

# UNIVERSITÉ PARIS-SUD 11

## ECOLE DOCTORALE :

INNOVATION THÉRAPEUTIQUE : DU FONDAMENTAL A L'APPLIQUÉ

PÔLE : INGENIERIE DES PROTEINES ET CIBLES THERAPEUTIQUES

## DISCIPLINE :

STRUCTURE, FONCTION ET INGENIERIE DES PROTEINES

ANNÉE 2012 - 2013

SÉRIE DOCTORAT N° 1219

## THÈSE DE DOCTORAT

soutenue le 17/01/2013

par

**Olivier PREUX**

**Assemblage et maturation de la capsidie du bactériophage  
T5 : analyse des processus d'expansion et de décoration**

**Directeur de thèse :** Pascale BOULANGER

CR (IBBMC, CNRS/Univ. Paris-Sud - Orsay)

### **Composition du jury :**

*Président du jury :*

Herman VAN TILBEURGH

Professeur (IBBMC, CNRS/Univ. Paris-Sud - Orsay)

*Rapporteurs :*

Guy SCHOEHN

DR (IBS, CEA/Univ. J.F. - Grenoble)

Frank GABEL

Ingénieur-Chercheur (IBS, CEA - Grenoble)

*Examineurs :*

Paulo TAVARES

DR (LVMS, CNRS - Gif-sur-Yvette)

Laurent DEBARBIEUX

Chercheur (Institut Pasteur - Paris)

*Membre invité :*

Dominique DURAND

DR (IBBMC, CNRS/Univ. Paris-Sud - Orsay)



# Remerciements

Je tiens à remercier toutes les personnes grâce à qui ce travail a pu être réalisé.

Je remercie tout d'abord Guy Schoehn et Frank Gabel de m'avoir fait l'honneur d'accepter de juger mon travail. Merci aux examinateurs Paulo Tavares et Laurent Debarbieux ainsi qu'à Dominique Durand pour leur participation à la soutenance, et merci à Herman van Tilbeurgh pour avoir accepté de la présider.

Je remercie Pascale Boulanger pour la qualité de son encadrement tout au long de mes trois années de doctorat, effectuées dans l'équipe de Virologie Bactérienne de l'Institut de Biochimie et de Biophysique Moléculaire et Cellulaire. Je remercie les directeurs, passé et présent, de cet institut, Lucienne Letellier et Michel Desmadril pour m'avoir permis d'y conduire mon doctorat. Je remercie tous les membres de cette équipe que j'ai eu la chance de côtoyer. En particulier, je remercie très chaleureusement Jeannine Drouin-Wahbi, qui a énormément contribué à la réussite de ce travail. Je remercie également Marta De Frutos et Lucienne Letellier pour les nombreuses et fructueuses discussions scientifiques et extra-scientifiques que nous avons eues au cours de ces trois longues années. Je remercie Aurélie Bertin et Claire Boulogne pour leur aide précieuse en microscopie électronique. Merci à Madalena Renouard pour ses nombreux conseils et ses nombreuses aides tout au long de ma thèse. Je remercie également Alexis Huet qui a initié ce projet et ma fait découvrir ce sujet pendant mon stage de M2.

Je remercie à nouveau Dominique Durand, ainsi que Patrice Vachette, pour leur encadrement attentionné sur les expériences de SAXS ainsi que sur l'analyse des données

obtenues. Ses expériences se sont déroulées en grande partie sur la ligne SWING du synchrotron SOLEIL, grâce au travail de Javier Perez et Didier Trévarin, que je remercie également.

Je remercie tous les membres de l'équipe Fonction et Architecture des Assemblages Macromoléculaires m'ayant permis de réaliser les expériences de cristallogénèse et de MALS : en premier lieu Inès Gallay, mais également Mounira Tiouajni, Zaineb Fourati-Kammoun, Nouredine Lazar, Bruno Collinet et Herman van Tilbeurgh.

Je remercie Duquerroy Stéphane, qui m'a permis d'effectuer les tentatives de résolution de la structure de la capsidie par cristallographie.

Je remercie Magali Aumont-Nicaise et Michel Desmadril, qui m'ont permis d'effectuer des expériences de DSC et m'ont permis d'analyser les données obtenues.

Je remercie également tous les participants des réunions doc/post-doc, tous les joueurs de foot du vendredi, Alexandre Ghazi pour les nombreuses et passionnantes conversations qui ont enrichi mes trois dernières années, ainsi que tous les membres de l'IBBMC avec qui j'ai eu le plaisir d'échanger.

En dernier lieu je tiens à remercier tous mes amis qui m'ont aidé à traverser toutes les épreuves rencontrées pendant mon doctorat, en particulier Jacques-Antoine Maisonobe qui par sa relecture attentive a contribué à la qualité de ce manuscrit, et toute ma famille, qui m'a soutenu durant ces trois ans, et plus spécialement mes parents qui ont permis d'améliorer l'orthographe et le style de cet ouvrage.

<b>Remerciements .....</b>	<b>3</b>
<b>Liste des abréviations utilisées .....</b>	<b>8</b>
<b>INTRODUCTION.....</b>	<b>9</b>
<b>I - Les bactériophages .....</b>	<b>10</b>
1- Présentation.....	10
2- Historique de la découverte des bactériophages .....	10
3- Applications .....	11
4- L'importance écologique des bactériophages .....	13
5- Classification et morphologie des phages.....	15
5-1 Classification.....	15
5-2 Morphologie des phages caudés.....	18
5-3 La Capside des phages caudés .....	20
5-3-1 La symétrie icosaédrique .....	20
5-3-2 Le nombre de triangulation.....	21
5-3-3 Cas des capsides oblongues .....	22
6- Assemblage et maturation de la capside .....	24
6-1 Cycle de vie du phage .....	24
6-2 Assemblage du virion.....	26
6-2-1 Formation de la procapside.....	27
6-2-2 Maturation de la capside.....	29
6-2-3 Etapes finales de l'assemblage du virion.....	32
6-3 L'expansion de la capside des phages caudés.....	34
6-3-1 Déclenchement de l'expansion et lien avec l'encapsulation de l'ADN.....	36
6-3-2 Les étapes d'expansion du phage HK97.....	37
7- Le bactériophage T5.....	42
7-1 Présentation.....	42
7-2 L'assemblage in vitro de la procapside.....	45
<b>II - Introduction méthodologique .....</b>	<b>47</b>
1- La diffusion des rayons X aux petits angles.....	47
1-1 Rappels historiques .....	47
1-2 Théorie de la diffusion des rayons X aux petits angles.....	48
1-2-1 La diffusion par deux électrons .....	48
1-2-2 La diffusion par N électrons .....	50
1-2-3 Diffusion d'une particule en solution .....	51
1-2-4 La loi de Guinier.....	52
1-2-5 La fonction de distribution des distances.....	55
1-2-6 Le problème des radicaux libres .....	57
1-2-7 Reconstitution de l'enveloppe moléculaire à l'aide du programme GASBOR .....	58
2- La décomposition en valeurs singulières (SVD).....	59
2-1 Définition .....	59
2-2 Exemple de la décomposition d'une matrice à coefficients réels .....	59
2-3 Application à l'étude d'un ensemble de courbes .....	60
2-3-1 Courbes théoriques sans bruit de fond expérimental .....	61
2-3-1 Courbes expérimentales.....	63

<b>III- Présentation de la thématique de recherche .....</b>	<b>66</b>
<b>ETUDE CINÉTIQUE DU PROCESSUS D'EXPANSION.....</b>	<b>69</b>
<b>I- Introduction .....</b>	<b>70</b>
1- Etudes cinétiques de l'expansion du phage HK97 .....	70
1-1 Etude de la première phase de l'expansion .....	70
1-2 Etude de la seconde phase de l'expansion .....	72
2- Présentation de la thématique de recherche .....	73
<b>II- Etude cinétique de l'expansion .....</b>	<b>74</b>
1- Résumé de l'article.....	74
2- Article .....	76
<b>III- Modèles cinétiques de l'expansion .....</b>	<b>124</b>
<b>ETUDE STRUCTURALE ET FONCTIONNELLE DE LA PROTEINE DE DECORATION PB10 .....</b>	<b>129</b>
<b>I- INTRODUCTION .....</b>	<b>130</b>
1- Les protéines HOC.....	132
1-1 Présentation.....	132
1-2 Données structurales des protéines HOC.....	132
1-3 Rôle biologique de HOC.....	136
2- Les fibres de la capsid de $\phi 29$ .....	137
3- La protéine gp12 du phage SPP1 .....	139
4- Applications biotechnologiques des protéines de décoration .....	140
5- Présentation de la thématique de recherche .....	141
<b>II- RESULTATS.....</b>	<b>142</b>
1- Clonage, production et purification de pb10.....	142
1-1 Production .....	142
1-2 Purification.....	142
2- Etude fonctionnelle de pb10 .....	146
2-1 Etude de la fixation de pb10 sur les capsides vides d'ADN .....	146
2-1-1 Etude de la fixation de pb10 par ultracentrifugation .....	146
2-1-2 Etude par migration sur gel d'agarose .....	147
2-1-3 Analyse de l'affinité et mise en évidence de la coopérativité de la fixation.....	148
2-2 Etude par calorimétrie différentielle à balayage du rôle de la protéine de décoration.....	151
3- Etude structurale de pb10.....	153
3-1 Etude par SAXS .....	153
3-1-1 Analyse de la courbe de diffusion et de la fonction de distribution des distances.....	153
3-1-2 Reconstitution de l'enveloppe moléculaire.....	154
3-1-3 Analyse des résultats de SAXS à la lumière de données bio-informatiques.....	156
3-2 Essais de cristallogénèse de pb10 .....	158
<b>III- Conclusion et perspectives.....</b>	<b>160</b>
1- Récapitulatif des résultats obtenus .....	160

2- Fixation de pb10 sur la capside.....	161
3- Données structurales .....	162
<b>IV- MATERIEL ET METHODES .....</b>	<b>164</b>
1- Clonage, production et purification de pb10.....	164
1-1 Clonage .....	164
1-2 Surexpression en système bactérien.....	164
1-3 Préparation des extraits cellulaires.....	165
1-4 Purification des protéines par chromatographie.....	165
1-4-1 Chromatographie d'affinité sur un métal immobilisé (IMAC).....	165
1-4-2 Chromatographie d'échange de cation .....	166
1-4-3 Chromatographie d'exclusion de taille (SEC).....	166
1-4-4 Concentration et contrôle de la qualité de l'échantillon protéique purifié.....	166
1-5 Contrôle de l'état d'oligomérisation et de l'intégrité de la protéine pb10 purifiée.....	167
1-5-1 SEC-MALS.....	167
1-5-2 Spectrométrie de masse (Q-TOF).....	167
2- Etudes Fonctionnelles et Structurales .....	168
2-1 Etude par ultracentrifugation de la fixation de pb10 sur les capsides vides d'ADN.....	168
2-2 Etude par migration sur gel d'agarose natif.....	168
2-3 Etude par calorimétrie différentielle à balayage .....	169
2-4 Etude structurale de pb10 par diffusion des rayons X aux petits angles.....	169
2-5 Essais de cristallogenèse de la protéine pb10 .....	170
3- Production, purification et expansion des capsides.....	171
3-1 Culture des bactéries et production des capsides .....	171
3-2 Purification des capsides vides.....	172
3-3 Expansion des capsides in vitro .....	172
3-4 Normalisation de la concentration en capside.....	172
 <b>CONCLUSION GÉNÉRALE .....</b>	 <b>175</b>
 <b>ANNEXES .....</b>	 <b>179</b>
 <b>Annexe 1 : Calcul de la concentration d'un échantillon de capside vide.....</b>	 <b>180</b>
 <b>Annexe 2 : La cristallogenèse des protéines .....</b>	 <b>182</b>
1- Théorie .....	182
2- Stratégie expérimentale.....	184
2-1 La méthode du criblage.....	184
2-2 La méthode de diffusion de la vapeur .....	185
2-3 L'optimisation des conditions de cristallogenèse.....	188
 <b>Annexe 3 : Calcul des concentrations des différentes espèces lors d'une transition irréversible.....</b>	 <b>189</b>
1- Transition irréversible en une étape .....	189
2- Transition irréversible en deux étapes.....	190
 <b>BIBLIOGRAPHIE .....</b>	 <b>193</b>

## Liste des abréviations utilisées

<b>ATP</b>	Adénosine Tri-Phosphate
<b>DLS</b>	Dynamic Light Scattering, diffusion dynamique de la lumière
<b>DSC</b>	Differential Scanning Calorimetry, calorimétrie différentielle à balayage
<b>DTT</b>	(2S,3S)-1,4-Bis-sulfanylbutane-2,3-diol
<b>EDTA</b>	acide Ethylène-Diamine-TétraAcétique
<b>HOC/SOC</b>	Highly antigenic Outer Capsid protein/Small Outer Capsid Protein
<b>HSQC</b>	Heteronuclear Single Quantum Coherence, cohérence quantique hétéronucléaire
<b>LB</b>	Luria Bertani
<b>MALDI-TOF</b>	Matrix-Assisted Laser Desorption/Ionization mass spectrometry – Time Of Flight
<b>NSD</b>	Normalized Spatial Discrepancy, Divergence Spatiale Normalisée
<b>PDB</b>	Protein Data Bank
<b>Q-TOF</b>	Quadripôle – Time Of Flight
<b>RMN</b>	Résonance Magnétique Nucléaire
<b>SAXS</b>	Small Angle X-ray Scattering, diffusion des rayons X aux petits angles
<b>SDS PAGE</b>	Sodium Dodecyl Sulfate PolyAcrylamide Gel Electrophoresis
<b>SEC-MALS</b>	Size Exclusion Chromatography – Multi Angle Light Scattering, Chromatographie d'exclusion stérique – Diffusion de la lumière à angles multiples
<b>SPR</b>	Surface Plasmon Resonance, résonance plasmonique de surface
<b>Tris</b>	Trishydroxyméthylaminométhane

# **INTRODUCTION**

# I - Les bactériophages

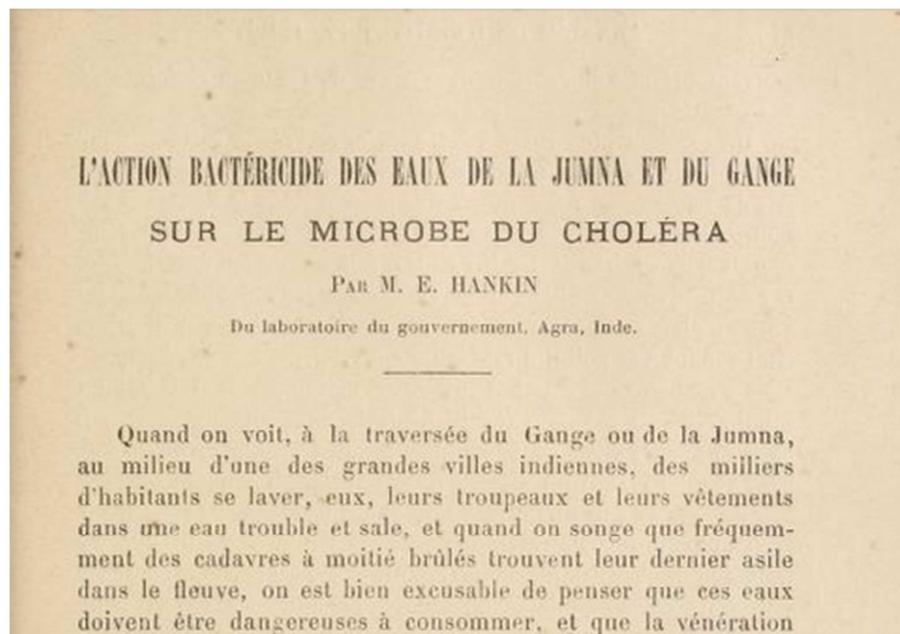
## 1- Présentation

Les bactériophages sont des virus infectant les bactéries. Comme tous les virus, ils ne possèdent pas d'activité métabolique propre et utilisent la machinerie biosynthétique de la cellule qu'ils infectent pour leur propre réplication. Ils sont présents partout où des bactéries sont présentes, c'est-à-dire dans tous les milieux terrestres et aquatiques, ainsi que dans toutes les espèces animales et végétales.

