for the CELO virus long-fibre head, molecular replacement phasing using human adenovirus fibre heads was not successful, presumably due to very low sequence homology and comparably low structural homology. Therefore, crystals were soaked overnight in 6 mM PIP (Di-μ-iodobis(ethylenediamine) diplatinum) and the best derivative crystal datasets extended to 2.5 Å resolution. SHELXD (Sheldrick, 2008) found 12 Pt sites in the asymmetric unit. Inspection showed that these corresponded to two trimers in the asymmetric unit and hence a solvent content of 65%. SHARP (De La Fortelle and Bricogne, 1997) was used to refine the sites and to calculate the initial phases. The map was improved with RESOLVE (Terwilliger, 2002) using initially 2-fold non-crystallographic (NCS) averaging (i.e. the two trimers) and subsequently 6-fold NCS (the six monomers). This allowed the construction of a first polyaniline model for ~80% of the protomer in the asymmetric unit. This model was used for molecular replacement with the 2 Å resolution native dataset with PhaserMR (McCoy et al., 2005). ARP/wARP (Perrakis et al., 1999) was then able to build 1120 residues out of 1422 residues in the asymmetric unit. Model visualization and subsequent building was done with Coot (Emsley and Cowtan, 2004) and refinement with REFMAC (Murshudov et al., 1997). Water molecules were added with ARP/wARP. 10 prominent electron density peaks were assigned to sulphate ions, a component present in the

crystallization solution. The final model contains residues 206–410 of each monomer. Density for the five N-terminal amino acids 201–205 was missing from the map. This region is probably flexible or unfolded when the fibre shaft is not complete. The geometry of the final model was checked with MolProbity (Davis et al., 2007) and 98% of all residues were in favoured regions and the remaining 2% were in allowed regions of the Ramachandran plot. The data-collection and refinement statistics are given in Table 2. The coordinates have been deposited in the protein structure database (<http://www.rcsb.org>) under the accession number 2VTW. The structure factors, including those for the derivative data have also been deposited.

Acknowledgments

We acknowledge the use of the facilities provided through the "Partnership for Structural Biology" such as the crystallization robot; the ESRF for use of beam-line ID14-EH4, Raimond Ravelli for help with data collection and initial data treatment and Hassan Belrhali for help with data collection. GS was supported in part by a Jeunes Chercheurs grant from the Agence Nationale pour la Recherche. MEB is a PhD student financed by the region Rhone Alpes. ES was supported by an qE-STAR q fellowship funded by the EC's FP6 Marie Curie Host fellowship for Early Stage Research Training under contract number MEST-CT-2004-504640. We thank Patrick Langlois, AFSSA, Ploufragan, France for providing the DNA clone with the CELO fibre head sequence.

References

- Arnberg, N., Edlund, K., Kidd, A.H., Wadell, G., 2000. Adenovirus type 37 uses sialic acid as a cellular receptor. *J Virol* 74 (1), 42–48.
- Bamford, D.H., Grimes, J.M., Stuart, D.I., 2005. What does structure tell us about virus evolution? *Curr. Opin. Struct. Biol.* 15 (6), 655–663.
- Bergelson, J.M., Cunningham, J.A., Droguett, G., Kurt-Jones, E.A., Krithivas, A., Hong, J.S., Horwitz, M.S., Crowell, R.L., Finberg, R.W., 1997. Isolation of a common receptor for Coxsackie B viruses and adenoviruses 2 and 5. *Science* 275 (5304), 1320–1323.
- Bewley, M.C., Springer, K., Zhang, Y.B., Freimuth, P., Flanagan, J.M., 1999. Structural analysis of the mechanism of adenovirus binding to its human cellular receptor. *CAR. Science* 286 (5444), 1579–1583.
- Bond, C.S., 2003. TopDraw: a sketchpad for protein structure topology cartoons. *Bioinformatics* 19 (2), 311–312.
- Burmeister, W.P., Gulligay, D., Cusack, S., Wadell, G., Arnberg, N., 2004. Crystal structure of species D adenovirus fiber knobs and their sialic acid binding sites. *J Virol* 78 (14), 7727–7736.
- Chiocca, S., Kurzbauer, R., Schaffner, G., Baker, A., Mautner, V., Cotten, M., 1996. The complete DNA sequence and genomic organization of the avian adenovirus CELO. *J. Virol.* 70 (5), 2939–2949.
- Chiou, S.H., Chow, K.C., Yang, C.H., Chiang, S.F., Lin, C.H., 2005. Discovery of Epstein-Barr virus (EBV)-encoded RNA signal and EBV nuclear antigen leader protein DNA sequence in pet dogs. *J. Gen. Virol.* 86 (Pt 4), 899–905.
- Cowen, B., Calnek, B.W., Menendez, N.A., Ball, R.F., 1978. Avian adenoviruses: effect on egg production, shell quality, and feed consumption. *Avian Dis.* 22 (3), 459–470.
- Davis, I.W., Leaver-Fay, A., Chen, V.B., Block, J.N., Kapral, G.J., Wang, X., Murray, L.W., Arendall III, W.B., Snoeyink, J., Richardson, J.S., Richardson, D.C., 2007. MolProbity: all-atom contacts and structure validation for proteins and nucleic acids. *Nucleic Acids Res.* 35, W375–W383 (Web Server issue).
- Davison, A.J., Benko, M., Harrach, B., 2003. Genetic content and evolution of adenoviruses. *J. Gen. Virol.* 84 (Pt 11), 2895–2908.
- De La Fortelle, E., Bricogne, G., 1997. Maximum-likelihood heavy-atom parameter refinement for multiple isomorphous replacement and multiwavelength anomalous diffraction methods. *Methods Enzymol.* 276, 472–494.
- DeLano, W.L., 2002. The PyMOL Molecular Graphics System. DeLano Scientific, San Carlos, CA.
- Dubose, R.J., Grumbles, L.C., 1959. The relationship between quail bronchitis virus and chicken embryo lethal orphan virus. *Avian Dis.* 3, 321–344.
- Durmort, C., Stehlin, C., Schoehn, G., Mittraki, A., Drouet, E., Cusack, S., Burmeister, W.P., 2001. Structure of the fiber head of Ad3, a non-CAR-binding serotype of adenovirus. *Virology* 285 (2), 302–312.
- Emsley, P., Cowtan, K., 2004. Coot: model-building tools for molecular graphics. *Acta Crystallogr. D. Biol. Crystallogr.* 60 (Pt 12 Pt 1), 2126–2132.
- Fabry, C.M., Rosa-Calatrava, M., Conway, J.F., Zubieta, C., Cusack, S., Ruigrok, R.W., Schoehn, G., 2005. A quasi-atomic model of human adenovirus type 5 capsid. *EMBO J.* 24 (9), 1645–1654.
- Felsenstein, J., 1989. Mathematics vs. evolution: mathematical evolutionary theory. *Science* 246 (4932), 941–942.
- Felsenstein, J., 1997. An alternating least squares approach to inferring phylogenies from pairwise distances. *Syst Biol.* 46 (1), 101–111.