## 2- Historique de la découverte des bactériophages

La première observation indirecte de phage est décrite en 1896 dans les annales de l'Institut Pasteur par un biologiste Anglais, H.Hankin (figure 1) (Hankin 1896). Cette étude décrit les propriétés antibactériennes d'une substance alors inconnue, présente dans les eaux de la Jumna et du Gange, en Inde, notamment sa propension à limiter la diffusion de la bactérie *Vibrio cholerae* responsable du choléra. En 1914, un microbiologiste anglais, Frederick Twort, observe au microscope des zones vitreuses dans ses colonies de bactéries. Après s'être rendu compte que ces zones étaient le résultat de la destruction des cellules bactériennes, il les a ensuite prélevées et transmises d'une colonie à l'autre. Son étude, publiée en 1915 dans *The Lancet* (Twort 1915), décrit donc un agent doté de propriétés antibactériennes, mais ne précise pas sa nature. Enfin en 1917, Félix d'Hérelle, un chercheur franco-canadien décrit l'isolement et les propriétés de ce qu'il considèrera dans un premier

temps comme un microbe antagoniste (d'Herelle 1917), et auquel il donna le nom de bactériophage (du grec *baktêria*: bâton et *phagein* : manger) (Duckworth 1976).



**Figure 1. Première publication décrivant une observation indirecte d'un bactériophage.** Adaptée de (Hankin 1896).

### 3- Applications

Dans ce même article, Félix d'Hérelle envisage leur utilisation en vue de lutter contre les infections bactériennes. Cette application, appelée phagothérapie, a été rapidement reconnue par un grand nombre de scientifiques comme étant une voie possible pour combattre les infections bactériennes. Ainsi, avant la deuxième Guerre Mondiale, des médicaments à base de phages sont commercialisés par de grands groupes pharmaceutiques, dont *Eli Lilly* aux États-Unis d'Amérique et *Robert & Carrière* en France. En 1934, Félix d'Hérelle fonde avec G.Eliava en Géorgie un institut de recherche sur les bactériophages. Néanmoins il doit quitter précipitamment ce pays quand, durant une des grandes purges staliniennes, un de ses collaborateurs est fusillé pour avoir « tenté d'empoisonner un puits avec des bactériophages ». Dans le bloc de l'ouest, les recherches sur la phagothérapie seront arrêtées dans les années

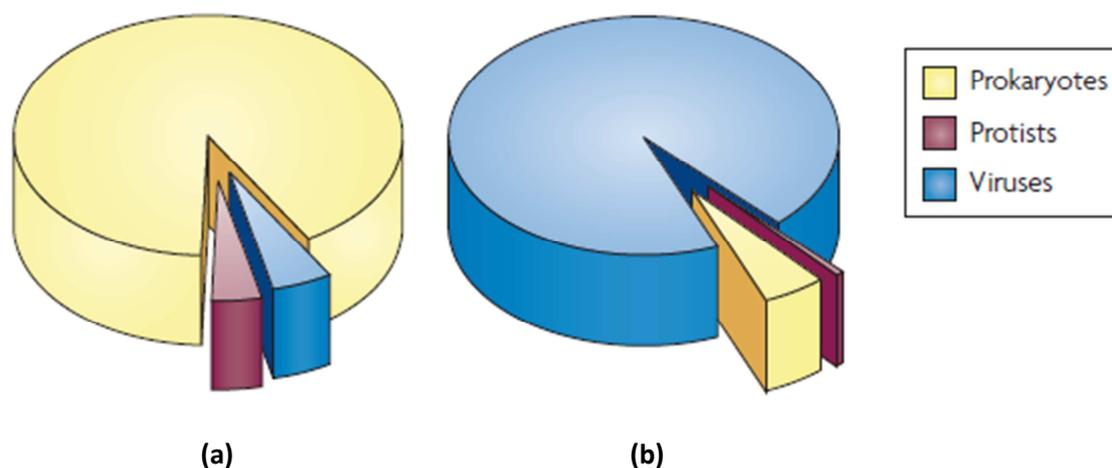
1940, date à laquelle les premiers antibiotiques sont utilisés de manière thérapeutique. Il est à noter que de nos jours, le nombre de bactéries résistantes aux antibiotiques augmentant fortement, les recherches visant à développer la thérapie phagique ou s'inspirant du mode d'action des phages suscitent un regain d'intérêt (Hanlon 2007; Kutateladze and Adamia 2010).

A partir des années 1950, la recherche sur les bactériophages a eu un rôle important dans l'essor de la biologie moléculaire. Ainsi, l'expérience de Hershey et Chase (Hershey and Chase 1952) a permis de confirmer que l'ADN était le support de l'information génétique. La première identification de l'ARN jouant le rôle d'intermédiaire lors de la synthèse des protéines, l'ARNm, est accomplie par E. Volkin et L. Astrachan en 1957, à l'aide de l'étude de l'infection d'*E. Coli* par le bactériophage T2. De même, les travaux de D. Dussoix et W. Arber sur les mécanismes par lesquels les bactéries deviennent résistantes à l'infection phagique permirent la découverte des premières enzymes de restriction en 1962 (Dussoix and Arber 1962). En 1977, le premier génome complet d'un organisme biologique est séquencé, il s'agit du génome du bactériophage  $\phi$ X174 (Sanger, Air et al. 1977).

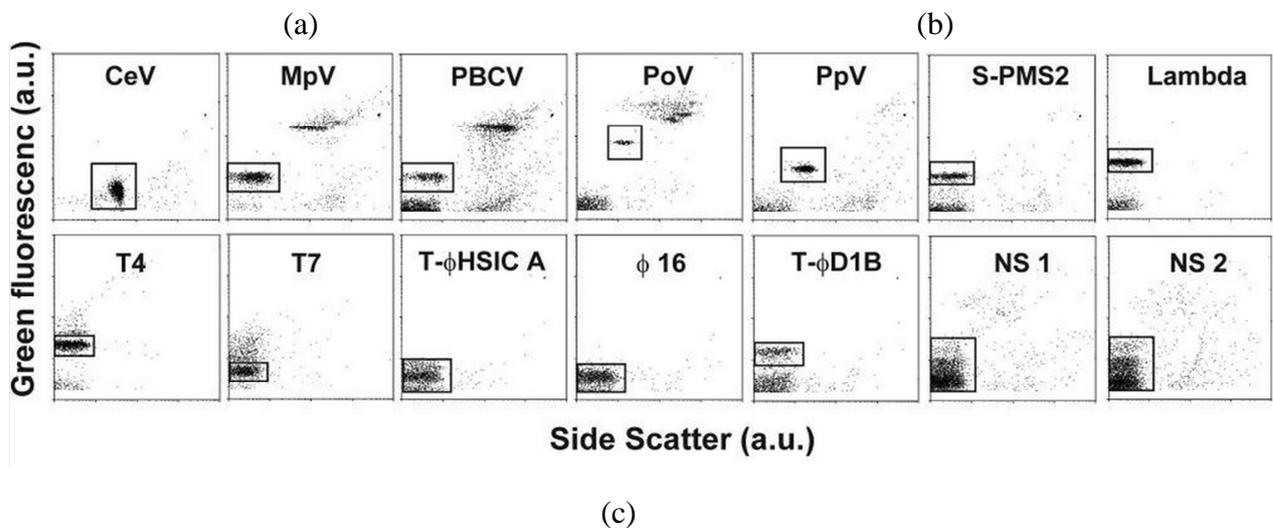
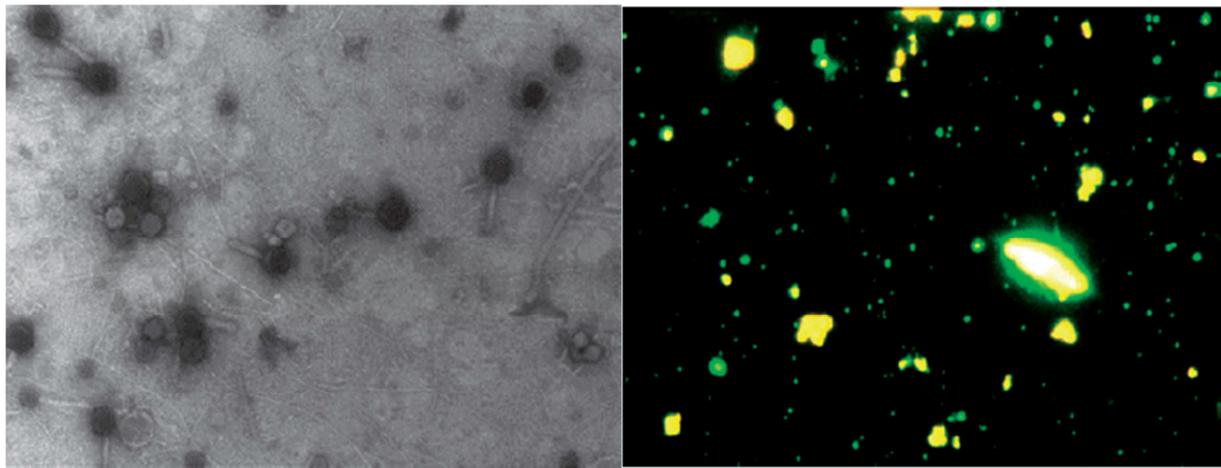
De nos jours, des phages ou des enzymes de phages sont couramment utilisés en biologie moléculaire (Livingston 1992) et en biotechnologie (Murphy 2012). Parmi les applications directes, nous pouvons noter la technique du *phage display*, l'utilisation du phage  $\lambda$  comme vecteur de clonage pour insérer de l'ADN dans les bactéries (Lech, Brent et al. 2001), l'utilisation de ligases (phage T4) (Pascal 2008), polymérase (phages T7 et T4) (Kubista, Andrade et al. 2006) (Finan, Torella et al. 2012) ou de différents promoteurs spécifiques (phages T7, T5 et Sf6) utilisés pour la surexpression contrôlée de protéines hétérologues chez les bactéries.

## 4- L'importance écologique des bactériophages

Les phages sont de très loin l'entité biologique la plus abondante sur Terre (Breitbart 2012). Leur concentration peut être estimée par microscopie électronique, par microscopie en épifluorescence ou par cytométrie en flux (figure 3). Elle varie d'environ  $10^8$  phages par millilitre dans les eaux côtières à  $3 \times 10^6$  phages par millilitre dans les zones bathypélagique et abyssopélagique (à plus de 1000 mètres de profondeur) de l'océan (Guixa-Boixereu, Vaque et al. 2002; Ortmann and Suttle 2005; Evans and Brussaard 2012). Leur nombre total est estimé à environ  $10^{30}$  à  $10^{31}$ , et le rapport entre le nombre de phages et celui des bactéries est d'environ 10 (figure 2). Il est fascinant de se représenter tous ces phages, chacun mesurant environ 100 nm de long, mis bout à bout : ils constitueraient une chaîne d'une longueur d'environ 10 à 100 millions d'années lumières, correspondant à 2 à 20 aller-retour vers la galaxie Andromède depuis la terre, ou encore à 100 à 1000 fois le diamètre de notre galaxie.



**Figure 2. Comparaison de la biomasse (a) et de l'abondance (b) des virus avec celles des procaryotes et eucaryotes dans les océans.** Les virus sont de très loin les entités biologiques les plus abondantes sur terre. Les phages représentent la très grande majorité de cette abondance. Toutefois, en raison de leur faible masse, ils ne correspondent qu'à environ 5% de la biomasse. Figure adaptée de (Suttle 2007).



**Figure 3. Techniques permettant d'estimer la concentration virale en milieu aquatique.**

(a) La microscopie électronique par transmission est l'unique méthode qui apporte des données à la fois sur l'abondance et sur la morphologie des particules virales (Borsheim, Bratbak et al. 1990). Cependant, à cause des étapes de concentration, de coloration par l'acétate d'uranyle ou par l'acide phospho-tungstique et de dépôt sur la grille, l'estimation de la concentration virale n'est pas fiable. (b) La microscopie en épifluorescence est la méthode actuellement la plus utilisée (Suttle 2010). Après concentration des virus, leurs acides nucléiques (ADN ou ARN) sont marqués avec un intercalant fluorescent et visualisés au microscope. Cette méthode permet aujourd'hui une grande fiabilité de la mesure de la concentration en particules phagiques (Wen, Ortmann et al. 2004; Suttle 2005). (c) L'utilisation de la cytométrie en flux est plus récente (Marie, Brussaard et al. 1999). Cette méthode à haut débit est très fiable (Brussaard 2004). Elle permet également la quantification de sous-populations de virus dont le signal de fluorescence ou de diffusion de la lumière diffère.

(a) et (b) : figures adaptées de (Suttle 2007) ; (c) : figure adaptée de (Brussaard 2004).

Les phages ont une grande importance dans la régulation, la dynamique et la biodiversité des populations bactériennes. On estime que chaque jour,  $10^{28}$  infections virales ont lieu dans les océans, tuant 20 à 40% des bactéries marines (Suttle 2007). L'influence des phages sur les populations bactériennes s'effectue de plusieurs manières. Premièrement, ils régulent les populations de bactéries à croissance la plus rapide, phénomène connu sous l'expression de « *killing the winner* » (Proctor and Fuhrman 1990; Weinbauer and Rassoulzadegan 2004). Ceci permet de préserver le grand nombre d'espèces et la diversité des génomes bactériens. Les virus sont aussi responsables de transferts horizontaux de gènes entre différentes espèces bactériennes, influençant ainsi la spéciation chez les bactéries.