- Fleischli, C., Sirena, D., Lesage, G., Havenga, M.J., Cattaneo, R., Greber, U.F., Hemmi, S., 2007. Species B adenovirus serotypes 3, 7, 11 and 35 share similar binding sites on the membrane cofactor protein CD46 receptor. *J. Gen. Virol.* 88 (Pt 11), 2925–2934.
- Fuschiotti, P., Schoehn, G., Fender, P., Fabry, C.M., Hewat, E.A., Chroboczek, J., Ruigrok, R.W., Conway, J.F., 2006. Structure of the dodecahedral penton particle from human adenovirus type 3. *J. Mol. Biol.* 356 (2), 510–520.
- Gaggar, A., Shayakhmetov, D.M., Lieber, A., 2003. CD46 is a cellular receptor for group B adenoviruses. *Nat. Med.* 9 (11), 1142–1408.
- Galtier, N., Gouy, M., Gautier, C., 1996. SEAVIEW and PHYLO_WIN: two graphic tools for sequence alignment and molecular phylogeny. *Comput. Appl. Biosci.* 12 (6), 543–548.
- Gelderblom, H., Maichle-Lauppe, I., 1982. The fibers of owl adenoviruses. *Arch. Virol.* 72 (4), 289–298.
- Gouet, P., Courcelle, E., Stuart, D.I., Metz, F., 1999. ESPript: analysis of multiple sequence alignments in PostScript. *Bioinformatics* 15 (4), 305–308.
- Guardado-Calvo, P., Llamas-Saiz, A.L., Langlois, P., van Raaij, M.J., 2006. Crystallization of the C-terminal head domain of the avian adenovirus CELO long fibre. *Acta Crystallogr., Sect. F Struct. Biol. Cryst. Commun.* 62 (Pt 5), 449–452.
- Guardado-Calvo, P., Llamas-Saiz, A.L., Fox, G.C., Langlois, P., van Raaij, M.J., 2007. Structure of the C-terminal head domain of the fowl adenovirus type 1 long fibre. *J. Gen. Virol.* 88 (Pt 9), 2407–2416.
- Hess, M., Cuzange, A., Ruigrok, R.W., Chroboczek, J., Jacrot, B., 1995. The avian adenovirus penton: two fibres and one base. *J. Mol. Biol.* 252 (4), 379–385.
- Holst, M., Kozack, R.E., Saied, F., Subramaniam, S., 1994. Protein electrostatics: rapid multigrid-based Newton algorithm for solution of the full nonlinear Poisson-Boltzmann equation. *J. Biomol. Struct. Dyn.* 11 (6), 1437–1445.
- Horwitz, M., 1996. Adenoviruses. In: Fields, B., Knipe, D. (Eds.), *Fields Virology*, vol. 2. Raven Press, Philadelphia, PA, pp. 2149–2171. 2 vols.
- Kabsch, W., 1988. Evaluation of single-crystal X-ray diffraction data from a position-sensitive detector. *J. Appl. Crystallogr.* 21, 916–924.
- Krissinel, E., Henrick, K., 2004. Secondary-structure matching (SSM), a new tool for fast protein structure alignment in three dimensions. *Acta Crystallogr., D Biol. Crystallogr.* 60 (Pt 12 Pt 1), 2256–2268.
- Laver, W.G., Younghusband, H.B., Wrigley, N.G., 1971. Purification and properties of chick embryo lethal orphan virus (an avian adenovirus). *Virology* 45 (3), 598–614.
- Lu, G., 2000. TOP: a new method for protein structure comparisons and similarity searches. *J. Appl. Crystallogr.* 33, 176–183.
- Madisch, I., Hofmayer, S., Moritz, C., Grintzalis, A., Hainmueller, J., Pring-Akerblom, P., Heim, A., 2007. Phylogenetic analysis and structural predictions of human adenovirus penton proteins as a basis for tissue-specific adenovirus vector design. *J. Virol.* 81 (15), 8270–8281.
- McCoy, A.J., Grosse-Kunstleve, R.W., Storoni, L.C., Read, R.J., 2005. Likelihood-enhanced fast translation functions. *Acta Crystallogr., D Biol. Crystallogr.* 61 (Pt 4), 458–464.
- Merkel, M.C., Huisken, J.T., Bamford, D.H., Goldman, A., Tuma, R., 2005. The structure of the bacteriophage PRD1 spike sheds light on the evolution of viral capsid architecture. *Mol. Cell* 18 (2), 161–170.
- Murshudov, G.N., Vagin, A.A., Dodson, E.J., 1997. Refinement of macromolecular structures by the maximum-likelihood method. *Acta Crystallogr., D Biol. Crystallogr.* 53 (Pt 3), 240–255.
- Payet, V., Arnauld, C., Picault, J.P., Jestin, A., Langlois, P., 1998. Transcriptional organization of the avian adenovirus CELO. *J. Virol.* 72 (11), 9278–9285.
- Perrakis, A., Morris, R., Lamzin, V.S., 1999. Automated protein model building combined with iterative structure refinement. *Nat. Struct. Biol.* 6 (5), 458–463.
- Persson, B.D., Reiter, D.M., Marttila, M., Mei, Y.F., Casasnovas, J.M., Arnberg, N., Stehle, T., 2007. Adenovirus type 11 binding alters the conformation of its receptor CD46. *Nat. Struct. Mol. Biol.* 14 (2), 164–166.
- Rekosh, D.M., Russell, W.C., Bellet, A.J., Robinson, A.J., 1977. Identification of a protein linked to the ends of adenovirus DNA. *Cell* 11 (2), 283–295.
- Roelvink, P.W., Lizonova, A., Lee, J.G., Li, Y., Bergelson, J.M., Finberg, R.W., Brough, D.E., Koveshi, L., Wickham, T.J., 1998. The coxsackievirus-adenovirus receptor protein can function as a cellular attachment protein for adenovirus serotypes from subgroups A, C, D, E, and F. *J. Virol.* 72 (10), 7909–7915.
- Segerman, A., Atkinson, J.P., Marttila, M., Dennerquist, V., Wadell, G., Arnberg, N., 2003. Adenovirus type 11 uses CD46 as a cellular receptor. *J. Virol.* 77 (17), 9183–9191.
- Seiradake, E., Cusack, S., 2005. Crystal structure of enteric adenovirus serotype 41 short fiber head. *J. Virol.* 79 (22), 14088–14094.
- Seiradake, E., Lortat-Jacob, H., Billet, O., Kremer, E.J., Cusack, S., 2006. Structural and mutational analysis of human Ad37 and canine adenovirus 2 fiber heads in complex with the D1 domain of coxsackie and adenovirus receptor. *J. Biol. Chem.* 281 (44), 33704–33716.
- Sheldrick, G.M., 2008. A short history of SHELX. *Acta Crystallogr., A* 64 (Pt 1), 112–122.
- Sirena, D., Lilienfeld, B., Eisenhut, M., Kalin, S., Boucke, K., Beerli, R.R., Vogt, L., Ruedl, C., Bachmann, M.F., Greber, U.F., Hemmi, S., 2004. The human membrane cofactor CD46 is a receptor for species B adenovirus serotype 3. *J. Virol.* 78 (9), 4454–4462.
- Stewart, P.L., Burnett, R.M., Cyrklaff, M., Fuller, S.D., 1991. Image reconstruction reveals the complex molecular organization of adenovirus. *Cell* 67 (1), 145–154.
- Stewart, P.L., Fuller, S.D., Burnett, R.M., 1993. Difference imaging of adenovirus: bridging the resolution gap between X-ray crystallography and electron microscopy. *EMBO J.* 12 (7), 2589–2599.
- Tan, P.K., Michou, A.I., Bergelson, J.M., Cotten, M., 2001. Defining CAR as a cellular receptor for the avian adenovirus CELO using a genetic analysis of the two viral fibre proteins. *J. Gen. Virol.* 82 (Pt 6), 1465–1472.
- Terwilliger, T.C., 2002. Statistical density modification with non-crystallographic symmetry. *Acta Crystallogr., D Biol. Crystallogr.* 58 (Pt 12), 2082–2086.
- Thompson, J.D., Higgins, D.G., Gibson, T.J., 1994. CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Res.* 22 (22), 4673–4680.
- Tomko, R.P., Xu, R., Phillipson, L., 1997. HCAR and MCAR: the human and mouse cellular receptors for subgroup C adenoviruses and group B coxsackieviruses. *Proc. Natl. Acad. Sci. U. S. A.* 94 (7), 3352–3356.
- van Raaij, M.J., Louis, N., Chroboczek, J., Cusack, S., 1999. Structure of the human adenovirus serotype 2 fiber head domain at 1.5 Å resolution. *Virology* 262 (2), 333–343.
- Wu, E., Fernandez, J., Fleck, S.K., Von Seggern, D.J., Huang, S., Nemerow, G.R., 2001. *Virology* 279, 78–89.
- Xia, D., Henry, L.J., Gerard, R.D., Deisenhofer, J., 1994. Crystal structure of the receptor-binding domain of adenovirus type 5 fiber protein at 1.7 Å resolution. *Structure* 2 (12), 1259–1270.
- Xu, L., Benson, S.D., Burnett, R.M., 2007. Nanoporous crystals of chicken embryo lethal orphan (CELO) adenovirus major coat protein, hexon. *J. Struct. Biol.* 157 (2), 424–431.
- Yates, V.J., Fry, D.E., 1957. Observations on a chicken embryo lethal orphan (CELO) virus. *Am. J. Vet. Res.* 18 (68), 657–660.
- Zubieta, C., Schoehn, G., Chroboczek, J., Cusack, S., 2005. The structure of the human adenovirus 2 penton. *Mol. Cell* 17 (1), 121–135.

6.3. Conclusion

L'adénovirus aviaire de sérotype 1 (CELO) est très différent des autres Ads, il comporte deux fibres liées à la même base de penton: une courte et une longue. Les récepteurs ainsi que la voie d'internalisation du virus ne sont pas encore élucidés. Au niveau de la séquence, les deux fibres et particulièrement la courte possèdent peu de similarité avec les fibres d'HAds. La structure cristallographique de la tête fibre courte du CELO a permis de mieux comprendre pourquoi cette protéine est incapable d'interagir avec le récepteur majeur des Ads (le CAR). En effet, la boucle impliquée dans cette interaction (boucle AB) est différente en longueur et en structure des autres boucles AB capables d'interagir avec le CAR. La surface électrostatique de la tête de fibre courte présente également une région fortement basique comparable à celle des têtes des fibres qui utilisent l'acide sialique comme récepteur. Cela suggère que cette protéine reconnaît ce même récepteur. Finalement la comparaison de la structure de la tête de fibre courte du CELO avec les structures de douze autres têtes de fibre de différents sérotypes d'Ads a permis de conclure qu'au cours de l'évolution, malgré une divergence de la séquence de la fibre, la structure a été globalement conservée. Bien que la tête de la fibre soit soumise à des pressions antigéniques intenses et accumule des mutations au cours de l'évolution, des arbres phylogénétiques basés sur les séquences ou les structures des têtes de fibres sont comparables à ceux obtenus pour la base de penton qui est elle hautement conservée. Ces arbres phylogénétiques ont abouti à trois conclusions majeures:

- Les deux fibres du CELO ont clairement un ancêtre commun.
- La tête de fibre courte d'HAd41 apparaît proche des têtes de fibre du CELO.
- La tête de fibre de l'adénovirus canin est plus proche des têtes de fibre des adénovirus humains que de celles du CELO.

La région fortement basique dans la partie apicale de la tête courte de fibre nous a amené à faire des tests d'interaction avec l'acide sialique par Biacore en collaboration avec le groupe du Dr. Hugues Lortat Jacob de l'IBS.

La glycophorine (GP) est une glycoprotéine riche en acide sialique, abondante à la surface des globules rouges. Pour mesurer l'interaction entre la tête de la fibre courte et l'acide sialique, la GP biotinillée a été immobilisée sur une surface de Biacore couverte de streptavidine. Différentes concentrations en protéine ont été injectées sur la surface et leur liaison a été mesurée par résonance plasmonique de surface (SPR). Aucune interaction n'a été mise en évidence. Il est donc possible que la tête de la fibre courte du CELO n'interagisse pas avec l'acide sialique. L'absence de l'interaction peut aussi être expliquée par la composition

du dérivé de l'acide sialique sur la glycophorine (Kaludov et al., 2001). Il faut noter l'existence de plus de 40 dérivés d'acide sialique différents entre les espèces animales et même entre différents tissus, il est donc possible que la fibre courte n'interagisse qu'avec un acide sialique présent à la surface des cellules d'oiseau mais pas avec l'acide sialique présent à la surface des globules rouges humain (généralement N-acétylneuraminique).

Conclusion Générale

L'étude des virus en général et des adénovirus en particulier, permet de mieux comprendre ces virus afin de mieux les contrôler et le cas échéant mieux les utiliser. L'adénovirus est largement étudié d'un point de vue fondamental et pour des applications en thérapie génique. Actuellement, l'utilisation de ces virus dans des essais de thérapie génique a encore accru l'intérêt que nous portons à leur structure. De nombreux efforts sont consacrés au ciblage spécifique des virus vers certains tissus cellulaires. Ceci est essentiel pour une application efficace et une diminution de la réponse immunitaire chez le patient (utilisation de petites doses). Il est assez aisé de détruire le tropisme naturel des adénovirus (en mutant la tête de la fibre), par contre le re-ciblage du virus est plus complexe. Une des méthodes les plus encourageantes consiste à greffer un peptide de ciblage à la surface du virus. Pour une greffe optimale sans gêner l'intégrité de la particule virale, la connaissance de la structure de la capsid à l'échelle atomique est indispensable. Une particule d'adénovirus humain comporte plus de 2700 polypeptides, avec une masse moléculaire d'environ 150×10^6 Da. La détermination de sa structure atomique reste un véritable challenge. Pour remédier à ce problème, nous étudions en parallèle la structure 3D de virus entier par MET et la structure des protéines de la capsid par cristallographie. La combinaison des données cristallographiques avec les données de microscopie électronique est pour l'instant la seule méthode nous permettant d'obtenir une information à l'échelle quasi-atomique pour le virus entier. Cette approche a déjà été utilisée au laboratoire et elle a permis l'obtention d'un modèle de l'HA5 à résolution quasi atomique (Fabry et al., 2005). Néanmoins, ce modèle ne comporte que les protéines majeures. Afin de compléter ce modèle je me suis intéressée à la structure atomique des protéines mineures de l'Ad.

Les essais de cristallisation des protéines mineures IX et IIIa n'ont malheureusement pas été fructueux. Ceci est probablement dû à une forte flexibilité et instabilité de ces protéines en dehors de la capsid. La protéine IX est la protéine mineure la plus étudiée de l'adénovirus. Récemment une structure 3D de l'adénovirus humain, à 6 Å de résolution par MET, a permis une attribution détaillée de la position de la protéine IX à la surface de la capsid. Cette attribution a été basée sur la densité électronique observée et sur les prédictions de structure secondaire de la protéine. Nous avons pu confirmer cette position définitivement avec des données d'immunodétection (localisation du domaine C-terminal de la protéine IX avec des Fabs spécifiques) et biochimiques (expériences d'oligomérisation). Ainsi nous avons

établi un modèle d'organisation de la protéine IX en réseau à la surface de la capsidie qui lui permet d'exercer sa fonction de ciment de la capsidie.

Le virus CELO a été proposé comme vecteur de transfert en thérapie génique (Kelleher and Vos, 1994; Michou et al., 1999). Celui-ci est en effet capable de contenir jusqu'à 43 kpb (théoriquement) de matériel génétique avec un risque limité de provoquer des réactions immunitaires importantes chez les patients par rapport aux vecteurs humains (à cause de l'immunité préexistante). Des connaissances fondamentales manquent cependant pour son utilisation rationnelle. Le virus CELO contient deux fibres liées à la même base de penton ce qui le différencie des autres adénovirus (Hess et al., 1995; Laver et al., 1971), sa voie d'internalisation n'est encore pas connue. La structure cristallographique de la tête de fibre courte que nous avons déterminée n'apporte pas de réponse par rapport à son rôle dans l'internalisation des particules virales. Néanmoins, la structure de la tête de fibre courte à l'échelle atomique permet la réalisation de mutations stratégiques pour identifier la nature de son récepteur ainsi que son site de liaison. Dans le cadre d'une application en thérapie génique, la structure servira à déterminer les sites de modification potentiels pour le reciblage du virus. La modification de la région fortement basique, par exemple (site potentiel de fixation d'un récepteur), permettra très probablement la modulation du tropisme naturel de la fibre courte pour un reciblage du virus.

Durant cette thèse je me suis aussi intéressée aux sous-particules virales de l'HA3 appelées dodécaèdres qui sont formés par un arrangement sphérique de 12 bases de penton. Ces particules non pathogènes et qui utilisent les HS comme récepteurs à la surface de la cellule ont été proposées comme vecteur de thérapie génique (Fender et al., 2008). Aujourd'hui nous disposons de la condition de cristallisation optimale pour une diffraction des cristaux du dodécaèdre seul et nous envisageons de résoudre la structure du dodécaèdre en complexe avec les HS. Ceci nous permettra d'élucider ce type d'interaction d'un point de vue fondamental et de mieux utiliser ces particules dans un but thérapeutique.

Références Bibliographiques

Abrescia, N. G., Cockburn, J. J., Grimes, J. M., Sutton, G. C., Diprose, J. M., Butcher, S. J., Fuller, S. D., San Martin, C., Burnett, R. M., Stuart, D. I., et al. (2004). Insights into assembly from structural analysis of bacteriophage PRD1. *Nature* *432*, 68-74.

Akalu, A., Liebermann, H., Bauer, U., Granzow, H., and Seidel, W. (1999). The subgenus-specific C-terminal region of protein IX is located on the surface of the adenovirus capsid. *J Virol* *73*, 6182-6187.

Albinsson, B., and Kidd, A. H. (1999). Adenovirus type 41 lacks an RGD alpha(v)-integrin binding motif on the penton base and undergoes delayed uptake in A549 cells. *Virus Res* *64*, 125-136.

Alexander, D. A., and Dimock, K. (2002). Sialic acid functions in enterovirus 70 binding and infection. *J Virol* *76*, 11265-11272.

Allard, A., and Wadell, G. (1988). Physical organization of the enteric adenovirus type 41 early region 1A. *Virology* *164*, 220-229.

Amalfitano, A., Hauser, M. A., Hu, H., Serra, D., Begy, C. R., and Chamberlain, J. S. (1998). Production and characterization of improved adenovirus vectors with the E1, E2b, and E3 genes deleted. *J Virol* *72*, 926-933.

Anderson, C. W. (1990). The proteinase polypeptide of adenovirus serotype 2 virions. *Virology* *177*, 259-272.

Anderson, C. W., Baum, P. R., and Gesteland, R. F. (1973). Processing of adenovirus 2-induced proteins. *J Virol* *12*, 241-252.

Aoki, K., Kato, M., Ohtsuka, H., Ishii, K., Nakazono, N., and Sawada, H. (1982). Clinical and aetiological study of adenoviral conjunctivitis, with special reference to adenovirus types 4 and 19 infections. *Br J Ophthalmol* *66*, 776-780.

Arnberg, N., Kidd, A. H., Edlund, K., Nilsson, J., Pring-Akerblom, P., and Wadell, G. (2002). Adenovirus type 37 binds to cell surface sialic acid through a charge-dependent interaction. *Virology* *302*, 33-43.

Arnberg, N., Mei, Y., and Wadell, G. (1997). Fiber genes of adenoviruses with tropism for the eye and the genital tract. *Virology* *227*, 239-244.

Athappilly, F. K., Murali, R., Rux, J. J., Cai, Z., and Burnett, R. M. (1994). The refined crystal structure of hexon, the major coat protein of adenovirus type 2, at 2.9 Å resolution. *J Mol Biol* *242*, 430-455.

Bai, M., Harfe, B., and Freimuth, P. (1993). Mutations that alter an Arg-Gly-Asp (RGD) sequence in the adenovirus type 2 penton base protein abolish its cell-rounding activity and delay virus reproduction in flat cells. *J Virol* *67*, 5198-5205.

Balakireva, L., Schoehn, G., Thouvenin, E., and Chroboczek, J. (2003). Binding of adenovirus capsid to dipalmitoyl phosphatidylcholine provides a novel pathway for virus entry. *J Virol* *77*, 4858-4866.

- Bamford, D. H., Burnett, R. M., and Stuart, D. I.** (2002). Evolution of viral structure. *Theor Popul Biol* 61, 461-470.
- Barton, E. S., Connolly, J. L., Forrest, J. C., Chappell, J. D., and Dermody, T. S.** (2001a). Utilization of sialic acid as a coreceptor enhances reovirus attachment by multistep adhesion strengthening. *J Biol Chem* 276, 2200-2211.
- Barton, E. S., Forrest, J. C., Connolly, J. L., Chappell, J. D., Liu, Y., Schnell, F. J., Nusrat, A., Parkos, C. A., and Dermody, T. S.** (2001b). Junction adhesion molecule is a receptor for reovirus. *Cell* 104, 441-451.
- Bauer, P. H., Cui, C., Liu, W. R., Stehle, T., Harrison, S. C., DeCaprio, J. A., and Benjamin, T. L.** (1999). Discrimination between sialic acid-containing receptors and pseudoreceptors regulates polyomavirus spread in the mouse. *J Virol* 73, 5826-5832.
- Benko, M., Elo, P., Ursu, K., Ahne, W., LaPatra, S. E., Thomson, D., and Harrach, B.** (2002). First molecular evidence for the existence of distinct fish and snake adenoviruses. *J Virol* 76, 10056-10059.
- Benko, M., and Harrach, B.** (1998). A proposal for a new (third) genus within the family Adenoviridae. *Arch Virol* 143, 829-837.
- Benko, M., and Harrach, B.** (2003). Molecular evolution of adenoviruses. *Curr Top Microbiol Immunol* 272, 3-35.
- Benson, S. D., Bamford, J. K., Bamford, D. H., and Burnett, R. M.** (2002). The X-ray crystal structure of P3, the major coat protein of the lipid-containing bacteriophage PRD1, at 1.65 Å resolution. *Acta Crystallogr D Biol Crystallogr* 58, 39-59.
- Bergelson, J. M., Cunningham, J. A., Droguett, G., Kurt-Jones, E. A., Krithivas, A., Hong, J. S., Horwitz, M. S., Crowell, R. L., and Finberg, R. W.** (1997). Isolation of a common receptor for Coxsackie B viruses and adenoviruses 2 and 5. *Science* 275, 1320-1323.
- Berk, A. J.** (1986). Functions of adenovirus E1A. *Cancer Surv* 5, 367-387.
- Bewley, M. C., Springer, K., Zhang, Y. B., Freimuth, P., and Flanagan, J. M.** (1999). Structural analysis of the mechanism of adenovirus binding to its human cellular receptor, CAR. *Science* 286, 1579-1583.
- Binger, M. H., and Flint, S. J.** (1984). Accumulation of early and intermediate mRNA species during subgroup C adenovirus productive infections. *Virology* 136, 387-403.
- Bischoff, J. R., Kirn, D. H., Williams, A., Heise, C., Horn, S., Muna, M., Ng, L., Nye, J. A., Sampson-Johannes, A., Fattaey, A., and McCormick, F.** (1996). An adenovirus mutant that replicates selectively in p53-deficient human tumor cells. *Science* 274, 373-376.
- Boudin, M. L., D'Halluin, J. C., Cousin, C., and Boulanger, P.** (1980). Human adenovirus type 2 protein IIIa. II. Maturation and encapsidation. *Virology* 101, 144-156.
- Boulanger, P., Lemay, P., Blair, G. E., and Russell, W. C.** (1979). Characterization of adenovirus protein IX. *J Gen Virol* 44, 783-800.
- Brockmann, D., and Esche, H.** (2003). The multifunctional role of E1A in the transcriptional regulation of CREB/CBP-dependent target genes. *Curr Top Microbiol Immunol* 272, 97-129.

- Brunger, A. T., Adams, P. D., Clore, G. M., DeLano, W. L., Gros, P., Grosse-Kunstleve, R. W., Jiang, J. S., Kuszewski, J., Nilges, M., Pannu, N. S., *et al.*** (1998). Crystallography & NMR system: A new software suite for macromolecular structure determination. *Acta Crystallogr D Biol Crystallogr* *54*, 905-921.
- Burmeister, W. P., Guilligay, D., Cusack, S., Wadell, G., and Arnberg, N.** (2004). Crystal structure of species D adenovirus fiber knobs and their sialic acid binding sites. *J Virol* *78*, 7727-7736.
- Burnett, R. M.** (1985). The structure of the adenovirus capsid. II. The packing symmetry of hexon and its implications for viral architecture. *J Mol Biol* *185*, 125-143.
- Burnett, R. M., Grutter, M. G., and White, J. L.** (1985). The structure of the adenovirus capsid. I. An envelope model of hexon at 6 Å resolution. *J Mol Biol* *185*, 105-123.
- Cahan, L. D., and Paulson, J. C.** (1980). Polyoma virus adsorbs to specific sialyloligosaccharide receptors on erythrocytes. *Virology* *103*, 505-509.
- Campos, S. K., and Barry, M. A.** (2007). Current advances and future challenges in Adenoviral vector biology and targeting. *Curr Gene Ther* *7*, 189-204.
- Campos, S. K., Parrott, M. B., and Barry, M. A.** (2004). Avidin-based targeting and purification of a protein IX-modified, metabolically biotinylated adenoviral vector. *Mol Ther* *9*, 942-954.
- Casasnovas, J. M., Larvie, M., and Stehle, T.** (1999). Crystal structure of two CD46 domains reveals an extended measles virus-binding surface. *Embo J* *18*, 2911-2922.
- Chatterjee, P. K., Vayda, M. E., and Flint, S. J.** (1985). Interactions among the three adenovirus core proteins. *J Virol* *55*, 379-386.
- Chatterjee, P. K., Vayda, M. E., and Flint, S. J.** (1986). Identification of proteins and protein domains that contact DNA within adenovirus nucleoprotein cores by ultraviolet light crosslinking of oligonucleotides ³²P-labelled in vivo. *J Mol Biol* *188*, 23-37.
- Chee-Sheung, C. C., and Ginsberg, H. S.** (1982). Characterization of a temperature-sensitive fiber mutant of type 5 adenovirus and effect of the mutation on virion assembly. *J Virol* *42*, 932-950.
- Chiou, S. K., Tseng, C. C., Rao, L., and White, E.** (1994). Functional complementation of the adenovirus E1B 19-kilodalton protein with Bcl-2 in the inhibition of apoptosis in infected cells. *J Virol* *68*, 6553-6566.
- Chroboczek, J., Viard, F., and D'Halluin, J. C.** (1986). Human adenovirus 2 temperature-sensitive mutant 112 contains three mutations in the protein IIIa gene. *Gene* *49*, 157-160.
- Chu, Y., Heistad, D., Cybulsky, M. I., and Davidson, B. L.** (2001). Vascular cell adhesion molecule-1 augments adenovirus-mediated gene transfer. *Arterioscler Thromb Vasc Biol* *21*, 238-242.
- Colby, W. W., and Shenk, T.** (1981). Adenovirus type 5 virions can be assembled in vivo in the absence of detectable polypeptide IX. *J Virol* *39*, 977-980.
- Crawford-Miksza, L., and Schnurr, D. P.** (1996). Analysis of 15 adenovirus hexon proteins reveals the location and structure of seven hypervariable regions containing serotype-specific residues. *J Virol* *70*, 1836-1844.
- Croyle, M. A., Stone, M., Amidon, G. L., and Roessler, B. J.** (1998). In vitro and in vivo assessment of adenovirus 41 as a vector for gene delivery to the intestine. *Gene Ther* *5*, 645-654.