## **5- Classification et morphologie des phages**

### **5-1 Classification**

Le très grand nombre de particules phagiques est associé à une très importante biodiversité. En effet, on estime qu'il y a en moyenne plus de 10 espèces de phages infectant chaque espèce de bactéries (Rohwer 2003). Le nombre d'espèces bactériennes étant estimé à  $10^7$  (Curtis, Sloan et al. 2002), il y aurait ainsi plus de  $10^8$  espèces de phages. Cependant, à cause des fortes contraintes structurales auxquelles sont soumises les protéines des phages, l'énorme diversité génomique des phages se traduit par une diversité structurale très limitée (Abrescia, Bamford et al. 2012).

Il n'existe pas de méthode universelle et systématique de classification des phages (Ackermann 2009). Cette classification est fondée principalement sur :

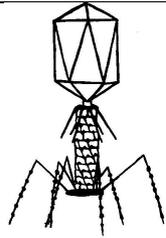
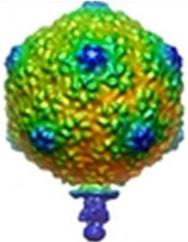
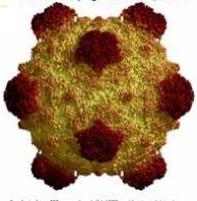
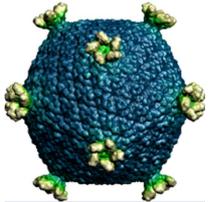
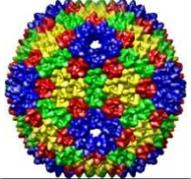
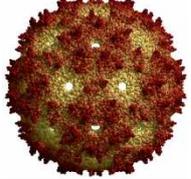
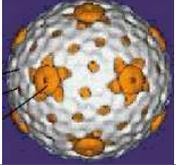
- leur morphologie (symétrie de la capsid (icosaédrique, hélicoïdale), présence d'une queue ou non, enveloppe de lipides entourant la capsid)

- la nature de leur génome : virus à ADN ou à ARN, acide nucléique linéaire ou circulaire, double ou simple brin
- les analogies de structure tridimensionnelle entre les protéines qui constituent la particule virale
- d'autres propriétés, comme le mode d'assemblage, ou le site de fixation sur la bactérie.

En revanche, la nature de leur hôte n'est peu ou pas prise en compte (Fauquet CM 2005). Au contraire des analogies structurales, le critère de l'homologie de séquence n'est utilisé que pour la classification en genre, cette homologie étant trop faible pour classer les phages dans les taxons supérieurs (Abrescia, Bamford et al. 2012).

Parmi les 100 millions d'espèces putatives de phages, seulement 5568 ont été décrites par microscopie électronique (Ackermann 2009). Parmi celles-ci, 5360 possèdent une queue et une capsidie icosaédrique. Ces phages appartiennent à l'ordre des *Caudovirales*, et sont couramment appelés phages caudés. Les autres phages décrits (208 espèces) sont polyédriques, filamenteux ou pléomorphes (figure 4).

Il est intéressant de noter qu'il existe de fortes homologies structurales entre des familles de virus infectant les trois domaines du vivant. Ainsi, les protéines majeures de capsidie des phages de la famille des *Herpesviridae*, comprenant notamment les virus responsables de l'Herpès et de la Varicelle chez l'Homme, ont des repliements très voisins aux protéines majeures de capsidie des phages de l'ordre des *Caudovirales* (Baker, Jiang et al. 2005). De plus, ce repliement n'est partagé par aucune autre famille de virus infectant les Eucaryotes ou les Bactéries. Ceci suggère une ascendance commune à ces 2 groupes de virus. Il a également été suggéré que certains virus infectant des Archées ont une ascendance commune avec les *Caudovirales* (Krupovic, Forterre et al. 2010).

Forme	Acide nucléique	Famille	Nombre de genres/espèces	Caractéristiques	Exemples	Morphologie
Phages caudés, capsidie icosaédrique (Ordre des <i>Caudovirales</i> )	ADN double brin linéaire	<i>Myoviridae</i>	6/1320	Queue contractile	T4	
		<i>Siphoviridae</i>	7/3229	Queue longue non contractile	$\lambda$ , T5, HK97, SPP1	
		<i>Podoviridae</i>	4/771	Queue courte	T7, $\phi$ 29, P22/L, sf6	
Polyédrique	ADN simple brin circulaire	<i>Microviridae</i>	4/40	Capsomères visibles, petite capsidie non enveloppée	$\phi$ x174	
	ADN double brin circulaire	<i>Corticoviridae</i>	1/3	Capsidie contenant plusieurs couches lipidiques	PM2	
	ADN double brin linéaire	<i>Tectiviridae</i>	1/19	Couche intérieure lipidique	PRD1	
	ARN simple brin linéaire	<i>Leviviridae</i>	2/39	Très petits génomes	Ms2	
	ARN double brin linéaire	<i>Cystoviridae</i>	1/3	Enveloppe lipidique	$\phi$ 6, $\phi$ 12	

Filamenteux	ADN simple brin circulaire	<i>Inoviridae</i>	2/67	Filaments longs	M13, fd	
	ADN double brin linéaire	<i>Lipo- thrixviridae</i>	4/7	Tiges longues avec enveloppe lipoprotéique	TTV1	
		<i>Rudiviridae</i>	1/3	Tiges droites sans enveloppe	SIRV-1	
Pléomorphe	ADN double brin circulaire	<i>Fuselloviridae</i>	1/11	Pas de capside, habitat : sources chaudes	SSV1	
		<i>Plasmaviridae</i>	1/5	Pas de capside		

**Figure 4. Classement et propriété des bactériophages et des virus d'Archaea.** Le nombre d'espèces indiqué dans ce tableau correspond au nombre d'espèces examinées par microscopie électronique, il est tiré de (Ackermann 2007). Les dimensions respectives des différents exemples de phages présentés dans la colonne « morphologie » ne sont pas respectées. Les exemples de structures sont tirés de : phage T5 (Effantin, Boulanger et al. 2006), T7 : (Agirrezabala, Martin-Benito et al. 2005),  $\phi$ x174 : (McKenna, Xia et al. 1992), PM2 : (Abrescia, Grimes et al. 2008), PRD1 : (Abrescia, Cockburn et al. 2004), Ms2 : (Golmohammadi, Valegard et al. 1993), fd : (Marvin 1990). Les images de microscopie électronique à coloration négative sont tirées de : phage TTV1 : (Zillig, Gropp et al. 1986), SIRV-1 : (Zillig, Arnold et al. 1998), SSV1 : (Arnold, She et al. 1999).  
Figure adaptée de (Ackermann 2009).

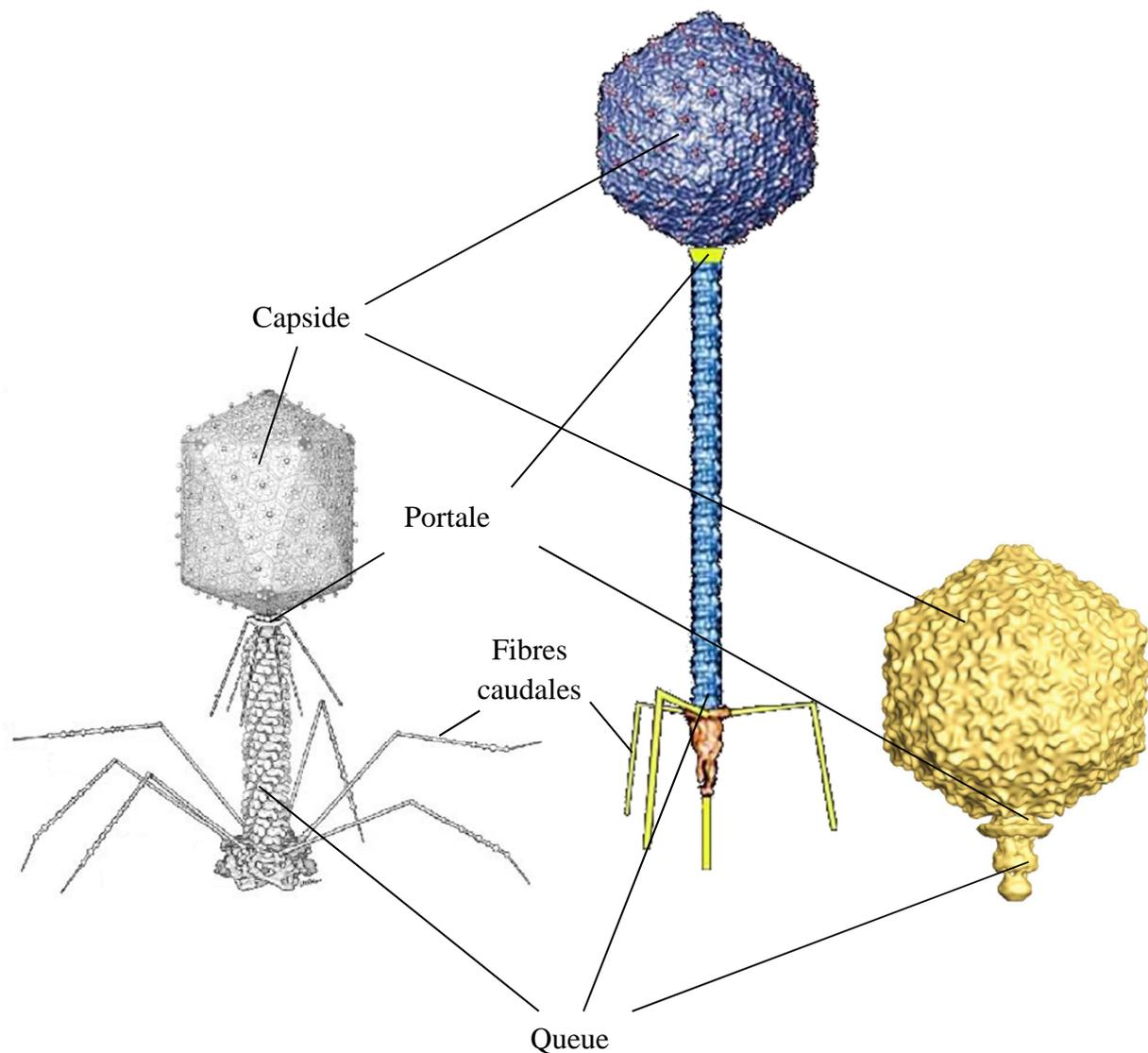
## 5-2 Morphologie des phages caudés

Les phages caudés possèdent une capsidie de symétrie icosaédrique ou pseudo-icosaédrique et une queue (figure 5). En revanche ils n'ont pas d'enveloppe lipidique. Leur capsidie contient un ADN double brin linéaire. Leur queue possède une symétrie hélicoïdale. Le site de fixation à l'hôte se trouve au bout de cette queue.

L'ordre des *caudovirales* comprend 3 familles :

- Les *siphoviridae* (~3200 observations (Ackermann 2009), soit 61% des espèces observées). Ils possèdent une queue longue, flexible et non contractile. La famille des *siphoviridae* comprend notamment les phages  $\lambda$ , HK 97, T5, SIO-2 et SPP1.

- Les *Myoviridae* (~1300 observations, soit 25% des espèces observées). Ils possèdent une queue longue contractile. La famille des *Myoviridae* comprend notamment les phages T4 et P2.
- Les *Podoviridae* (~750 observations, soit 14 % des espèces observées). Ils sont dotés d'une queue courte non contractile. Cette famille comporte, entre autres, les phages T7, P22 et  $\Phi$ 29.



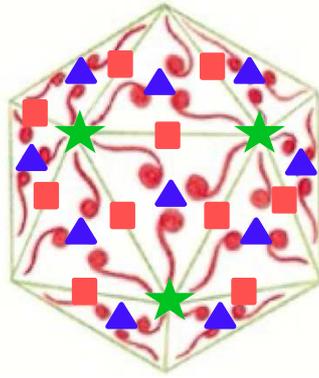
**Figure 5. Morphologie des phages caudés.** Schéma illustrant la morphologie des phages des différentes familles de l'ordre des *Caudovirales*. De gauche à droite : les phages T4 (*Myoviridae*, queue contractile), T5 (*Siphoviridae*, queue non contractile) et T7 (*Podoviridae*, queue courte).

## 5-3 La Capside des phages caudés

### 5-3-1 La symétrie icosaédrique

Les capsides des phages caudés sont constituées de multiples copies d'une (ou de deux pour le phage T4) protéine majeure de tête, qui forment leur architecture icosaédrique, d'un homo-dodécamère ayant une forme d'anneau formant le port d'entrée et de sortie de l'ADN, appelé portale, et parfois de protéines additionnelles, appelées protéines accessoires. L'organisation à symétrie icosaédrique de la capside est partagée par de nombreux virus d'Eucaryotes et d'Archaea : il s'agit du type de structure qui permet, avec un seul type de protéine, de maximiser le rapport volume sur surface, donc d'avoir le plus d'espace possible pour l'ADN en utilisant le moins possible de protéines.

Un icosaèdre (figure 6) est une figure géométrique dont la forme est proche d'une sphère. Il est composé de 20 triangles équilatéraux identiques et 12 sommets. Il possède un ensemble d'axes de symétrie par rotation passant par son centre : 6 axes d'ordre 5, 10 axes d'ordre 3 et 15 axes d'ordre 2. Un axe d'ordre 5 (respectivement 3) signifie que la rotation d'un angle  $\frac{2\pi}{5}$  (respectivement  $\frac{2\pi}{3}$ ) laisse la figure inchangée. Les axes d'ordre 5 sont situés à chaque sommet, les axes d'ordre 3 au centre des 20 faces triangulaires et les axes d'ordre 2 au milieu des 30 arêtes. Un icosaèdre est constitué de 60 unités asymétriques. A partir d'une unité asymétrique et en appliquant les symétries précédemment citées, on construit l'icosaèdre entier (figure 6).



**Figure 6. Représentation simplifiée d'une capsidie icosaédrique.** Le motif  représente une unité asymétrique, qui est, dans les virus de triangulation  $T=1$ , la protéine majeure de capsidie. Les axes d'ordre 5 sont représentés par une étoile verte, ceux d'ordre 3 par un triangle bleu, ceux d'ordre 2 par un carré rouge.