- D'Halluin, J. C., Cousin, C., and Boulanger, P.** (1982). Physical mapping of adenovirus type 2 temperature-sensitive mutations by restriction endonuclease analysis of interserotypic recombinants. *J Virol* *41*, 401-413.
- D'Halluin, J. C., Milleville, M., Martin, G. R., and Boulanger, P.** (1980). Morphogenesis of human adenovirus type 2 studied with fiber- and fiber and penton base-defective temperature-sensitive mutants. *J Virol* *33*, 88-99.
- Davison, A. J., Benko, M., and Harrach, B.** (2003). Genetic content and evolution of adenoviruses. *J Gen Virol* *84*, 2895-2908.
- Davison, E., Diaz, R. M., Hart, I. R., Santis, G., and Marshall, J. F.** (1997). Integrin alpha5beta1-mediated adenovirus infection is enhanced by the integrin-activating antibody TS2/16. *J Virol* *71*, 6204-6207.
- Davison, E., Kirby, I., Elliott, T., and Santis, G.** (1999). The human HLA-A*0201 allele, expressed in hamster cells, is not a high-affinity receptor for adenovirus type 5 fiber. *J Virol* *73*, 4513-4517.
- Davydova, J., Le, L. P., Gavrikova, T., Wang, M., Krasnykh, V., and Yamamoto, M.** (2004). Infectivity-enhanced cyclooxygenase-2-based conditionally replicative adenoviruses for esophageal adenocarcinoma treatment. *Cancer Res* *64*, 4319-4327.
- De Jong, J. C., Wermenbol, A. G., Verweij-Uijterwaal, M. W., Slaterus, K. W., Wertheim-Van Dillen, P., Van Doornum, G. J., Khoo, S. H., and Hierholzer, J. C.** (1999). Adenoviruses from human immunodeficiency virus-infected individuals, including two strains that represent new candidate serotypes Ad50 and Ad51 of species B1 and D, respectively. *J Clin Microbiol* *37*, 3940-3945.
- Dehecchi, M. C., Melotti, P., Bonizzato, A., Santacatterina, M., Chilosi, M., and Cabrini, G.** (2001). Heparan sulfate glycosaminoglycans are receptors sufficient to mediate the initial binding of adenovirus types 2 and 5. *J Virol* *75*, 8772-8780.
- Dehecchi, M. C., Tamanini, A., Bonizzato, A., and Cabrini, G.** (2000). Heparan sulfate glycosaminoglycans are involved in adenovirus type 5 and 2-host cell interactions. *Virology* *268*, 382-390.
- Dekker, J., Kanellopoulos, P. N., Loonstra, A. K., van Oosterhout, J. A., Leonard, K., Tucker, P. A., and van der Vliet, P. C.** (1997). Multimerization of the adenovirus DNA-binding protein is the driving force for ATP-independent DNA unwinding during strand displacement synthesis. *Embo J* *16*, 1455-1463.
- Ding, J., McGrath, W. J., Sweet, R. M., and Mangel, W. F.** (1996). Crystal structure of the human adenovirus proteinase with its 11 amino acid cofactor. *Embo J* *15*, 1778-1783.
- Dmitriev, I., Krasnykh, V., Miller, C. R., Wang, M., Kashentseva, E., Mikheeva, G., Belousova, N., and Curiel, D. T.** (1998). An adenovirus vector with genetically modified fibers demonstrates expanded tropism via utilization of a coxsackievirus and adenovirus receptor-independent cell entry mechanism. *J Virol* *72*, 9706-9713.
- Dmitriev, I. P., Kashentseva, E. A., and Curiel, D. T.** (2002). Engineering of adenovirus vectors containing heterologous peptide sequences in the C terminus of capsid protein IX. *J Virol* *76*, 6893-6899.

- Doucas, V., Ishov, A. M., Romo, A., Juguilon, H., Weitzman, M. D., Evans, R. M., and Maul, G. G.** (1996). Adenovirus replication is coupled with the dynamic properties of the PML nuclear structure. *Genes Dev* *10*, 196-207.
- Durmort, C., Stehlin, C., Schoehn, G., Mitraki, A., Drouet, E., Cusack, S., and Burmeister, W. P.** (2001). Structure of the fiber head of Ad3, a non-CAR-binding serotype of adenovirus. *Virology* *285*, 302-312.
- Élő, P., Farkas, S. L., Dán, A. L., and M., K. G.** (2002). The p32K Structural Protein of the Adenovirus Might Have Bacterial Relatives. *Journal of Molecular Evolution* *56*, 175-180.
- Emsley, P., and Cowtan, K.** (2004). Coot: model-building tools for molecular graphics. *Acta Crystallogr D Biol Crystallogr* *60*, 2126-2132.
- Evans, J. D., and Hearing, P.** (2003). Distinct roles of the Adenovirus E4 ORF3 protein in viral DNA replication and inhibition of genome concatenation. *J Virol* *77*, 5295-5304.
- Everitt, E., Lutter, L., and Philipson, L.** (1975). Structural proteins of adenoviruses. XII. Location and neighbor relationship among proteins of adenovirion type 2 as revealed by enzymatic iodination, immunoprecipitation and chemical cross-linking. *Virology* *67*, 197-208.
- Everitt, E., Sundquist, B., Pettersson, U., and Philipson, L.** (1973). Structural proteins of adenoviruses. X. Isolation and topography of low molecular weight antigens from the virion of adenovirus type 2. *Virology* *52*, 130-147.
- Fabry, C. M., Rosa-Calatrava, M., Conway, J. F., Zubietta, C., Cusack, S., Ruigrok, R. W., and Schoehn, G.** (2005). A quasi-atomic model of human adenovirus type 5 capsid. *Embo J* *24*, 1645-1654.
- Falgout, B., and Ketner, G.** (1988). Characterization of adenovirus particles made by deletion mutants lacking the fiber gene. *J Virol* *62*, 622-625.
- Favier, A. L., Burmeister, W. P., and Chroboczek, J.** (2004). Unique physicochemical properties of human enteric Ad41 responsible for its survival and replication in the gastrointestinal tract. *Virology* *322*, 93-104.
- Favier, A. L., Schoehn, G., Jaquinod, M., Harsi, C., and Chroboczek, J.** (2002). Structural studies of human enteric adenovirus type 41. *Virology* *293*, 75-85.
- Fender, P., Boussaid, A., Mezin, P., and Chroboczek, J.** (2005). Synthesis, cellular localization, and quantification of penton-dodecahedron in serotype 3 adenovirus-infected cells. *Virology* *340*, 167-173.
- Fender, P., Kidd, A. H., Brebant, R., Oberg, M., Drouet, E., and Chroboczek, J.** (1995). Antigenic sites on the receptor-binding domain of human adenovirus type 2 fiber. *Virology* *214*, 110-117.
- Fender, P., Ruigrok, R. W., Gout, E., Buffet, S., and Chroboczek, J.** (1997). Adenovirus dodecahedron, a new vector for human gene transfer. *Nat Biotechnol* *15*, 52-56.
- Fender, P., Schoehn, G., Foucaud-Gamen, J., Gout, E., Garcel, A., Drouet, E., and Chroboczek, J.** (2003). Adenovirus dodecahedron allows large multimeric protein transduction in human cells. *J Virol* *77*, 4960-4964.

- Fender, P., Schoehn, G., Perron-Sierra, F., Tucker, G. C., and Lortat-Jacob, H.** (2008). Adenovirus dodecahedron cell attachment and entry are mediated by heparan sulfate and integrins and vary along the cell cycle. *Virology* 371, 155-164.
- Fisher, K. D., Stallwood, Y., Green, N. K., Ulbrich, K., Mautner, V., and Seymour, L. W.** (2001). Polymer-coated adenovirus permits efficient retargeting and evades neutralising antibodies. *Gene Ther* 8, 341-348.
- Fleischli, C., Sirena, D., Lesage, G., Havenga, M. J., Cattaneo, R., Greber, U. F., and Hemmi, S.** (2007). Species B adenovirus serotypes 3, 7, 11 and 35 share similar binding sites on the membrane cofactor protein CD46 receptor. *J Gen Virol* 88, 2925-2934.
- Fleischli, C., Verhaagh, S., Havenga, M., Sirena, D., Schaffner, W., Cattaneo, R., Greber, U. F., and Hemmi, S.** (2005). The distal short consensus repeats 1 and 2 of the membrane cofactor protein CD46 and their distance from the cell membrane determine productive entry of species B adenovirus serotype 35. *J Virol* 79, 10013-10022.
- Flint, S. J., and Gonzalez, R. A.** (2003). Regulation of mRNA production by the adenoviral E1B 55-kDa and E4 Orf6 proteins. *Curr Top Microbiol Immunol* 272, 287-330.
- Furcinitti, P. S., van Oostrum, J., and Burnett, R. M.** (1989). Adenovirus polypeptide IX revealed as capsid cement by difference images from electron microscopy and crystallography. *Embo J* 8, 3563-3570.
- Fuschiotti, P., Fender, P., Schoehn, G., and Conway, J. F.** (2006a). Development of the dodecahedral penton particle from adenovirus 3 for therapeutic application. *J Gen Virol* 87, 2901-2905.
- Fuschiotti, P., Schoehn, G., Fender, P., Fabry, C. M., Hewat, E. A., Chroboczek, J., Ruigrok, R. W., and Conway, J. F.** (2006b). Structure of the dodecahedral penton particle from human adenovirus type 3. *J Mol Biol* 356, 510-520.
- Gaggar, A., Shayakhmetov, D. M., and Lieber, A.** (2003). CD46 is a cellular receptor for group B adenoviruses. *Nat Med* 9, 1408-1412. Epub 2003 Oct 1419.
- Galinier, R., Gout, E., Lortat-Jacob, H., Wood, J., and Chroboczek, J.** (2002). Adenovirus protein involved in virus internalization recruits ubiquitin-protein ligases. *Biochemistry* 41, 14299-14305.
- Gardlik, R., Palffy, R., Hodosy, J., Lukacs, J., Turna, J., and Celec, P.** (2005). Vectors and delivery systems in gene therapy. *Med Sci Monit* 11, RA110-121.
- Garnett, C. T., Erdman, D., Xu, W., and Gooding, L. R.** (2002). Prevalence and quantitation of species C adenovirus DNA in human mucosal lymphocytes. *J Virol* 76, 10608-10616.
- Geraghty, R. J., Krummenacher, C., Cohen, G. H., Eisenberg, R. J., and Spear, P. G.** (1998). Entry of alphaherpesviruses mediated by poliovirus receptor-related protein 1 and poliovirus receptor. *Science* 280, 1618-1620.
- Ghosh-Choudhury, G., Haj-Ahmad, Y., and Graham, F. L.** (1987). Protein IX, a minor component of the human adenovirus capsid, is essential for the packaging of full length genomes. *Embo J* 6, 1733-1739.

- Giannakis, E., Jokiranta, T. S., Ormsby, R. J., Duthy, T. G., Male, D. A., Christiansen, D., Fischetti, V. A., Bagley, C., Loveland, B. E., and Gordon, D. L.** (2002). Identification of the streptococcal M protein binding site on membrane cofactor protein (CD46). *J Immunol* *168*, 4585-4592.
- Greber, U. F.** (2002). Signalling in viral entry. *Cell Mol Life Sci* *59*, 608-626.
- Greber, U. F., Suomalainen, M., Stidwill, R. P., Boucke, K., Ebersold, M. W., and Helenius, A.** (1997). The role of the nuclear pore complex in adenovirus DNA entry. *Embo J* *16*, 5998-6007.
- Greber, U. F., and Way, M.** (2006). A superhighway to virus infection. *Cell* *124*, 741-754.
- Greber, U. F., Webster, P., Weber, J., and Helenius, A.** (1996). The role of the adenovirus protease on virus entry into cells. *Embo J* *15*, 1766-1777.
- Greber, U. F., Willetts, M., Webster, P., and Helenius, A.** (1993). Stepwise dismantling of adenovirus 2 during entry into cells. *Cell* *75*, 477-486.
- Greenstone, H. L., Santoro, F., Lusso, P., and Berger, E. A.** (2002). Human Herpesvirus 6 and Measles Virus Employ Distinct CD46 Domains for Receptor Function. *J Biol Chem* *277*, 39112-39118.
- Greenwald, R. J., Freeman, G. J., and Sharpe, A. H.** (2005). The B7 family revisited. *Annu Rev Immunol* *23*, 515-548.
- Greve, J. M., Davis, G., Meyer, A. M., Forte, C. P., Yost, S. C., Marlor, C. W., Kamarck, M. E., and McClelland, A.** (1989). The major human rhinovirus receptor is ICAM-1. *Cell* *56*, 839-847.
- Guardado-Calvo, P., Llamas-Saiz, A. L., Fox, G. C., Langlois, P., and van Raaij, M. J.** (2007). Structure of the C-terminal head domain of the fowl adenovirus type 1 long fiber. *J Gen Virol* *88*, 2407-2416.
- Gustafsson, D. J., Segerman, A., Lindman, K., Mei, Y. F., and Wadell, G.** (2006). The Arg279Gln [corrected] substitution in the adenovirus type 11p (Ad11p) fiber knob abolishes EDTA-resistant binding to A549 and CHO-CD46 cells, converting the phenotype to that of Ad7p. *J Virol* *80*, 1897-1905.
- Hannan, C., Raptis, L. H., Dery, C. V., and Weber, J.** (1983). Biological and structural studies with an adenovirus type 2 temperature-sensitive mutant defective for uncoating. *Intervirology* *19*, 213-223.
- Harada, J. N., Shevchenko, A., Shevchenko, A., Pallas, D. C., and Berk, A. J.** (2002). Analysis of the adenovirus E1B-55K-anchored proteome reveals its link to ubiquitination machinery. *J Virol* *76*, 9194-9206.
- Havenga, M. J., Lemckert, A. A., Ophorst, O. J., van Meijer, M., Germeraad, W. T., Grimbergen, J., van Den Doel, M. A., Vogels, R., van Deutekom, J., Janson, A. A., *et al.* (2002). Exploiting the natural diversity in adenovirus tropism for therapy and prevention of disease. *J Virol* *76*, 4612-4620.**
- Hendrickson, W. A., Horton, J. R., and LeMaster, D. M.** (1990). Selenomethionyl proteins produced for analysis by multiwavelength anomalous diffraction (MAD): a vehicle for direct determination of three-dimensional structure. *Embo J* *9*, 1665-1672.
- Hess, M., Cuzange, A., Ruigrok, R. W., Chroboczek, J., and Jacrot, B.** (1995). The avian adenovirus penton: two fibres and one base. *J Mol Biol* *252*, 379-385.

- Hierholzer, J. C.** (1992). Adenoviruses in the immunocompromised host. *Clin Microbiol Rev* 5, 262-274.
- Hierholzer, J. C., Wigand, R., Anderson, L. J., Adrian, T., and Gold, J. W.** (1988). Adenoviruses from patients with AIDS: a plethora of serotypes and a description of five new serotypes of subgenus D (types 43-47). *J Infect Dis* 158, 804-813.
- Hong, S. S., Habib, N. A., Franqueville, L., Jensen, S., and Boulanger, P. A.** (2003). Identification of adenovirus (ad) penton base neutralizing epitopes by use of sera from patients who had received conditionally replicative ad (add11520) for treatment of liver tumors. *J Virol* 77, 10366-10375.
- Hong, S. S., Karayan, L., Tournier, J., Curiel, D. T., and Boulanger, P. A.** (1997). Adenovirus type 5 fiber knob binds to MHC class I alpha2 domain at the surface of human epithelial and B lymphoblastoid cells. *EMBO J* 16, 2294-2306.
- Howitt, J., Bewley, M. C., Graziano, V., Flanagan, J. M., and Freimuth, P.** (2003). Structural basis for variation in adenovirus affinity for the cellular coxsackievirus and adenovirus receptor. *J Biol Chem* 278, 26208-26215.
- Huang, S., Kamata, T., Takada, Y., Ruggeri, Z. M., and Nemerow, G. R.** (1996). Adenovirus interaction with distinct integrins mediates separate events in cell entry and gene delivery to hematopoietic cells. *J Virol* 70, 4502-4508.
- Huang, S., Reddy, V., Dasgupta, N., and Nemerow, G. R.** (1999). A single amino acid in the adenovirus type 37 fiber confers binding to human conjunctival cells. *J Virol* 73, 2798-2802.
- Hurst, H. C.** (1995). Transcription factors 1: bZIP proteins. *Protein Profile* 2, 101-168.
- Jin, Y. M., Pardoe, I. U., Burness, A. T., and Michalak, T. I.** (1994). Identification and characterization of the cell surface 70-kilodalton sialoglycoprotein(s) as a candidate receptor for encephalomyocarditis virus on human nucleated cells. *J Virol* 68, 7308-7319.
- Kabsch, W.** (1988). Evaluation of single-crystal X-ray diffraction data from a position-sensitive detector. *Journal of Applied Crystallography* 21, 916-924.
- Kaludov, N., Brown, K. E., Walters, R. W., Zabner, J., and Chiorini, J. A.** (2001). Adeno-associated virus serotype 4 (AAV4) and AAV5 both require sialic acid binding for hemagglutination and efficient transduction but differ in sialic acid linkage specificity. *J Virol* 75, 6884-6893.
- Kelleher, Z. T., and Vos, J. M.** (1994). Long-term episomal gene delivery in human lymphoid cells using human and avian adenoviral-assisted transfection. *Biotechniques* 17, 1110-1117.
- Kirby, I., Davison, E., Beavil, A. J., Soh, C. P., Wickham, T. J., Roelvink, P. W., Kovesdi, I., Sutton, B. J., and Santis, G.** (2000). Identification of contact residues and definition of the CAR-binding site of adenovirus type 5 fiber protein. *J Virol* 74, 2804-2813.
- Kojaoghlanian, T., Flomenberg, P., and Horwitz, M. S.** (2003). The impact of adenovirus infection on the immunocompromised host. *Rev Med Virol* 13, 155-171.
- Kremer, E. J.** (2005). Gene transfer to the central nervous system: current state of the art of the viral vectors. *Current Genomics* 6, 13-39.
- Krempl, C., Ballesteros, M. L., Zimmer, G., Enjuanes, L., Klenk, H. D., and Herrler, G.** (2000). Characterization of the sialic acid binding activity of transmissible gastroenteritis coronavirus by analysis of haemagglutination-deficient mutants. *J Gen Virol* 81, 489-496.