### 5-3-2 Le nombre de triangulation

La plus simple capsidie à symétrie icosaédrique est obtenue avec 60 copies de la protéine majeure de tête, qui est l'unité asymétrique de l'icosaèdre. Cette protéine est alors organisée en 12 pentamères, situés aux sommets de l'icosaèdre. Ces pentamères sont en tous points équivalents, tant chimiquement que par rapport à leurs interactions respectives. Les protéines majeures de capsidie de tous les phages et virus à capsidie icosaédrique sont constituées d'un nombre comparable d'acides aminés, environ 300. Les capsides de plus grande taille sont donc composées de plus de protéines majeures. L'unité asymétrique est alors composée d'un nombre de protéine supérieur à 1. Ce nombre est appelé le nombre de triangulation (Caspar and Klug 1962). La taille des capsides est donc caractérisée par le nombre de triangulation  $T$ . Ce nombre indique également le nombre d'hexamères entre deux pentamères à l'aide de la formule :

$$T = h^2 + k^2 + h * k$$

où  $h$  et  $k$  sont deux entiers positifs et non nuls, définis sur la figure 7. En effet, en plus des pentamères formant les angles sommitaux, les capsides dont le nombre de triangulation est supérieur à 1 possèdent des hexamères, plans, formant les faces de l'icosaèdre. Dans ce cas, les protéines majeures d'une unité asymétrique sont chimiquement identiques mais ne forment pas exactement les mêmes interactions avec les autres protéines majeures. Par exemple une protéine majeure située dans un pentamère interagit avec 6 autres protéines majeures, alors qu'une protéine majeure située dans un hexamère interagit avec 7 autres protéines majeures. Cette légère variation de l'environnement physico-chimique des protéines majeures de capsid est appelée quasi-équivalence (Caspar and Klug 1962).

### 5-3-3 Cas des capsides oblongues

Parmi les 5530 phages caudés étudiés par microscopie électronique, 780 ont une capsid allongée, soit environ 15 % (Ackermann 2007). On parle alors de symétrie pseudo-icosaédrique. Dans ce cas, il n'y a pas un mais deux nombres de triangulation. En effet, les 20 faces d'un icosaèdre peuvent se décomposer en 10 faces formant les extrémités et 10 faces formant la section intermédiaire. Le premier nombre de triangulation correspond à la section terminale de la capsid. Comme pour les capsides icosaédriques, il vaut :

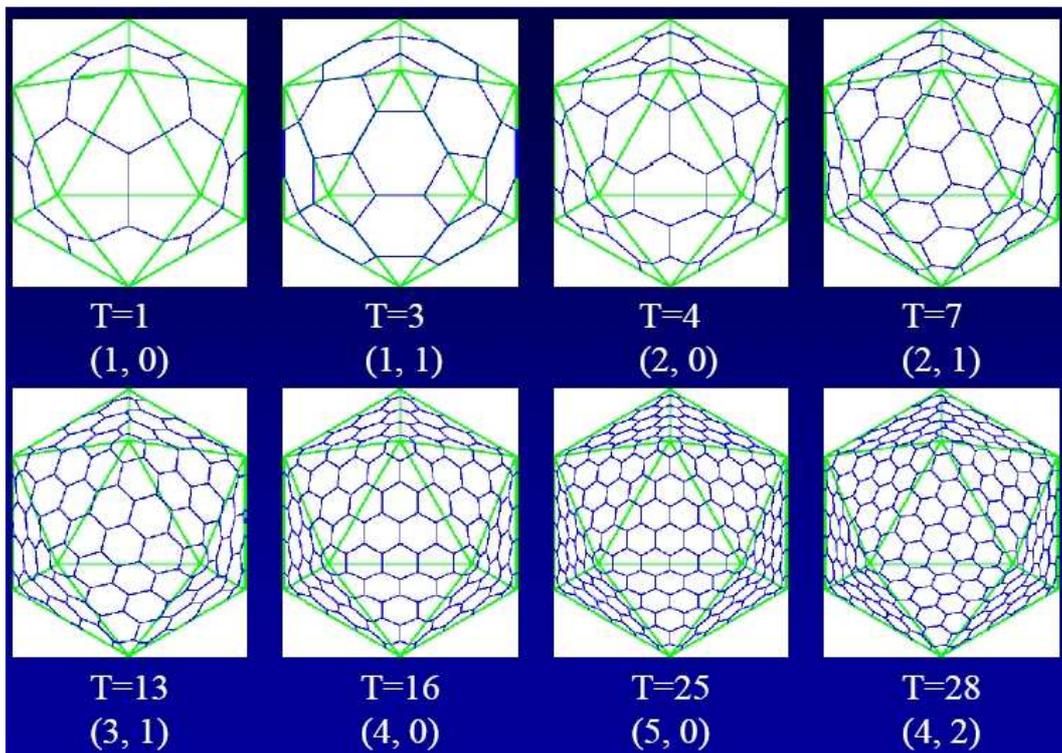
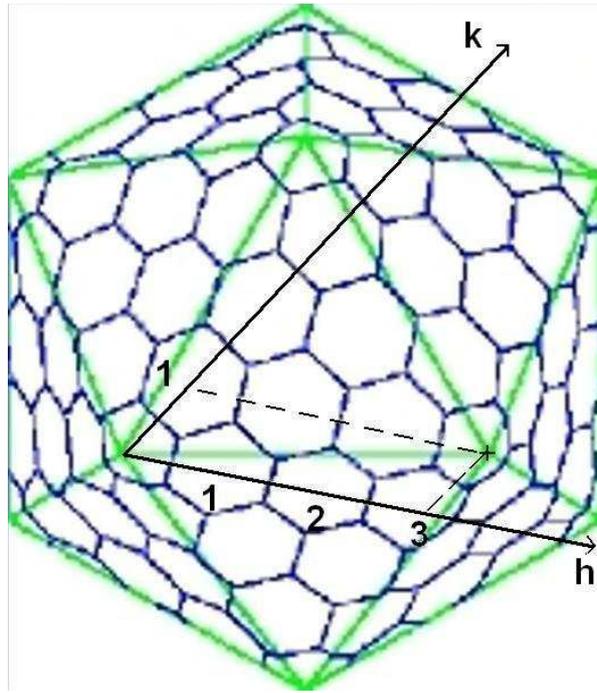
$$T = h^2 + k^2 + h * k$$

Le deuxième correspond à la section intermédiaire. Il se définit comme étant égal à :

$$Q = h_1 h_2 + h_1 k_2 + k_1 k_2$$

Le nombre de protéines majeures de capsid  $N$  se déduit des deux nombres de triangulation :

$$N = 30(T + Q)$$



**Figure 7. Le nombre de triangulation T.** (a) Exemple d'icosaèdre avec un nombre de triangulation égal à 13. 2 pentamères sont au minimum séparés par 3 hexamères dans l'axe des  $h$  et 1 hexamère dans l'axe des  $k$ , d'où son nombre de triangulation. (b) Quelques exemples d'icosaèdres avec différents nombres de triangulation. Entre parenthèse est indiqué le couple  $(h,k)$ .

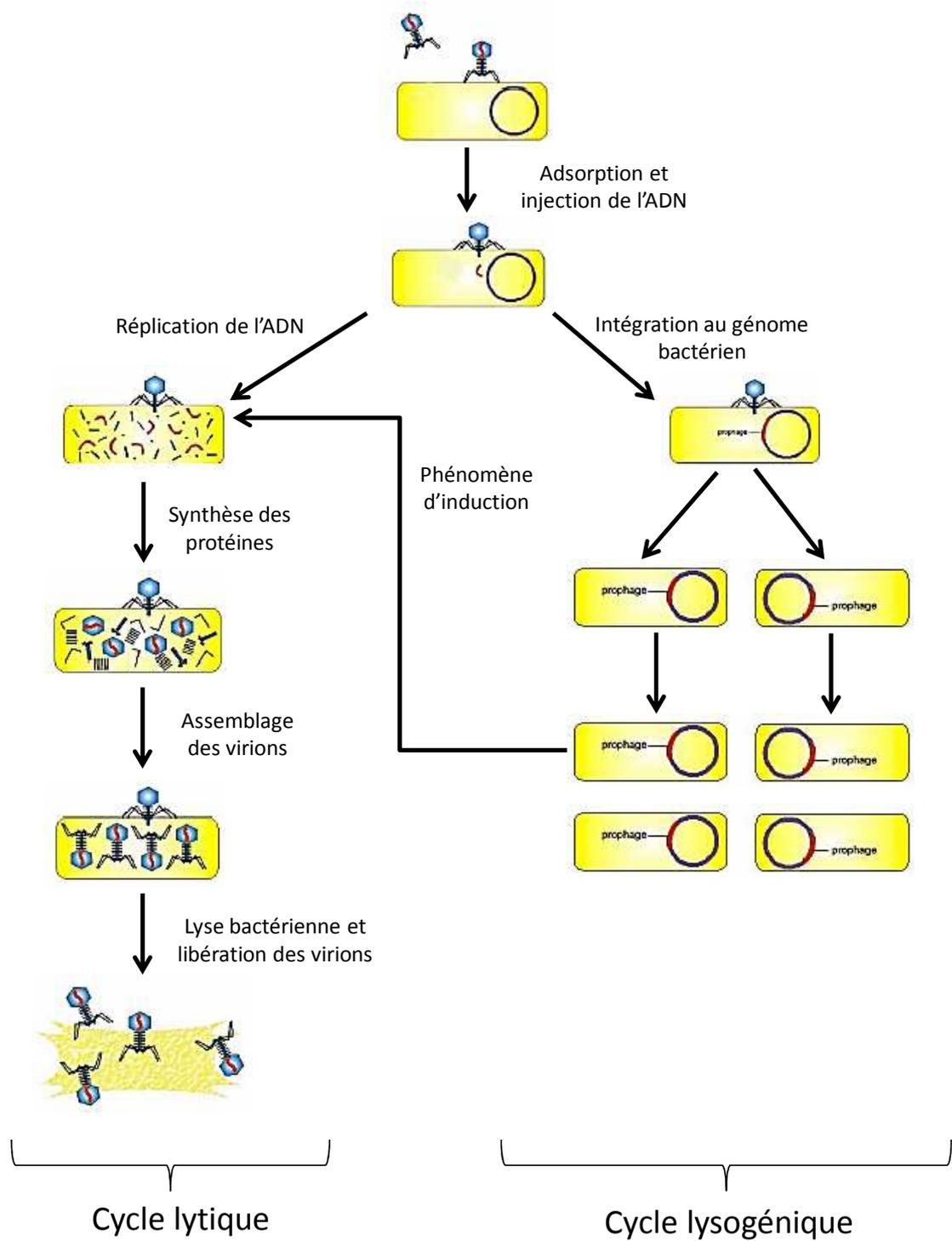
## 6- Assemblage et maturation de la capsid

### 6-1 Cycle de vie du phage

Le cycle infectieux commence par la fixation du bactériophage sur son récepteur membranaire, puis par l'éjection de l'ADN de la capsid et son transfert dans le cytoplasme bactérien. À ce stade, on distingue deux types de bactériophages : les phages lytiques et les phages tempérés. Les phages lytiques n'effectuent que le cycle lytique, alors que les phages tempérés, suivant les conditions d'infection, peuvent effectuer les cycles lytiques ou lysogéniques (figure 8) (Dodd, Shearwin et al. 2005).

Lors d'un cycle lytique la machinerie biosynthétique de la bactérie est détournée au profit de l'expression du génome viral, permettant sa réplication et la synthèse des protéines virales puis l'assemblage de nouvelles particules infectieuses. La libération des virions s'effectue ensuite par lyse bactérienne.

Lors d'un cycle lysogénique, le génome viral est intégré au génome bactérien : il est alors appelé prophage. Ce prophage est incapable de se répliquer de manière autonome. Une protéine appelée répresseur empêche la transcription des gènes viraux. Le prophage se réplique en même temps que le génome bactérien. Cependant, sous certaines conditions de stress, le génome viral est excisé et « sort » du chromosome bactérien, ce qui enclenche le cycle lytique (Oppenheim, Kobilier et al. 2005).



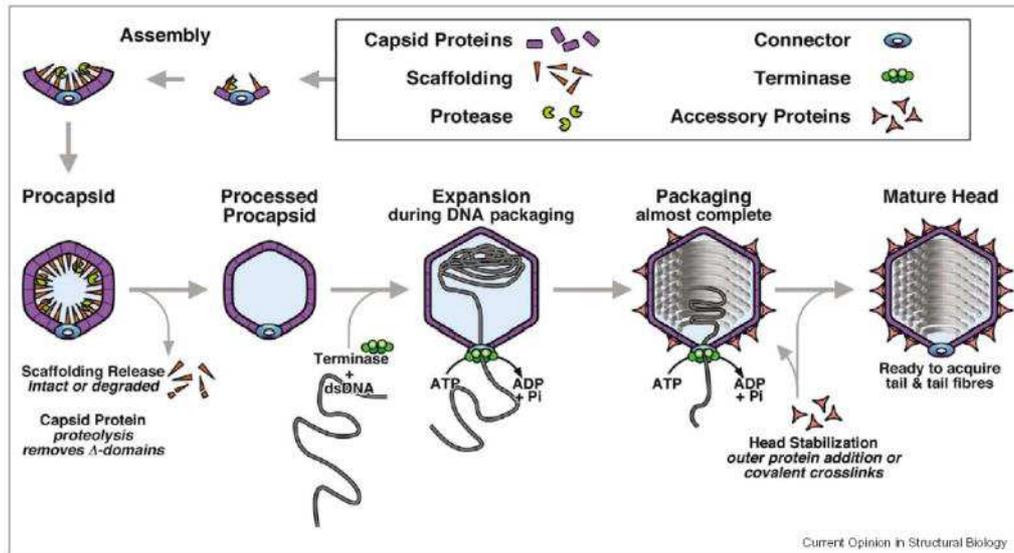
**Figure 8. Schéma général du cycle de vie d'un phage.** Les phages tempérés peuvent effectuer les cycles lytiques et lysogènes, au contraire des phages lytiques qui n'effectuent que le cycle lytique.

## 6-2 Assemblage du virion

Le processus d'assemblage d'un virion infectieux est en grande partie similaire pour tous les phages de l'ordre des *Caudovirales* (figure 9) (Aksyuk and Rossmann 2011). Il implique plusieurs processus :

- la formation d'une proto-capside vide d'ADN, appelée procapside.
- le retrait de la protéine ou du domaine « d'échafaudage » ayant permis la formation de cette procapside
- l'accroissement du volume de la capside par un réarrangement conformationnel des protéines majeures de capside. Ce changement structural est appelé expansion.
- l'entrée de l'ADN dans la capside, ce processus étant concomitant avec l'expansion de la capside
- un éventuel renforcement de la stabilité de la capside, par ajout de protéines accessoires ou par pontage covalent entre les capsomères. Chez certains phages il y a également fixation de protéines appelées protéines de décoration. Le rôle de ces protéines n'est pas connu.
- la fermeture de la capside par une ou plusieurs protéines appelées protéines de fermeture.