- Lanciotti, J., Song, A., Doukas, J., Sosnowski, B., Pierce, G., Gregory, R., Wadsworth, S., and O'Riordan, C.** (2003). Targeting adenoviral vectors using heterofunctional polyethylene glycol FGF2 conjugates. *Mol Ther* 8, 99-107.
- Laver, W. G., Youngusband, H. B., and Wrigley, N. G.** (1971). Purification and properties of chick embryo lethal orphan virus (an avian adenovirus). *Virology* 45, 598-614.
- Le, L. P., Everts, M., Dmitriev, I. P., Davydova, J. G., Yamamoto, M., and Curiel, D. T.** (2004). Fluorescently labeled adenovirus with pIX-EGFP for vector detection. *Mol Imaging* 3, 105-116.
- Leen, A. M., and Rooney, C. M.** (2005). Adenovirus as an emerging pathogen in immunocompromised patients. *Br J Haematol* 128, 135-144.
- Lehmborg, E., Traina, J. A., Chakel, J. A., Chang, R. J., Parkman, M., McCaman, M. T., Murakami, P. K., Lahidji, V., Nelson, J. W., Hancock, W. S., et al.** (1999). Reversed-phase high-performance liquid chromatographic assay for the adenovirus type 5 proteome. *J Chromatogr B Biomed Sci Appl* 732, 411-423.
- Leopold, P. L., Ferris, B., Grinberg, I., Worgall, S., Hackett, N. R., and Crystal, R. G.** (1998). Fluorescent virions: dynamic tracking of the pathway of adenoviral gene transfer vectors in living cells. *Hum Gene Ther* 9, 367-378.
- Leslie, A. G. W.** (1992). Recent changes to the MOSFLM package for processing film and image plate data. *Joint CCP4 + ESF-EAMCB Newsletter on Protein Crystallography* 26.
- Li, E., Brown, S. L., Stupack, D. G., Puente, X. S., Cheresch, D. A., and Nemerow, G. R.** (2001). Integrin alpha(v)beta1 is an adenovirus coreceptor. *J Virol* 75, 5405-5409.
- Li, E., Stupack, D., Bokoch, G. M., and Nemerow, G. R.** (1998). Adenovirus endocytosis requires actin cytoskeleton reorganization mediated by Rho family GTPases. *J Virol* 72, 8806-8812.
- Li, Q. G., Hambræus, J., and Wadell, G.** (1991). Genetic relationship between thirteen genome types of adenovirus 11, 34, and 35 with different tropisms. *Intervirology* 32, 338-350.
- Liebermann, H., Lotz, K., Mentel, R., Bauer, U., and Seidel, W.** (2001). Mapping of linear epitopes on fibre knob of human adenovirus serotype 5. *Virus Res* 73, 145-151.
- Liebermann, H., Lotz, K., and Seidel, W.** (2002). Mapping of epitopes on the fiber knobs of human adenovirus serotypes 8 and 15. *Intervirology* 45, 59-66.
- Liu, G. Q., Babiss, L. E., Volkert, F. C., Young, C. S., and Ginsberg, H. S.** (1985). A thermolabile mutant of adenovirus 5 resulting from a substitution mutation in the protein VIII gene. *J Virol* 53, 920-925.
- Lonberg-Holm, K., Crowell, R. L., and Philipson, L.** (1976). Unrelated animal viruses share receptors. *Nature* 259, 679-681.
- Lupas, A.** (1996a). Coiled coils: new structures and new functions. *Trends Biochem Sci* 21, 375-382.
- Lupas, A.** (1996b). Prediction and analysis of coiled-coil structures. *Methods Enzymol* 266, 513-525.
- Lutz, P., Rosa-Calatrava, M., and Keding, C.** (1997). The product of the adenovirus intermediate gene IX is a transcriptional activator. *J Virol* 71, 5102-5109.

- Mangel, W. F., Baniecki, M. L., and McGrath, W. J.** (2003). Specific interactions of the adenovirus proteinase with the viral DNA, an 11-amino-acid viral peptide, and the cellular protein actin. *Cell Mol Life Sci* 60, 2347-2355.
- Mangel, W. F., McGrath, W. J., Toledo, D. L., and Anderson, C. W.** (1993). Viral DNA and a viral peptide can act as cofactors of adenovirus virion proteinase activity. *Nature* 361, 274-275.
- Marttila, M., Persson, D., Gustafsson, D., Liszewski, M. K., Atkinson, J. P., Wadell, G., and Arnberg, N.** (2005). CD46 is a cellular receptor for all species B adenoviruses except types 3 and 7. *J Virol* 79, 14429-14436.
- Mathias, P., Wickham, T., Moore, M., and Nemerow, G.** (1994). Multiple adenovirus serotypes use alpha v integrins for infection. *J Virol* 68, 6811-6814.
- Matthews, B. W.** (1968). Solvent content of protein crystals. *J Mol Biol* 33, 491-497.
- Matthews, D. A., and Russell, W. C.** (1994). Adenovirus protein-protein interactions: hexon and protein VI. *J Gen Virol* 75 (Pt 12), 3365-3374.
- Matthews, D. A., and Russell, W. C.** (1995). Adenovirus protein-protein interactions: molecular parameters governing the binding of protein VI to hexon and the activation of the adenovirus 23K protease. *J Gen Virol* 76 (Pt 8), 1959-1969.
- Matthews, D. A., and Russell, W. C.** (1998). Adenovirus core protein V interacts with p32--a protein which is associated with both the mitochondria and the nucleus. *J Gen Virol* 79 (Pt 7), 1677-1685.
- Mei, Y. F., and Wadell, G.** (1996). Epitopes and hemagglutination binding domain on subgenus B:2 adenovirus fibers. *J Virol* 70, 3688-3697.
- Meier, O., Boucke, K., Hammer, S. V., Keller, S., Stidwill, R. P., Hemmi, S., and Greber, U. F.** (2002). Adenovirus triggers macropinocytosis and endosomal leakage together with its clathrin-mediated uptake. *J Cell Biol* 158, 1119-1131.
- Meier, O., and Greber, U. F.** (2004). Adenovirus endocytosis. *J Gene Med* 6 Suppl 1, S152-163.
- Michou, A. I., Lehrmann, H., Saltik, M., and Cotten, M.** (1999). Mutational analysis of the avian adenovirus CELO, which provides a basis for gene delivery vectors. *J Virol* 73, 1399-1410.
- Mirza, A., and Weber, J.** (1980). Infectivity and uncoating of adenovirus cores. *Intervirology* 13, 307-311.
- Molin, M., Bouakaz, L., Berenjian, S., and Akusjarvi, G.** (2002). Unscheduled expression of capsid protein IIIa results in defects in adenovirus major late mRNA and protein expression. *Virus Res* 83, 197-206.
- Murakami, S., Sakurai, F., Kawabata, K., Okada, N., Fujita, T., Yamamoto, A., Hayakawa, T., and Mizuguchi, H.** (2007). Interaction of penton base Arg-Gly-Asp motifs with integrins is crucial for adenovirus serotype 35 vector transduction in human hematopoietic cells. *Gene Ther* 14, 1525-1533.
- Murshudov, G. N., Vagin, A. A., and Dodson, E. J.** (1997). Refinement of macromolecular structures by the maximum-likelihood method. *Acta Crystallogr D Biol Crystallogr* 53, 240-255.
- Mysiak, M. E., Holthuizen, P. E., and van der Vliet, P. C.** (2004). The adenovirus priming protein pTP contributes to the kinetics of initiation of DNA replication. *Nucleic Acids Res* 32, 3913-3920.

- Nakano, M. Y., Boucke, K., Suomalainen, M., Stidwill, R. P., and Greber, U. F.** (2000). The first step of adenovirus type 2 disassembly occurs at the cell surface, independently of endocytosis and escape to the cytosol. *J Virol* *74*, 7085-7095.
- Naniche, D., Varior-Krishnan, G., Cervoni, F., Wild, T. F., Rossi, B., Rabourdin-Combe, C., and Gerlier, D.** (1993). Human membrane cofactor protein (CD46) acts as a cellular receptor for measles virus. *J Virol* *67*, 6025-6032.
- Neurath, A. R., Rubin, B. A., and Stasny, J. T.** (1968). Cleavage by formamide of intercapsomer bonds in adenovirus types 4 and 7 virions and hemagglutinins. *J Virol* *2*, 1086-1095.
- Norrby, E.** (1966). The relationship between the soluble antigens and the virion of adenovirus type 3. II. Identification and characterization of an incomplete hemagglutinin. *Virology* *30*, 608-617.
- Norrby, E.** (1968). Identification of soluble components of adenovirus type 11. *J Gen Virol* *2*, 123-133.
- Norrby, E., and Ankerst, J.** (1969). Biological characterization of structural components of Adenovirus type 12. *J Gen Virol* *5*, 183-194.
- Norrby, E., Bartha, A., Boulanger, P., Dreizin, R. S., Ginsberg, H. S., Kalter, S. S., Kawamura, H., Rowe, W. P., Russell, W. C., Schlesinger, W., and Wigand, R.** (1976). Adenoviridae. *Intervirology* *7*, 117-125.
- Norrby, E., Nyberg, B., Skaaret, P., and Lengyel, A.** (1967). Separation and characterization of soluble adenovirus type 9 components. *J Virol* *1*, 1101-1108.
- Norrby, E., and Skaaret, P.** (1968). Comparison between soluble components of adenovirus types 3 and 16 and of the intermediate strain 3-16 (the San Carlos agent). *Virology* *36*, 201-211.
- Norrby, E., and Wadell, G.** (1967). Soluble components of adenovirus type 4. *Virology* *31*, 592-600.
- Ojkic, D., and Nagy, E.** (2000). The complete nucleotide sequence of fowl adenovirus type 8. *J Gen Virol* *81*, 1833-1837.
- Otwinowski, Z., and Minor, W.** (1997a). *Macromolecular Crystallography, Part A*, edited by C.W. Carter Jr & R. M. Sweet. *Methods in Enzymology* *276*, 307-326.
- Otwinowski, Z., and Minor, W.** (1997b). Processing of X-ray diffraction data collected in oscillation mode. *Methods in Enzymology* *276*, 307-326.
- Parks, R. J.** (2005). Adenovirus protein IX: a new look at an old protein. *Mol Ther* *11*, 19-25.
- Parks, R. J., Chen, L., Anton, M., Sankar, U., Rudnicki, M. A., and Graham, F. L.** (1996). A helper-dependent adenovirus vector system: removal of helper virus by Cre-mediated excision of the viral packaging signal. *Proc Natl Acad Sci U S A* *93*, 13565-13570.
- Perez-Romero, P., Gustin, K. E., and Imperiale, M. J.** (2006). Dependence of the encapsidation function of the adenovirus L1 52/55-kilodalton protein on its ability to bind the packaging sequence. *J Virol* *80*, 1965-1971.
- Persson, B. D., Reiter, D. M., Marttila, M., Mei, Y. F., Casasnovas, J. M., Arnberg, N., and Stehle, T.** (2007). Adenovirus type 11 binding alters the conformation of its receptor CD46. *Nat Struct Mol Biol* *14*, 164-166.