Ensuite, chez les *Podoviridae*, les protéines de queue sont attachées à la capside de manière séquentielle (Camacho, Jimenez et al. 1979; Lander, Khayat et al. 2009). Chez les *Siphoviridae* et les *Myoviridae* la queue est assemblée de manière indépendante puis attachée à la capside, au niveau de la portale.



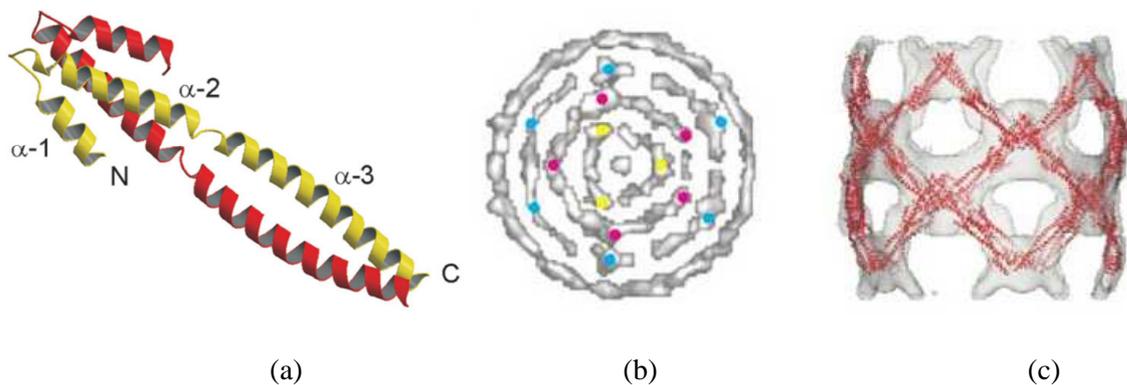
**Figure 9. Schéma général des étapes de l'assemblage des capsides de phage caudé ainsi que les protéines impliquées.** Figure adaptée de (Steven, Heymann et al. 2005).

### 6-2-1 Formation de la procapside

La procapside est composée de la portale, de multiples copies de la protéine majeure de capsid ainsi que de multiples copies d'une protéine appelée protéine d'échafaudage. Chez la plupart des phages, les sous-unités de la protéine majeure de capsid ne peuvent pas s'auto-assembler en capsid à la géométrie adéquate, sans une protéine d'échafaudage nécessaire pour « monter » cet assemblage (Fane and Prevelige 2003). En effet, pour le phage  $\phi 29$ , une mutation de la protéine d'échafaudage conduit à la formation de capsides isométriques, et non allongées, qui sont donc trop petites pour encapsider l'ensemble du génome phagique (Choi, Morais et al. 2006). Chez le phage T4, l'absence de cette protéine conduit à la formation de structures cylindriques de grande dimension, non viable pour la suite de l'assemblage du virion (Kellenberger 1990). Pour le phage P22 cette absence induit d'une part un assemblage de particules aberrantes et d'autre part une vitesse d'assemblage beaucoup plus faible (Casjens and King 1974). Il a été récemment montré que les interactions entre les protéines majeures étaient plus faibles dans la procapside de P22 que dans ces particules aberrantes. La

protéine d'échafaudage, en interagissant sur plusieurs sites de fixation de la protéine majeure, permet de stabiliser la forme procapside (Zlotnick, Suhanovsky et al. 2012). La structure de la protéine d'échafaudage du phage  $\phi 29$  a été déterminée par cristallographie (Morais, Kanamaru et al. 2003). Elle indique que cette protéine est un homodimère structuré en coiled-coil qui interagit avec la protéine majeure de tête *via* des interactions ioniques (figure 10).

Chez les phages T5 et HK97 la protéine d'échafaudage n'est pas codée par un gène séparé des autres gènes de capsid, mais cette fonction étant assurée par une extension N-terminale de la protéine majeure de capsid appelé domaine d'échafaudage (Duda, Martincic et al. 1995; Huet, Conway et al. 2010).



**Figure 10. Echafaudage de la procapside du phage  $\phi 29$ .** (a) Structure cristalline du dimère de la protéine gp7, protéine d'échafaudage de  $\phi 29$ . (b) Coupe équatoriale de la reconstruction à partir d'images de cryo-microscopie électronique de la procapside de  $\phi 29$ . Cette coupe montre l'échafaudage de la procapside formant des cercles concentriques. (c) En gris : coupe axiale de cette reconstruction. En rouge : structure cristalline de gp7 inséré dans cette reconstruction.

Figure adaptée de (Morais, Kanamaru et al. 2003).

Les mécanismes par lesquels la protéine d'échafaudage agit sur la forme de la procapside sont malgré tout mal connus, de même que les étapes menant à la formation de la procapside. Il est probable que l'assemblage soit initié depuis la portale (Valpuesta and

Carrascosa 1994). Cette structure, par laquelle l'ADN entre et sort de la capsid, est formée par douze copies de la protéine portale (Johnson and Chiu 2007). Chez T4, elle interagit avec la membrane bactérienne via une protéine échafaudage supplémentaire, et l'assemblage s'effectue à partir de ce point de nucléation (Michaud, Zachary et al. 1989). Cependant, cet ancrage à la membrane des procapsides durant leur assemblage n'a été démontré chez aucun autre phage. De plus, certains phages, comme P22, T7, P4 (Dokland, Wang et al. 2002) et HK97 (Duda, Martincic et al. 1995) peuvent former des procapsides sans la présence de la portale. Mais bien sûr, ces procapsides ne peuvent pas encapsider l'ADN et donc former des virions infectieux.

### **6-2-2 Maturation de la capsid**

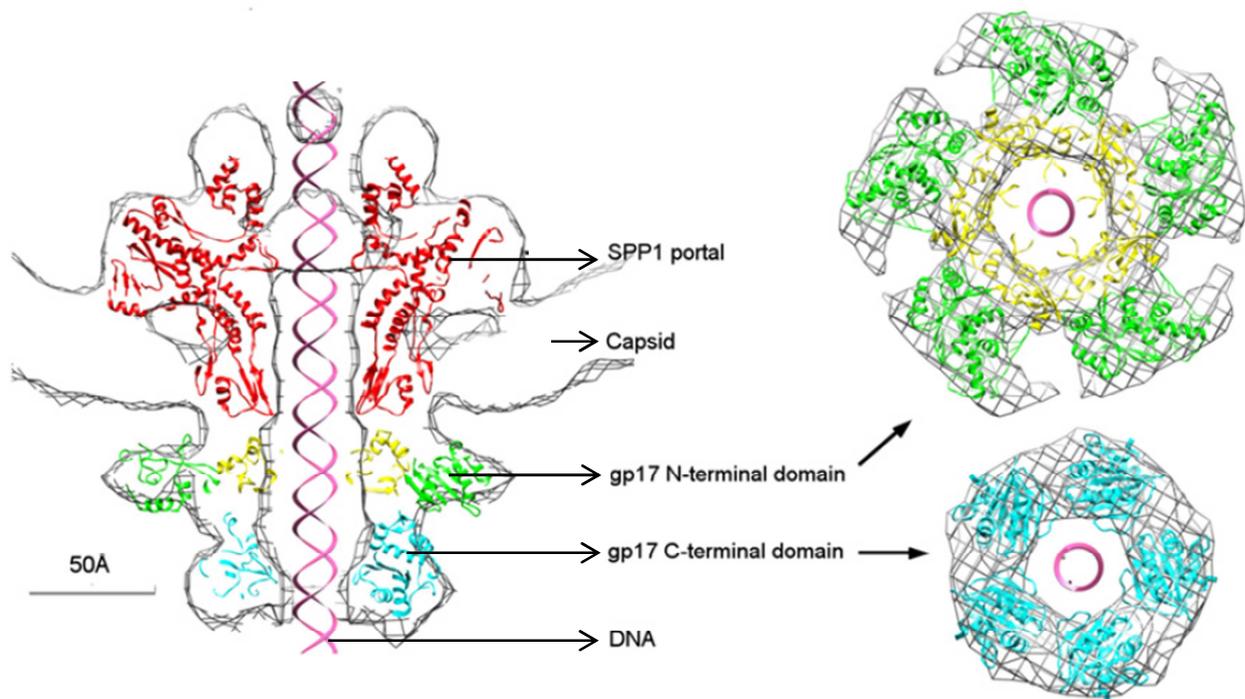
Après la formation de la procapsid, les domaines ou protéines échafaudages sont éliminés de la capsid. Dans le cas des bactériophages  $\phi 29$  (Bjornsti, Reilly et al. 1983), T7 (Cerritelli, Conway et al. 2003) et P22 ces protéines sortent intactes de la capsid et chez P22 elles peuvent être réutilisées plusieurs fois pour l'assemblage d'autres procapsides (King and Casjens 1974). Chez le phage  $\phi 29$ , la sortie de cette protéine pourrait être induite par ses interactions avec l'ADN, et serait donc concomitante avec l'encapsidation de celui-ci (Morais, Kanamaru et al. 2003). Chez les phages HK97, P2,  $\lambda$ , T4 et T5 les domaines ou protéines échafaudages sont découpées par une protéase du phage et éliminées avant l'encapsidation du génome et l'expansion de la capsid (Aksyuk and Rossmann 2011).

Le transfert de l'ADN dans la capsid est assuré par un complexe enzymatique appelé terminase. Elle comprend deux sous-unités : une sous-unité appelée petite terminase qui reconnaît l'ADN à encapsider et d'une sous-unité catalytique appelée grande terminase dont l'activité ATPase fournit l'énergie pour transporter l'ADN dans la capsid à travers la portale.

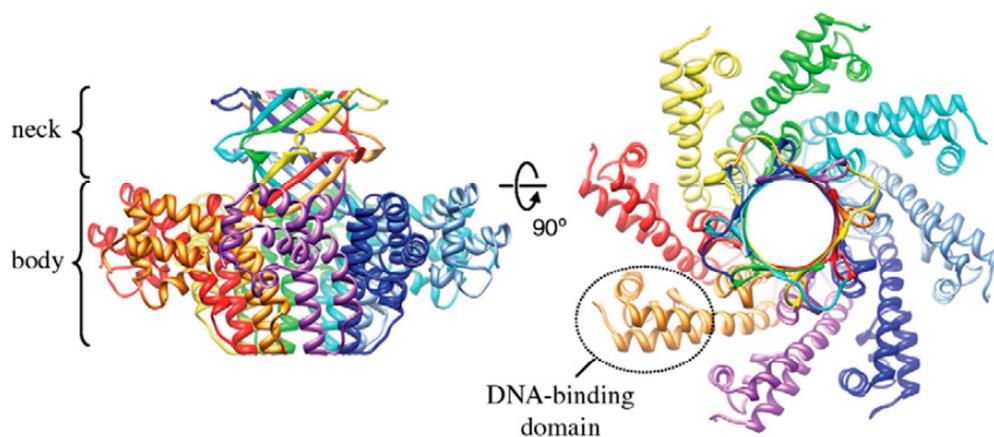
La force produite par les terminases des phages  $\phi 29$ ,  $\lambda$  et T4, mesurée par pinces optiques, est supérieure à 60 pN ; la vitesse d'encapsidation *in vitro* quant à elle varie selon la quantité de génome à encapsider, de 180 ( $\phi 29$ , taille du génome : 19,3 kbp ) à 1800 (T4, taille du génome : 168,9 kbp) paires de bases par seconde, permettant une encapsidation du génome entier en 2 à 3 minutes (Casjens 2011; Smith 2011). La grande terminase possède également une activité DNase, qui clive les concatémères d'ADN produits par la réplication de l'ADN viral, permettant l'entrée d'une unité de génome dans chaque capsid (Rao and Feiss 2008).

La structure atomique de la protéine constituant la grande terminase du phage T4, gp17, a été déterminé par cristallographie et la reconstruction par cryo-microscopie électronique de la procapsid de T4 complexé avec la terminase montre que cette dernière est un pentamère de gp17 figure 11 (a). La structure de la petite terminase d'un phage de la famille des *Podoviridae*, sf6, a également été résolue par cristallographie (figure 11(b)) (Zhao, Finch et al. 2010). Elle suggère qu'elle forme un anneau situé en dessous de la grande terminase, qui elle-même se fixe à la portale.

L'étape d'expansion consiste en un réarrangement structural de la capsid qui se traduit par un accroissement d'environ 50% du volume intérieur de la capsid. Les détails de cette étape, qui est nécessaire à l'encapsidation de tout le génome du phage, sont développés dans la section 6-3.



(a)



(b)

**Figure 11. Structures cristallographiques d'une petite et d'une grande terminase.** (a) Les structures cristallographiques de la portales du phage SPP1 (en rouge) et de la grande terminase du phage T4 (domaine N-terminal en vert et jaune, domaine C-terminal en cyan) sont insérées dans la carte de densité électronique du complexe procapside-grande terminase. Cette carte de densité électronique a été obtenue par cryo-microscopie. La grande terminase correspond à un pentamère de la protéine gp17. En haut à droite, vu en coupe du domaine N-terminal de ces protéines, ce domaine ATPase fournit l'énergie nécessaire à l'encapsidation. En bas à droite, le domaine C-terminale permet de couper l'ADN concamérique une fois l'encapsidation du génome complet achevé. (b) Structure cristallographique (résolution : 1,65 Å) de la petite terminase du phage sf6, un phage appartenant à la famille des *Podoviridae*. Les huit monomères de la protéine gp1 sont coloriés avec une couleur différente. Cette structure montre un anneau dont le centre forme un tunnel par lequel pourrait traverser l'ADN. (a) : figure adaptée de : (Sun, Kondabagil et al. 2008) ; (b) : figure adaptée de (Zhao, Finch et al. 2010).