- Pichla-Gollon, S. L., Drinker, M., Zhou, X., Xue, F., Rux, J. J., Gao, G. P., Wilson, J. M., Ertl, H. C., Burnett, R. M., and Bergelson, J. M.** (2007). Structure-based identification of a major neutralizing site in an adenovirus hexon. *J Virol* *81*, 1680-1689.
- Pieniasek, N. J., Slemenda, S. B., Pieniasek, D., Velarde, J., Jr., and Luftig, R. B.** (1990). Human enteric adenovirus type 41 (Tak) contains a second fiber protein gene. *Nucleic Acids Res* *18*, 1901.
- Prage, L., Pettersson, U., Hoglund, S., Lonberg-Holm, K., and Philipson, L.** (1970). Structural proteins of adenoviruses. IV. Sequential degradation of the adenovirus type 2 virion. *Virology* *42*, 341-358.
- Ramachandran, G. N., and Sasisekharan, V.** (1968). Conformation of polypeptides and proteins. *Adv Protein Chem* *23*, 283-438.
- Rao, L., Perez, D., and White, E.** (1996). Lamin proteolysis facilitates nuclear events during apoptosis. *J Cell Biol* *135*, 1441-1455.
- Rein, D. T., Breidenbach, M., and Curiel, D. T.** (2006). Current developments in adenovirus-based cancer gene therapy. *Future Oncol* *2*, 137-143.
- Rekosh, D. M., Russell, W. C., Bellet, A. J., and Robinson, A. J.** (1977). Identification of a protein linked to the ends of adenovirus DNA. *Cell* *11*, 283-295.
- Roberts, M. M., White, J. L., Grutter, M. G., and Burnett, R. M.** (1986). Three-dimensional structure of the adenovirus major coat protein hexon. *Science* *232*, 1148-1151.
- Roberts, R. J., O'Neill, K. E., and Yen, C. T.** (1984). DNA sequences from the adenovirus 2 genome. *J Biol Chem* *259*, 13968-13975.
- Roelvink, P. W., Lizonova, A., Lee, J. G., Li, Y., Bergelson, J. M., Finberg, R. W., Brough, D. E., Kovesdi, I., and Wickham, T. J.** (1998). The coxsackievirus-adenovirus receptor protein can function as a cellular attachment protein for adenovirus serotypes from subgroups A, C, D, E, and F. *J Virol* *72*, 7909-7915.
- Rosa-Calatrava, M., Grave, L., Puvion-Dutilleul, F., Chatton, B., and Kedinger, C.** (2001). Functional analysis of adenovirus protein IX identifies domains involved in capsid stability, transcriptional activity, and nuclear reorganization. *J Virol* *75*, 7131-7141.
- Rosa-Calatrava, M., Puvion-Dutilleul, F., Lutz, P., Dreyer, D., de The, H., Chatton, B., and Kedinger, C.** (2003). Adenovirus protein IX sequesters host-cell promyelocytic leukaemia protein and contributes to efficient viral proliferation. *EMBO Rep* *4*, 969-975.
- Rowe, W. P., Huebner, R. J., Gilmore, L. K., Parrott, R. H., and Ward, T. G.** (1953). Isolation of a cytopathogenic agent from human adenoids undergoing spontaneous degeneration in tissue culture. *Proc Soc Exp Biol Med* *84*, 570-573.
- Ruigrok, R. W., Barge, A., Albiges-Rizo, C., and Dayan, S.** (1990). Structure of adenovirus fibre. II. Morphology of single fibres. *J Mol Biol* *215*, 589-596.
- Ruigrok, R. W., Barge, A., Mittal, S. K., and Jacrot, B.** (1994). The fibre of bovine adenovirus type 3 is very long but bent. *J Gen Virol* *75* (Pt 8), 2069-2073.
- Russell, W. C.** (2000). Update on adenovirus and its vectors. *J Gen Virol* *81*, 2573-2604.

- Rux, J. J., and Burnett, R. M.** (2000). Type-specific epitope locations revealed by X-ray crystallographic study of adenovirus type 5 hexon. *Mol Ther* *1*, 18-30.
- Rux, J. J., and Burnett, R. M.** (2004). Adenovirus structure. *Hum Gene Ther* *15*, 1167-1176.
- Rux, J. J., Kuser, P. R., and Burnett, R. M.** (2003). Structural and phylogenetic analysis of adenovirus hexons by use of high-resolution x-ray crystallographic, molecular modeling, and sequence-based methods. *J Virol* *77*, 9553-9566.
- Saban, S. D., Silvestry, M., Nemerow, G. R., and Stewart, P. L.** (2006). Visualization of alpha-helices in a 6-angstrom resolution cryoelectron microscopy structure of adenovirus allows refinement of capsid protein assignments. *J Virol* *80*, 12049-12059.
- San Martin, C., and Burnett, R. M.** (2003). Structural studies on adenoviruses. *Curr Top Microbiol Immunol* *272*, 57-94.
- Sargent, K. L., Meulenbroek, R. A., and Parks, R. J.** (2004a). Activation of adenoviral gene expression by protein IX is not required for efficient virus replication. *J Virol* *78*, 5032-5037.
- Sargent, K. L., Ng, P., Eveleigh, C., Graham, F. L., and Parks, R. J.** (2004b). Development of a size-restricted pIX-deleted helper virus for amplification of helper-dependent adenovirus vectors. *Gene Ther* *11*, 504-511.
- Schoehn, G., El Bakkouri, M., Fabry, C. M., Billet, O., Estrozi, L. F., Le, L., Curiel, D. T., Kajava, A. V., Ruigrok, R. W., and Kremer, E. J.** (2008). 3D structure of canine adenovirus serotype 2 capsid. *J Virol*.
- Schoehn, G., Fender, P., Chroboczek, J., and Hewat, E. A.** (1996). Adenovirus 3 penton dodecahedron exhibits structural changes of the base on fibre binding. *Embo J* *15*, 6841-6846.
- Schulick, A. H., Vassalli, G., Dunn, P. F., Dong, G., Rade, J. J., Zamarron, C., and Dichek, D. A.** (1997). Established immunity precludes adenovirus-mediated gene transfer in rat carotid arteries. Potential for immunosuppression and vector engineering to overcome barriers of immunity. *J Clin Invest* *99*, 209-219.
- Segerman, A., Atkinson, J. P., Marttila, M., Dennerquist, V., Wadell, G., and Arnberg, N.** (2003). Adenovirus type 11 uses CD46 as a cellular receptor. *J Virol* *77*, 9183-9191.
- Seiradake, E., and Cusack, S.** (2005). Crystal structure of enteric adenovirus serotype 41 short fiber head. *J Virol* *79*, 14088-14094.
- Seiradake, E., Lortat-Jacob, H., Billet, O., Kremer, E. J., and Cusack, S.** (2006). Structural and mutational analysis of human Ad37 and canine adenovirus 2 fiber heads in complex with the D1 domain of coxsackie and adenovirus receptor. *J Biol Chem* *281*, 33704-33716.
- Shah, A. H., and Lipton, H. L.** (2002). Low-neurovirulence Theiler's viruses use sialic acid moieties on N-linked oligosaccharide structures for attachment. *Virology* *304*, 443-450.
- Shayakhmetov, D. M., Papayannopoulou, T., Stamatoyannopoulos, G., and Lieber, A.** (2000). Efficient gene transfer into human CD34(+) cells by a retargeted adenovirus vector. *J Virol* *74*, 2567-2583.
- Short, J. J., Pereboev, A. V., Kawakami, Y., Vasu, C., Holterman, M. J., and Curiel, D. T.** (2004). Adenovirus serotype 3 utilizes CD80 (B7.1) and CD86 (B7.2) as cellular attachment receptors. *Virology* *322*, 349-359.

- Short, J. J., Vasu, C., Holterman, M. J., Curiel, D. T., and Pereboev, A.** (2006). Members of adenovirus species B utilize CD80 and CD86 as cellular attachment receptors. *Virus Res* *122*, 144-153.
- Shu, W., Ji, H., and Lu, M.** (1999). Trimerization specificity in HIV-1 gp41: analysis with a GCN4 leucine zipper model. *Biochemistry* *38*, 5378-5385.
- Sirena, D., Lilienfeld, B., Eisenhut, M., Kalin, S., Boucke, K., Beerli, R. R., Vogt, L., Ruedl, C., Bachmann, M. F., Greber, U. F., and Hemmi, S.** (2004). The human membrane cofactor CD46 is a receptor for species B adenovirus serotype 3. *J Virol* *78*, 4454-4462.
- Smith, T. A., Idamakanti, N., Rollence, M. L., Marshall-Neff, J., Kim, J., Mulgrew, K., Nemerow, G. R., Kaleko, M., and Stevenson, S. C.** (2003). Adenovirus serotype 5 fiber shaft influences in vivo gene transfer in mice. *Hum Gene Ther* *14*, 777-787.
- Souquere-Besse, S., Pichard, E., Filhol, O., Legrand, V., Rosa-Calatrava, M., Hovanessian, A. G., Cochet, C., and Puvion-Dutilleul, F.** (2002). Adenovirus infection targets the cellular protein kinase CK2 and RNA-activated protein kinase (PKR) into viral inclusions of the cell nucleus. *Microsc Res Tech* *56*, 465-478.
- Staunton, D. E., Merluzzi, V. J., Rothlein, R., Barton, R., Marlin, S. D., and Springer, T. A.** (1989). A cell adhesion molecule, ICAM-1, is the major surface receptor for rhinoviruses. *Cell* *56*, 849-853.
- Stevenson, S. C., Rollence, M., White, B., Weaver, L., and McClelland, A.** (1995). Human adenovirus serotypes 3 and 5 bind to two different cellular receptors via the fiber head domain. *J Virol* *69*, 2850-2857.
- Stewart, P. L., Burnett, R. M., Cyrklaff, M., and Fuller, S. D.** (1991). Image reconstruction reveals the complex molecular organization of adenovirus. *Cell* *67*, 145-154.
- Stewart, P. L., Fuller, S. D., and Burnett, R. M.** (1993). Difference imaging of adenovirus: bridging the resolution gap between X-ray crystallography and electron microscopy. *Embo J* *12*, 2589-2599.
- Stoner, G. D., Williams, B., Kniazeff, A., and Shimkin, M. B.** (1973). Effect of neuraminidase pretreatment on the susceptibility of normal and transformed mammalian cells to bovine enterovirus 261. *Nature* *245*, 319-320.
- Suomalainen, M., Nakano, M. Y., Keller, S., Boucke, K., Stidwill, R. P., and Greber, U. F.** (1999). Microtubule-dependent plus- and minus end-directed motilities are competing processes for nuclear targeting of adenovirus. *J Cell Biol* *144*, 657-672.
- Sutjipto, S., Ravindran, S., Cornell, D., Liu, Y. H., Horn, M., Schlupe, T., Hutchins, B., and Vellekamp, G.** (2005). Characterization of empty capsids from a conditionally replicating adenovirus for gene therapy. *Hum Gene Ther* *16*, 109-125.
- Suzuki, T., Portner, A., Scroggs, R. A., Uchikawa, M., Koyama, N., Matsuo, K., Suzuki, Y., and Takimoto, T.** (2001). Receptor specificities of human respiroviruses. *J Virol* *75*, 4604-4613.
- Suzuki, Y., Ito, T., Suzuki, T., Holland, R. E., Jr., Chambers, T. M., Kiso, M., Ishida, H., and Kawaoka, Y.** (2000). Sialic acid species as a determinant of the host range of influenza A viruses. *J Virol* *74*, 11825-11831.

- Takahashi, E., Cohen, S. L., Tsai, P. K., and Sweeney, J. A.** (2006). Quantitation of adenovirus type 5 empty capsids. *Anal Biochem* 349, 208-217.
- Tan, P. K., Michou, A. I., Bergelson, J. M., and Cotten, M.** (2001). Defining CAR as a cellular receptor for the avian adenovirus CELO using a genetic analysis of the two viral fibre proteins. *J Gen Virol* 82, 1465-1472.
- Tollefson, A. E., Ryerse, J. S., Scaria, A., Hermiston, T. W., and Wold, W. S.** (1996). The E3-11.6-kDa adenovirus death protein (ADP) is required for efficient cell death: characterization of cells infected with adp mutants. *Virology* 220, 152-162.
- Toogood, C. I., Crompton, J., and Hay, R. T.** (1992). Antipeptide antisera define neutralizing epitopes on the adenovirus hexon. *J Gen Virol* 73 (Pt 6), 1429-1435.
- Top, F. H., Jr., Grossman, R. A., Bartelloni, P. J., Segal, H. E., Dudding, B. A., Russell, P. K., and Buescher, E. L.** (1971). Immunization with live types 7 and 4 adenovirus vaccines. I. Safety, infectivity, antigenicity, and potency of adenovirus type 7 vaccine in humans. *J Infect Dis* 124, 148-154.
- Toth, K., Spencer, J. F., Dhar, D., Sagartz, J. E., Buller, R. M., Painter, G. R., and Wold, W. S.** (2008). Hexadecyloxypropyl-cidofovir, CMX001, prevents adenovirus-induced mortality in a permissive, immunosuppressed animal model. *Proc Natl Acad Sci U S A* 105, 7293-7297.
- Tsuruta, Y., Pereboeva, L., Glasgow, J. N., Rein, D. T., Kawakami, Y., Alvarez, R. D., Rocconi, R. P., Siegal, G. P., Dent, P., Fisher, P. B., and Curiel, D. T.** (2007). A mosaic fiber adenovirus serotype 5 vector containing reovirus sigma 1 and adenovirus serotype 3 knob fibers increases transduction in an ovarian cancer ex vivo system via a coxsackie and adenovirus receptor-independent pathway. *Clin Cancer Res* 13, 2777-2783.
- Tsuzuki, J., and Luftig, R. B.** (1983). The adenovirus type 5 capsid protein IIIa is phosphorylated during an early stage of infection of HeLa cells. *Virology* 129, 529-533.
- Uhnoo, I., Wadell, G., Svensson, L., and Johansson, M. E.** (1984). Importance of enteric adenoviruses 40 and 41 in acute gastroenteritis in infants and young children. *J Clin Microbiol* 20, 365-372.
- Uncapher, C. R., DeWitt, C. M., and Colonno, R. J.** (1991). The major and minor group receptor families contain all but one human rhinovirus serotype. *Virology* 180, 814-817.
- Utagawa, E. T., Miyamura, K., Mukoyama, A., and Kono, R.** (1982). Neuraminidase-sensitive erythrocyte receptor for enterovirus type 70. *J Gen Virol* 63 (Pt 1), 141-148.
- van Breukelen, B., Brenkman, A. B., Holthuisen, P. E., and van der Vliet, P. C.** (2003). Adenovirus type 5 DNA binding protein stimulates binding of DNA polymerase to the replication origin. *J Virol* 77, 915-922.
- van Oostrum, J., and Burnett, R. M.** (1985). Molecular composition of the adenovirus type 2 virion. *J Virol* 56, 439-448.
- van Oostrum, J., Smith, P. R., Mohraz, M., and Burnett, R. M.** (1987). The structure of the adenovirus capsid. III. Hexon packing determined from electron micrographs of capsid fragments. *J Mol Biol* 198, 73-89.

- van Raaij, M. J., Chouin, E., van der Zandt, H., Bergelson, J. M., and Cusack, S.** (2000). Dimeric structure of the coxsackievirus and adenovirus receptor D1 domain at 1.7 Å resolution. *Structure Fold Des* 8, 1147-1155.
- van Raaij, M. J., Louis, N., Chroboczek, J., and Cusack, S.** (1999a). Structure of the human adenovirus serotype 2 fiber head domain at 1.5 Å resolution. *Virology* 262, 333-343.
- van Raaij, M. J., Mitraki, A., Lavigne, G., and Cusack, S.** (1999b). A triple beta-spiral in the adenovirus fibre shaft reveals a new structural motif for a fibrous protein. *Nature* 401, 935-938.
- Vellekamp, G., Porter, F. W., Sutjipto, S., Cutler, C., Bondoc, L., Liu, Y. H., Wylie, D., Cannon-Carlson, S., Tang, J. T., Frei, A., et al.** (2001). Empty capsids in column-purified recombinant adenovirus preparations. *Hum Gene Ther* 12, 1923-1936.
- Vellinga, J., Rabelink, M. J., Cramer, S. J., van den Wollenberg, D. J., Van der Meulen, H., Leppard, K. N., Fallaux, F. J., and Hoeben, R. C.** (2004). Spacers increase the accessibility of peptide ligands linked to the carboxyl terminus of adenovirus minor capsid protein IX. *J Virol* 78, 3470-3479.
- Vellinga, J., Uil, T. G., de Vrij, J., Rabelink, M. J., Lindholm, L., and Hoeben, R. C.** (2006). A system for efficient generation of adenovirus protein IX-producing helper cell lines. *J Gene Med* 8, 147-154.
- Vellinga, J., van den Wollenberg, D. J., van der Heijdt, S., Rabelink, M. J., and Hoeben, R. C.** (2005). The coiled-coil domain of the adenovirus type 5 protein IX is dispensable for capsid incorporation and thermostability. *J Virol* 79, 3206-3210.
- Vives, R. R., Lortat-Jacob, H., Chroboczek, J., and Fender, P.** (2004). Heparan sulfate proteoglycan mediates the selective attachment and internalization of serotype 3 human adenovirus dodecahedron. *Virology* 321, 332-340.
- Waddington, S. N., McVey, J. H., Bhella, D., Parker, A. L., Barker, K., Atoda, H., Pink, R., Buckley, S. M., Greig, J. A., Denby, L., et al.** (2008). Adenovirus serotype 5 hexon mediates liver gene transfer. *Cell* 132, 397-409.
- Wadell, G., and Norrby, E.** (1969). The soluble hemagglutinins of adenoviruses belonging to Rosen's subgroup 3. II. The slowly sedimenting hemagglutinin. *Arch Gesamte Virusforsch* 26, 53-62.
- Walters, R. W., Freimuth, P., Moninger, T. O., Ganske, I., Zabner, J., and Welsh, M. J.** (2002). Adenovirus fiber disrupts CAR-mediated intercellular adhesion allowing virus escape. *Cell* 110, 789-799.
- Wang, C. Y., Cusack, J. C., Jr., Liu, R., and Baldwin, A. S., Jr.** (1999). Control of inducible chemoresistance: enhanced anti-tumor therapy through increased apoptosis by inhibition of NF-kappaB. *Nat Med* 5, 412-417.
- Wang, H., Liaw, Y. C., Stone, D., Kalyuzhniy, O., Amiraslanov, I., Tuve, S., Verlinde, C. L., Shayakhmetov, D., Stehle, T., Roffler, S., and Lieber, A.** (2007). Identification of CD46 binding sites within the adenovirus serotype 35 fiber knob. *J Virol* 81, 12785-12792.
- Wang, K., Guan, T., Cheresch, D. A., and Nemerow, G. R.** (2000). Regulation of adenovirus membrane penetration by the cytoplasmic tail of integrin beta5. *J Virol* 74, 2731-2739.
- Webster, A., Hay, R. T., and Kemp, G.** (1993). The adenovirus protease is activated by a virus-coded disulphide-linked peptide. *Cell* 72, 97-104.

- Wickham, T. J., Filardo, E. J., Cheresch, D. A., and Nemerow, G. R.** (1994). Integrin alpha v beta 5 selectively promotes adenovirus mediated cell membrane permeabilization. *J Cell Biol* 127, 257-264.
- Wickham, T. J., Mathias, P., Cheresch, D. A., and Nemerow, G. R.** (1993). Integrins alpha v beta 3 and alpha v beta 5 promote adenovirus internalization but not virus attachment. *Cell* 73, 309-319.
- Wickham, T. J., Roelvink, P. W., Brough, D. E., and Kovesdi, I.** (1996a). Adenovirus targeted to heparan-containing receptors increases its gene delivery efficiency to multiple cell types. *Nat Biotechnol* 14, 1570-1573.
- Wickham, T. J., Segal, D. M., Roelvink, P. W., Carrion, M. E., Lizonova, A., Lee, G. M., and Kovesdi, I.** (1996b). Targeted adenovirus gene transfer to endothelial and smooth muscle cells by using bispecific antibodies. *J Virol* 70, 6831-6838.
- Wiethoff, C. M., Wodrich, H., Gerace, L., and Nemerow, G. R.** (2005). Adenovirus protein VI mediates membrane disruption following capsid disassembly. *J Virol* 79, 1992-2000.
- Wodrich, H., Guan, T., Cingolani, G., Von Seggern, D., Nemerow, G., and Gerace, L.** (2003). Switch from capsid protein import to adenovirus assembly by cleavage of nuclear transport signals. *Embo J* 22, 6245-6255.
- Wold, W. S., and Gooding, L. R.** (1989). Adenovirus region E3 proteins that prevent cytolysis by cytotoxic T cells and tumor necrosis factor. *Mol Biol Med* 6, 433-452.
- Wu, E., Fernandez, J., Fleck, S. K., Von Seggern, D. J., Huang, S., and Nemerow, G. R.** (2001). A 50-kDa membrane protein mediates sialic acid-independent binding and infection of conjunctival cells by adenovirus type 37. *Virology* 279, 78-89.
- Wu, E., Pache, L., Von Seggern, D. J., Mullen, T. M., Mikiyas, Y., Stewart, P. L., and Nemerow, G. R.** (2003). Flexibility of the adenovirus fiber is required for efficient receptor interaction. *J Virol* 77, 7225-7235.
- Wu, E., Trauger, S. A., Pache, L., Mullen, T. M., von Seggern, D. J., Siuzdak, G., and Nemerow, G. R.** (2004). Membrane cofactor protein is a receptor for adenoviruses associated with epidemic keratoconjunctivitis. *J Virol* 78, 3897-3905.
- Wu, H., Seki, T., Dmitriev, I., Uil, T., Kashentseva, E., Han, T., and Curiel, D. T.** (2002). Double modification of adenovirus fiber with RGD and polylysine motifs improves coxsackievirus-adenovirus receptor-independent gene transfer efficiency. *Hum Gene Ther* 13, 1647-1653.
- Xia, D., Henry, L. J., Gerard, R. D., and Deisenhofer, J.** (1994). Crystal structure of the receptor-binding domain of adenovirus type 5 fiber protein at 1.7 Å resolution. *Structure* 2, 1259-1270.
- Xu, L., Benson, S. D., and Burnett, R. M.** (2007). Nanoporous crystals of chicken embryo lethal orphan (CELO) adenovirus major coat protein, hexon. *J Struct Biol* 157, 424-431.
- Zhang, W., and Arcos, R.** (2005). Interaction of the adenovirus major core protein precursor, pVII, with the viral DNA packaging machinery. *Virology* 334, 194-202.
- Zhao, H., Granberg, F., Elfineh, L., Pettersson, U., and Svensson, C.** (2003). Strategic attack on host cell gene expression during adenovirus infection. *J Virol* 77, 11006-11015.
- Zhou, L., Luo, Y., Wu, Y., Tsao, J., and Luo, M.** (2000). Sialylation of the host receptor may modulate entry of demyelinating persistent Theiler's virus. *J Virol* 74, 1477-1485.

Zubieta, C., Blanchoin, L., and Cusack, S. (2006). Structural and biochemical characterization of a human adenovirus 2/12 penton base chimera. *Febs J* 273, 4336-4345.

Zubieta, C., Schoehn, G., Chroboczek, J., and Cusack, S. (2005). The structure of the human adenovirus 2 penton. *Mol Cell* 17, 121-135.