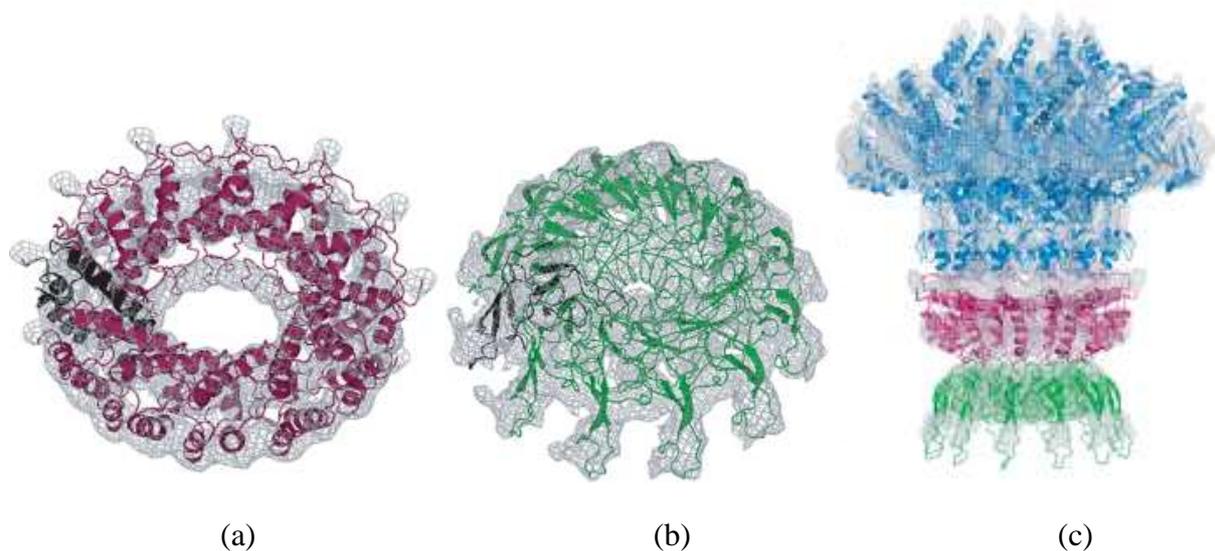
Après les étapes d'encapsidation de l'ADN et d'expansion, il existe chez certains phages un processus de stabilisation de la capsid. Chez HK97, cette stabilisation est effectuée par un pontage auto-catalytique qui forme des liaisons covalentes entre les protéines des capsomères adjacents, générant un réseau de type « *chain mail* » ou cote de mailles, indispensable à la viabilité du phage. Chez les phages T4, L et  $\lambda$ , les protéines SOC, DEC et gpD participent à sa stabilité en se fixant aux interfaces entre trois hexamères et en les liant ensemble (figure 34). Cependant, si la présence de gpD est indispensable pour l'encapsidation de tout le génome de  $\lambda$  (Fuller, Raymer et al. 2007), les protéines SOC et DEC ne sont indispensables que pour la stabilité des particules virales à des pHs et des températures extrêmes.

Certains phages fixent, à la surface de leur capsid, des protéines appelées protéines de décoration, qui ne sont pas indispensables à l'infectiosité et pour lesquelles aucun effet stabilisateur n'a été mis en évidence. La décoration des capsides est traitée dans la troisième partie de cette thèse.

### **6-2-3 Etapes finales de l'assemblage du virion**

La dernière étape de l'assemblage de la capsid est sa fermeture par des protéines spécifiques qui constituent également le site de fixation de la queue, terminant ainsi l'assemblage du virion infectieux. Le cas le mieux décrit est celui du phage SPP1 pour lequel deux homo-dodécamères appelés connecteur et bouchon (Orlova, Gowen et al. 2003; White, Sherman et al. 2012) . Ces deux dodécamères forment avec la portale un complexe appelé "barrière" (*note : « gatekeeper » en anglais*), qui permet d'empêcher la sortie prématurée de l'ADN hors de la capsid (figure 12). La capsid phage P22 est, elle, fermée par deux

protéines connecteurs (gp4 et gp10) et une protéine bouchon (gp26) (Olia, Bhardwaj et al. 2007).



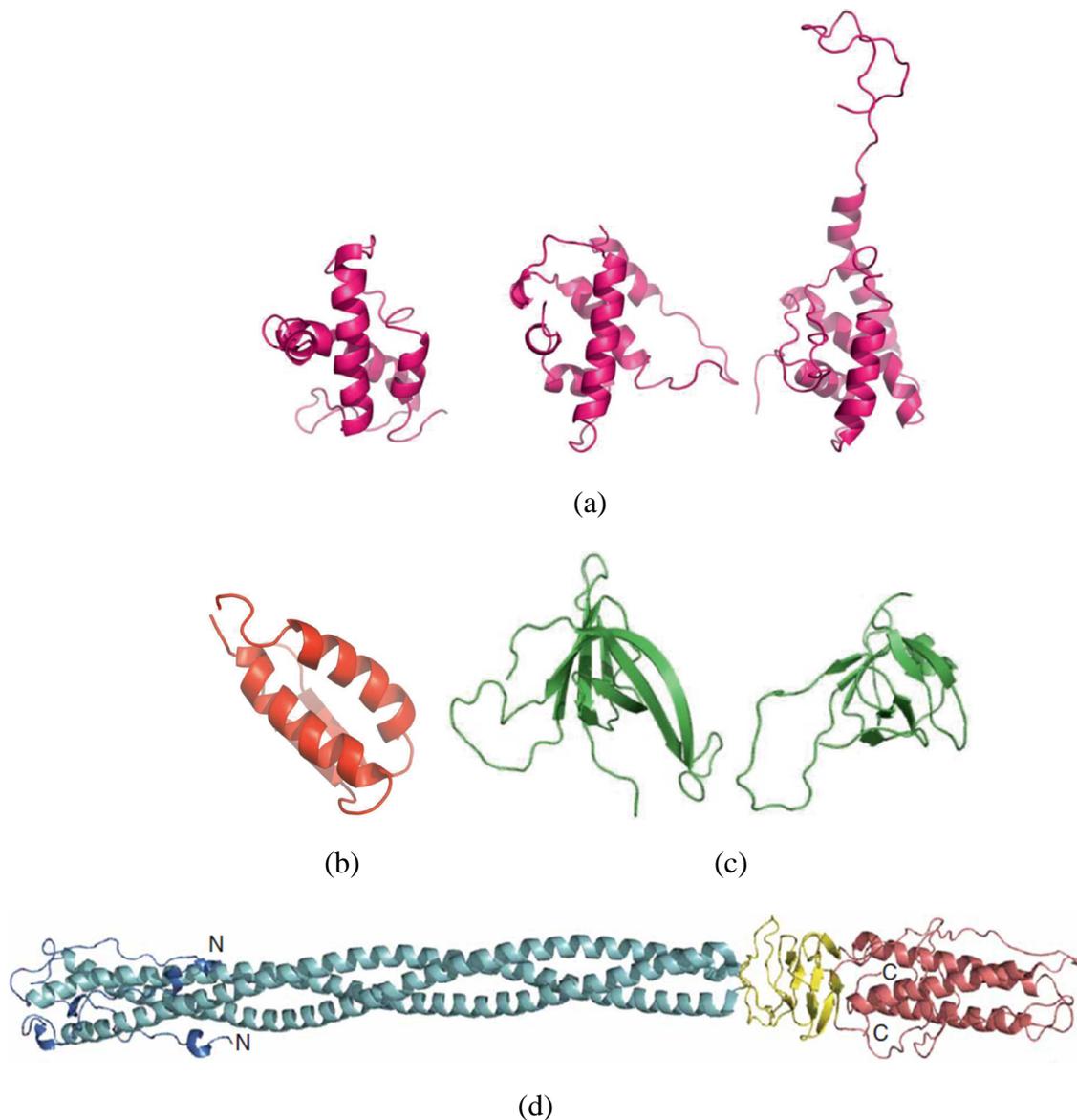
**Figure 12. Représentation des protéines formant la portale, le connecteur et le bouchon du phage SPP1.** (a) Représentation de la protéine gp15 formant un anneau dodécamérique, le connecteur. La structure de la protéine gp15 seule a été obtenue par RMN. Ce modèle du connecteur est inséré et contraint dans sa carte de densité électronique obtenue par cryo-microscopie électronique. (b) Représentation de la protéine gp16 formant un anneau dodécamérique, le bouchon. La structure de la protéine gp16 seule a été obtenue par RMN. Ce modèle du connecteur est inséré et contraint dans sa carte de densité électronique obtenue par cryo-microscopie électronique. (c) Vue d'ensemble des structures pseudo-atomiques de la portale (en bleu), du connecteur (en magenta) et du bouchon (en vert) obtenus par RMN et contraints séparément dans leur carte de densité électronique correspondante. Figure adaptée de (Lhuillier, Gallopain et al. 2009).

Il existe 4 familles de protéines de fermeture ayant des repliements similaires. La première comprend la protéine gp15 du phage SPP1, la protéine gp6 du phage HK97 et la protéine gp4. Les 3 protéines connues de cette famille sont des connecteurs constituées de 3 ou 4 hélices  $\alpha$  entourant un cœur hydrophobe (figure 13 (a)). La deuxième famille est composée de petites protéines analogues à la protéine gpW, protéine bouchon du phage  $\lambda$  (Cardarelli, Pell et al. 2010). Cette protéine est formée de deux hélices  $\alpha$  reliées par un brin  $\beta$

(figure 13 (b)) (Maxwell, Yee et al. 2001). Les protéines de la troisième famille partagent un repliement en tonneau  $\beta$  ainsi que de grandes boucles non structurées et très mobiles (figure 13 (c)). Au contraire de la protéine gp16 qui est le bouchon de la capsid de SPP1 (Lhuillier, Gallopin et al. 2009), la protéine gpFII n'est pas nécessaire pour éviter la sortie de l'ADN de la capsid du phage  $\lambda$ . La fonction putative de cette protéine qui se fixe sur la capsid pleine de  $\lambda$  après la protéine gpW est de former une interface nécessaire à la fixation de la queue (Cardarelli, Pell et al. 2010). La quatrième famille de protéines de fermeture est composée de protéines de phage à queue courte. Ces protéines s'organisent en trimères pour former de fines aiguilles qui ont deux fonctions : celle de fermer la capsid et celle de percer l'enveloppe bactérienne de l'hôte et ainsi de former un conduit pour le passage de l'ADN phagique vers le cytoplasme bactérien (Bhardwaj, Walker-Kopp et al. 2009). La protéine gp26 du phage P22 est l'exemple le mieux caractérisé de cette famille de protéines ; elle est formée par deux hélices  $\alpha$  surenroulées (*coiled-coil*) séparées par une hélice  $\beta$  et mesure 240 Å de long (figure 13 (d)).

### **6-3 L'expansion de la capsid des phages caudés**

L'étape d'expansion a été étudiée chez des phages appartenant aux trois familles des Caudovirales : chez HK97 (Hendrix and Johnson 2012), et  $\lambda$  (Dokland and Murialdo 1993) (*siphoviridae*), chez T4 (Ray, Oram et al. 2009) (*myoviridae*) et chez P22 (Jiang, Li et al. 2003) (*podoviridae*). C'est une étape vraisemblablement incontournable dans la formation des capsides de tous les phages caudés.



**Figure 13. Structures atomiques de protéines appartenant aux 4 familles de protéines de fermeture.** (a) Structures atomiques des monomères des protéines gp15 (gauche), gp6 (centre) et gp4 (droite) des phages SPP1, HK97 et P22. La structure de gp15 a été obtenue par RMN (Cf. figure 12) (Lhuillier, Gallopin et al. 2009), les structures de gp6 et gp4 ont été déterminées par cristallographie (Cardarelli, Lam et al. 2010) (Olia, Prevelige et al. 2011). (b) Structure atomique du monomère de la protéine gpW du phage  $\lambda$ , obtenue par RMN (Maxwell, Yee et al. 2001). (c) Structures atomiques des monomères des protéines gp16 (gauche) et gpFII (droite) des phages SPP1 et  $\lambda$ . Ces structures ont été obtenues par RMN (Lhuillier, Gallopin et al. 2009) (Maxwell, Yee et al. 2002). (d) Structure cristallographique du trimère de gp26, formant le bouchon de la capside du phage P22 ainsi que l'aiguille permettant de percer l'enveloppe bactérienne de *Salmonella enterica* lors de son infection (Olia, Casjens et al. 2007). Codes d'accèsion pdb : gp15 (SPP1) : 2KBZ, gp6 (HK97) : 3JVO, gp4 (P22) : 3LJ4, gpW ( $\lambda$ ) : 1HYW, gp16 (SPP1) : 2KCA, gpFII ( $\lambda$ ) : 1K0H, gp26 (P22) : 2POH. Figure adaptée de (Tavares, Zinn-Justin et al. 2012).

De nombreuses études cinétiques et structurales du phénomène d'expansion ont été réalisées *in vitro* sur le système modèle de la capsidite du phage HK97. Cette procapsidite modèle est uniquement constituée de la protéine majeure de capsidite gp5 et elle ne contient pas la protéine portante. Elle est obtenue par co-expression chez *E. coli* de gp5 et de la protéase de maturation qui clive le domaine d'échafaudage. Cette procapsidite est appelée P-2. L'expansion peut être déclenchée *in vitro* par de nombreux traitements chimiques (urée, glycérol, chloroforme...) (Duda, Hempel et al. 1995), et physico-chimiques, en particulier l'acidification à pH 4 (Hendrix and Duda 1998). Elle peut également être induite par chauffage sur la procapsidite du phage P22 (Parker and Prevelige 1998).

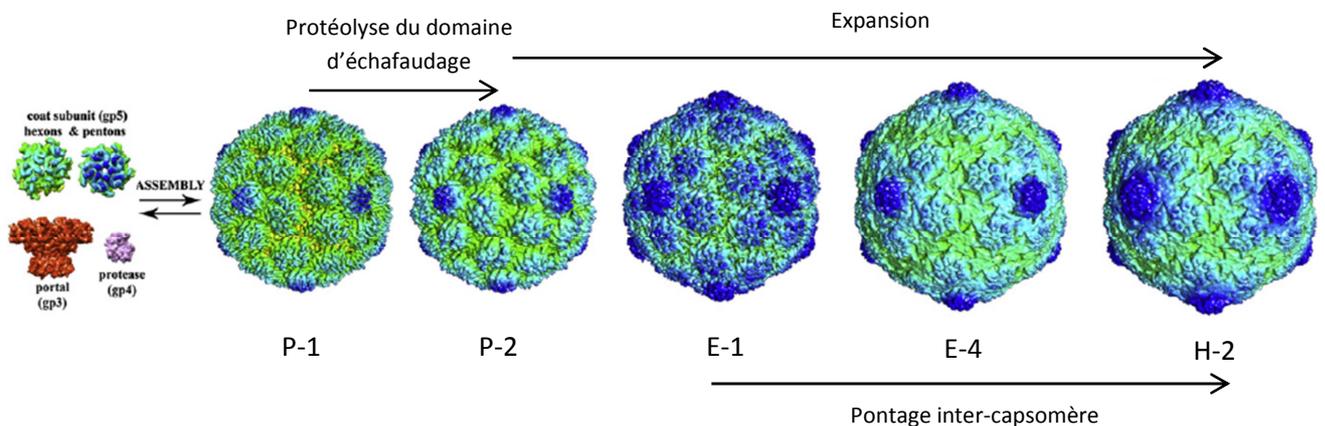
### **6-3-1 Déclenchement de l'expansion et lien avec l'encapsidation de l'ADN**

Des expériences d'encapsidation effectuées *in vitro* avec le phage  $\lambda$  montrent une forte baisse de la vitesse d'encapsidation lorsqu'environ 30% du génome est encapsidé (Fuller, Raymer et al. 2007). Cette baisse de vitesse a été interprétée comme une conséquence de l'expansion de la procapsidite. Cette interprétation est étayée par une étude antérieure, qui a montré par microscopie électronique que l'expansion intervient quand 10 à 50% de l'ADN phagique est encapsidé (Murialdo 1991). L'expansion serait donc initiée par l'encapsidation de l'ADN. La structure de la procapsidite après clivage du domaine d'échafaudage du phage HK97 (Prohead II, ou P-2) a été résolue par cristallographie (Gertsman, Gan et al. 2009). Elle montre que la surface interne de la procapsidite est composée majoritairement d'acides aminés acides, ce qui crée une charge négative globale. L'entrée dans la procapsidite de l'ADN, chargé négativement, pourrait très vraisemblablement créer une perturbation qui induirait son expansion. Par ailleurs, il a été montré que l'expansion *in vitro* de la procapsidite de P22 était inhibée par une forte concentration en chlorure de sodium, ce qui suggère que cette expansion est induite par des interactions électrostatiques (Parker and Prevelige 1998). De plus, *in vivo*,

l'initiation de l'encapsidation de l'ADN de T4 est nécessaire et suffisante pour provoquer l'expansion des capsides (Jardine, McCormick et al. 1998). Dans l'état actuel des connaissances, l'hypothèse la plus probable est donc que le phénomène d'expansion soit provoqué par la répulsion électrostatique entre l'ADN et les acides aminés tapissant l'intérieur de la capside.

L'hypothèse que l'expansion pourrait fournir une partie de l'énergie nécessaire pour encapsider l'ADN a également été proposée (Serwer 2003). Cependant, les capsides expansées du phage T4 sont compétentes pour l'encapsidation *in vitro*, cette étape n'aurait donc pas besoin d'un éventuel apport énergétique venant de l'expansion (Jardine, McCormick et al. 1998).

### 6-3-2 Les étapes d'expansion du phage HK97



**Figure 14. Vue d'ensemble de l'assemblage et de l'expansion de la capside du phage HK97.** Les différents intermédiaires dans la formation de la capside expansée (H-2, à droite) à partir des protéines isolées (à gauche) sont, dans l'ordre : la procapside avec (P-1) et sans (P-2) le domaine d'échafaudage, le premier état intermédiaire de l'expansion (E-1) et l'état dit « ballon » (E-4). L'étape d'expansion permet de faire passer le diamètre de la capside de 540 Å (P-2) à 660 Å (H-2), soit un doublement de volume. Figure adaptée de (Veesler, Quispe et al. 2012).

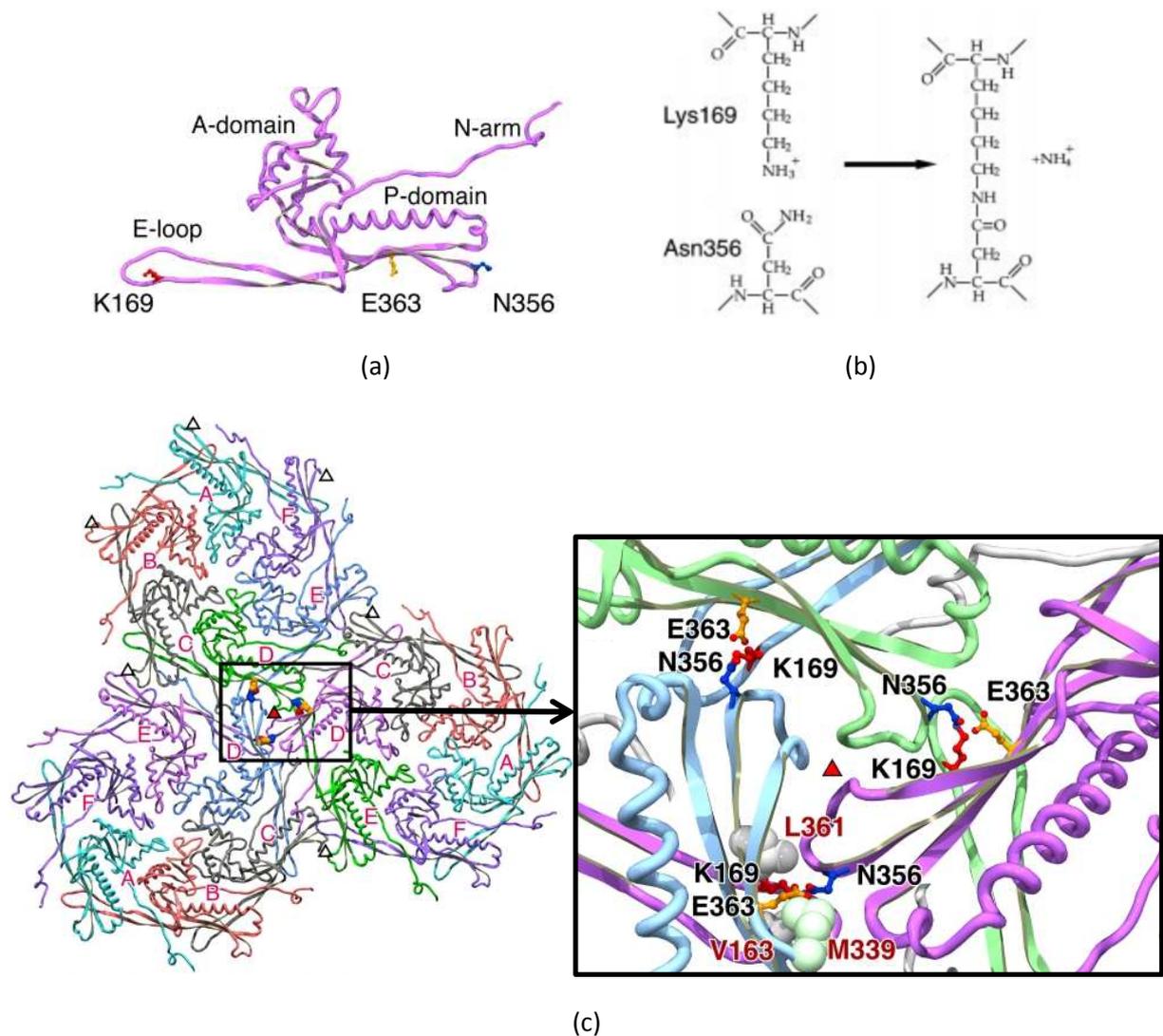
Le processus de maturation *in vitro* de la capsidite du phage HK97 est particulièrement bien décrit. La première transition après acidification transforme la procapsidite P-2 en un état intermédiaire appelé EI-1 (figure 14). Cet état EI-1 est plus grand et plus anguleux que l'état P-2. Des études par diffusion des rayons X aux petits angles (SAXS) ont montré que la transition intervient de manière stochastique sur la population de procapsidites, sans état intermédiaire détectable. A partir de l'état EI-1 s'établit un réseau de pontages covalents entre les protéines des capsomères adjacents (figure 15). A partir d'un nombre critique de pontages formés (~60%), la capsidite est soumise à un changement structural important consistant en une nouvelle augmentation de sa taille et conduisant à une forme arrondie. Cette capsidite est appelée ballon ou EI-4. A ce stade tous les pontages ont été effectués, sauf ceux qui relient les pentamères aux hexamères. *In vitro*, la formation de ces pontages doit être initiée par une neutralisation à pH7. Après formation des derniers pontages, la capsidite acquiert sa forme définitive, elle est appelée H-2.

En plus de la structure de la procapsidite P-2, les structures atomiques des formes EI-4 et H-2 ont été résolues par cristallographie (Gan, Speir et al. 2006). De plus, la structure de l'état intermédiaire E-1 a été récemment déterminée par cryo-microscopie à une résolution de 9.3 Å (Veesler, Quispe et al. 2012). Ces données permettent de visualiser les changements structuraux qui ont lieu durant l'expansion. Au niveau de la protéine majeure de capsidite, il y a peu de changement de structure secondaire, à l'exception notable de torsions de 3 brins  $\beta$  et d'une flexion concomitante au déroulement partiel d'une hélice  $\alpha$  (figure 16 (a)) (Gertsman, Gan et al. 2009).

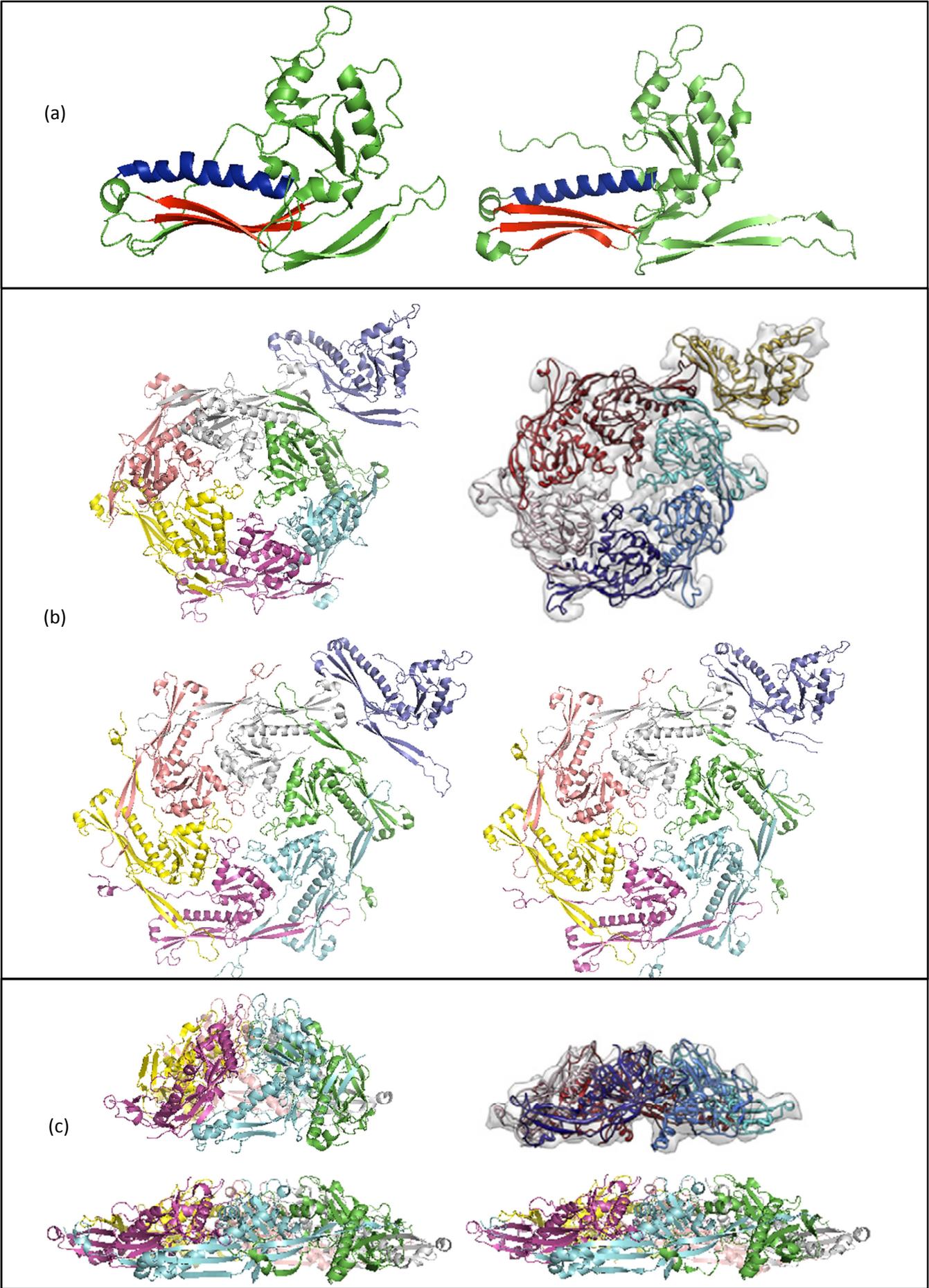
Les changements de structure quaternaire des capsomères sont en revanche saisissants (figure 16 (b) et (c)). Dans la procapsidite P-2 les protéines majeures sont orientées de façon radiale par rapport à la surface de la capsidite, alors que dans la capsidite expansée H-2 ils sont tangentiels à cette surface. Il en résulte que l'épaisseur de H-2 est beaucoup plus faible que

celle de P-2 (20 Å contre 40 Å). Les hexamères de H-2 étant aplatis, ils occupent une surface plus importante que celle de P-2, permettant à la capsidie expansée d'avoir un volume presque deux fois supérieur à celui de la procapsidie. D'autre part, dans la procapsidie P-2 le principe de quasi équivalence n'est pas respecté : les hexamères n'ont qu'un pseudo axe de symétrie d'ordre 2 et sont donc obliques. L'apparition de l'axe de symétrie d'ordre 6 au centre des hexamères s'effectue dès la transition vers l'état intermédiaire EI-1. En revanche les interactions entre les différents hexamères sont remarquablement stables durant toutes les étapes de l'expansion, ce qui participerait à la stabilité des capsides malgré les importants réarrangements structuraux induits par ce processus.

Le processus d'expansion a également été étudié d'un point de vue structural par cryo-microscopie électronique et reconstruction sur les phages  $\lambda$  (Lander, Evilevitch et al. 2008), T7 (Agirrezabala, Velazquez-Muriel et al. 2007; Ionel, Velazquez-Muriel et al. 2011) et P22 (Chen, Baker et al. 2011). Comme pour le phage HK97, ces reconstructions montrent un important amincissement de la capsidie durant l'expansion. Les modèles quasi-atomiques des capsides avant et après expansion ont été déterminés à partir de ces reconstructions et de la structure de la protéine majeure de capsidie de HK97, gp5. Ces modèles indiquent notamment que la symétrisation et l'aplatissement des hexamères durant l'expansion, processus ayant été observés pour le phage HK97, sont également présents pour les phages  $\lambda$ , T7, P22. Cependant, aucune structure cristallographique de capsidie ou de protéine majeure de capsidie n'a pu être obtenue chez d'autres phages caudés que HK97.



**Figure 15. Pontage inter-capsomère de la capsid de HK97.** (a) Structure cristallographique de la protéine majeure de capsid de HK97. Cette structure est issue de la structure de la capsid expansée (H-2). Les acides aminés glutamate E363, lysine K169, et asparagine N356 participent au pontage inter-capsomère ayant lieu pendant l'expansion. Ils sont respectivement coloriés en orange, rouge et bleu. (b) Schéma de la réaction chimique du pontage entre la lysine K169 et l'asparagine N356. Cette réaction est catalysée par le glutamate E363. (c) A gauche : vue d'ensemble de trois capsomères adjacents. Les triangles représentent les quasi-axes de symétries d'ordre 3. A droite, vue zoomée de l'interface entre les trois capsomères, montrant les acides aminés impliqués dans le pontage. (a) et (c) : figure adaptée de (Dierkes, Peebles et al. 2009), (b) : figure adaptée de (Wikoff, Liljas et al. 2000).



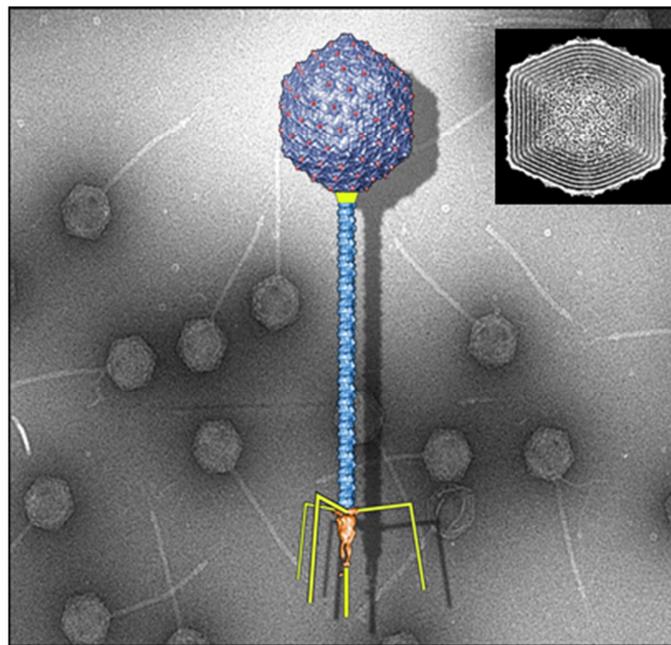
**Figure 16. Changements structuraux de gp5 et des hexamères de gp5 durant l'expansion.** (a) Représentation de la structure atomique de gp5 issue de la procapside (P-2) et de la capsidie expansée (H-2). Les trois brins  $\beta$  subissant une torsion lors de l'expansion sont coloriés en rouge ; l'hélice  $\alpha$  subissant une flexion concomitante à son déroulement partiel est coloriée en bleue (b) Vue d'une unité asymétrique (c'est à dire sept protéines majeures gp5 : six formant un hexamère et une issue d'un pentamère) de la capsidie de HK97 à différentes étapes de l'expansion. De gauche à droite et de haut en bas : procapsidie (P-2), état intermédiaire (EI-1), état ballon (EI-4) et capsidie expansée (H-2). (c) Les mêmes structures sont vues de profil. La structure de l'état intermédiaire EI-1 a été obtenue en contraignant la structure d'une protéine majeure issue de la P-2 dans la carte de densité électronique de l'état intermédiaire. Cette carte de densité a été obtenue par reconstruction à partir d'images de cryo-microscopie électronique (résolution : 9.3 Å). Les deux représentations correspondantes ont été adaptées de (Veesler, Quispe et al. 2012). Les autres structures ont été obtenues par cristallographie (P-2 : (Gertsman, Gan et al. 2009)), EI-4 et H2 : (Gan, Speir et al. 2006)). Figure créée avec le logiciel Pymol, en utilisant les coordonnées de la banque de données PDB (P-2 : 3E8K, EI-4 : 2FRP et H-2 : 2FT1).

## 7- Le bactériophage T5

### 7-1 Présentation

T5 est un phage lytique qui infecte *E. coli*. Il appartient à l'ordre des *Caudovirales*. Il possède une queue longue non contractile ce qui le classe dans la famille des *Syphoviridae* (figures 5 et 17). Cette queue mesure 2500 Å de longueur et est formée par l'empilement de 40 trimères de la protéine majeure de queue, pb6 (Effantin, Boulanger et al. 2006). Cette organisation est atypique parmi les phages dont la queue présente habituellement une symétrie hexamérique. La partie distale de la queue est une structure complexe formée d'une partie conique et d'une fibre centrale au bout de laquelle se situe la protéine pb5, qui se fixe au récepteur bactérien FhuA, induisant l'ancrage du phage dans l'enveloppe bactérienne et l'éjection de l'ADN (Letellier, Boulanger et al. 2004). T5 possède une capsidie icosaédrique d'environ 900 Å de diamètre contenant son génome, un ADN double brin de 121 kpb,

compacté en structure quasi cristalline. Cette capsid, de triangulation  $T=13$ , est composée de 775 copies de la protéine majeure de tête, pb8, réparties en 11 pentamères et 120 hexamères, et de la portale, un oligomère formé de 12 copies de la protéine pb7. La structure des capsides expansées vides et pleines d'ADN a été résolue par cryo-microscopie et reconstruction d'image avec respectivement une résolution de 19 et 20 Å (Effantin, Boulanger et al. 2006) (figure 19 (a) et (b)).

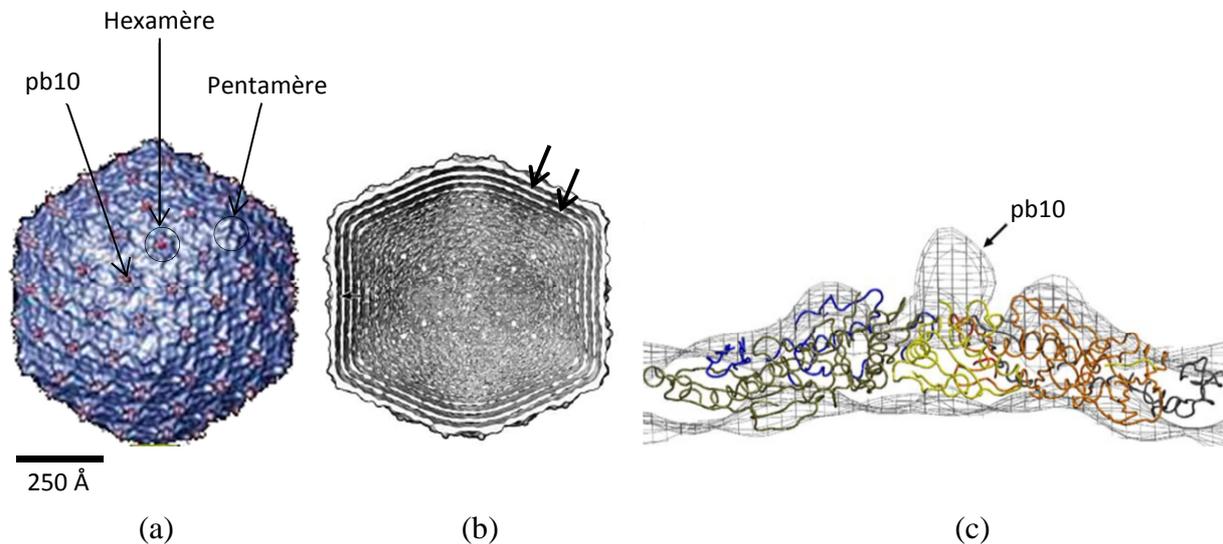


**Figure 17. Présentation structurale du phage T5.** *Au premier plan* : représentation du phage T5 à partir des structures résolues de manière indépendante de la capsid, du tube hélicoïdal et de la partie distale de la queue. Les structures de la capsid et du tube ont été déterminées à l'aide de cryo-microscopie électronique, celle de la pointe distale par microscopie électronique à coloration négative (Effantin, Boulanger et al. 2006). Le connecteur ainsi que les différentes fibres de queue sont représentées schématiquement en jaune. *En arrière-plan* : Cliché de microscopie électronique par coloration négative de plusieurs phages T5. *En haut à droite* : Vue de la section centrale de la reconstruction tridimensionnelle de la capsid pleine d'ADN de T5, permettant de visualiser l'ADN compacté sous une forme quasi-cristalline. Figure adaptée de (Effantin, Boulanger et al. 2006).

pb8 et la protéine majeure de capsid de HK97 gp5 ne présentent que 18% d'identité de séquence, mais possèdent très vraisemblablement le même repliement (figure 18). Ce repliement est commun à plusieurs phages appartenant aux 3 familles de l'ordre des



après l'assemblage de la procapside. Bien que ces deux domaines ne soient pas homologues, ils sont tous les deux prédits avec un repliement en hélices surenroulées (*coiled-coil*) (figure 20 (a)).

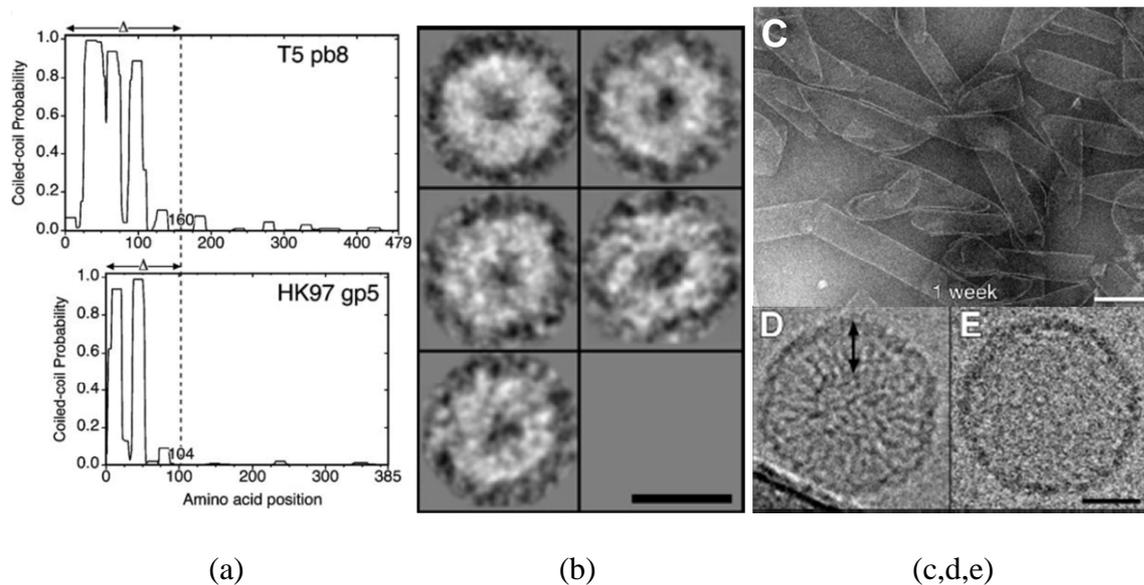


**Figure 19. Structure de la capsid de T5.** (a) Reconstruction tridimensionnelle de la capsid pleine d'ADN du phage T5 à partir de clichés de cryo-microscopie électronique (résolution : 20 Å). Les 775 copies de la protéine majeure de capsid pb8 sont coloriées en bleu, les 120 copies de la protéine de décoration pb10, situées au centre de chaque hexamère, sont coloriées en rouge. (b) Vue en coupe de la reconstruction précédente. Les flèches noires pointent vers les couches de l'ADN compacté dans un état quasi-cristallin. (c) Représentation des capsomères de la capsid de HK97 insérés dans la carte de densité électronique de la capsid de T5 (en gris). La densité non occupée par gp5, indiquée par la flèche noire, correspond à la protéine de décoration. Figure adaptée de (Effantin, Boulanger et al. 2006).

## 7-2 L'assemblage *in vitro* de la procapsid

La protéine pb8 dotée de son domaine d'échafaudage (pb8p), produite par surexpression chez *E. Coli* et purifiée par chromatographie d'affinité se présente à l'état de capsomère (hexamère et/ou pentamère) (Huet, Conway et al. 2010) (figure 20 (b)). L'auto-assemblage *in vitro* de ces capsomères est régulé par la concentration en ions monovalents et divalents ; suivant ces concentrations cet assemblage peut conduire à la formation de tubes ou

de procapsides (figure 20 (c) et (d)). Comme pour les phages HK97, P22, T4 et  $\phi$ 29, la séquence de la protéine pb8p contient donc toute l'information nécessaire pour s'assembler en procapside, sans la présence de la portale.



**Figure 20. Domaine d'échafaudage et assemblage in vitro de pb8.** (a) Prédiction de domaines en hélices surenroulées (*coiled-coil*) des protéines pb8 (haut) et gp5 (bas). Les 160 (pb8) et 104 (gp5) premiers acides aminés correspondent au  $\Delta$ -domaine. (b) Capsomères de pb8p visualisés par microscopie électronique à coloration négative. La barre représente 50 nm. (c) Visualisation par microscopie électronique à coloration négative de tubes de pb8. Ces tubes sont formés après incubation pendant une semaine des capsomères de pb8p dans une solution contenant 100 mM de NaCl. La barre représente 100 nm. (d) Visualisation par cryo-microscopie électronique d'une procapside de T5, formée par incubation pendant 24 heures des capsomères de pb8p dans une solution contenant 10 mM Tris pH 7.6 et 40 mM  $\text{CaCl}_2$ . La double flèche noire montre une structure radiale mesurant 20nm de long, le domaine d'échafaudage. Cette procapside est comparée avec la procapside sans le domaine d'échafaudage (e), obtenue par infection d'*Escherichia Coli* avec un mutant de T5 incapable d'encapsider l'ADN (mutant AMN5, (Zweig and Cummings 1973)). La barre représente 25 nm. Figure adaptée de (Huet, Conway et al. 2010).

## **II - Introduction méthodologique**

### **1- La diffusion des rayons X aux petits angles**

La diffusion des rayons X aux petits angles est une méthode couramment utilisée en biologie structurale pour étudier, à basse résolution, les macromolécules ou les assemblages de macromolécules en solution. Elle permet de déterminer la masse molaire, la taille et la forme des objets étudiés. Elle est complémentaire des méthodes structurales à haute résolution, comme la cristallographie. Contrairement à cette dernière, elle ne nécessite pas l'obtention de cristaux. Elle permet notamment d'acquérir des informations sur des changements structuraux en réponse à des variations de paramètres externes et permet une étude résolue dans le temps de ces changements structuraux.

#### **1-1 Rappels historiques**

La diffusion des rayons X aux petits angles dédiée à la caractérisation d'objets biologiques est aujourd'hui une méthode bien établie qui a nécessité plusieurs décennies de développement. Les rayons X, à la base de cette technique, ont été découverts en 1895 par le physicien allemand Röntgen, ce qui lui vaudra le premier prix Nobel de physique en 1901. Dans les années 1930, André Guinier mit au point une chambre de diffraction des rayons X, appelée "chambre de Guinier", permettant d'étudier la diffusion des rayons X au voisinage immédiat du faisceau incident, c'est-à-dire aux « petits angles » (Guinier 1938). Les premières expériences de SAXS effectuées sur des échantillons biologiques ne permettaient que la détermination de paramètres simples, comme le rayon de giration (Guinier and Fournet

1955). Cette technique a ensuite beaucoup progressé avec l'utilisation du rayonnement synchrotron (années 1970) puis avec l'apparition d'ordinateurs de bureau de puissance suffisante (années 1980-90) et le développement de programmes permettant l'interprétation des données de SAXS et la modélisation structurale à partir de ces données, notamment la série ATSAS, initiée par D.Svergun à l'EMBL de Hambourg (<http://www.embl-hamburg.de/biosaxs/>) (Petoukhov, Franke et al. 2012).

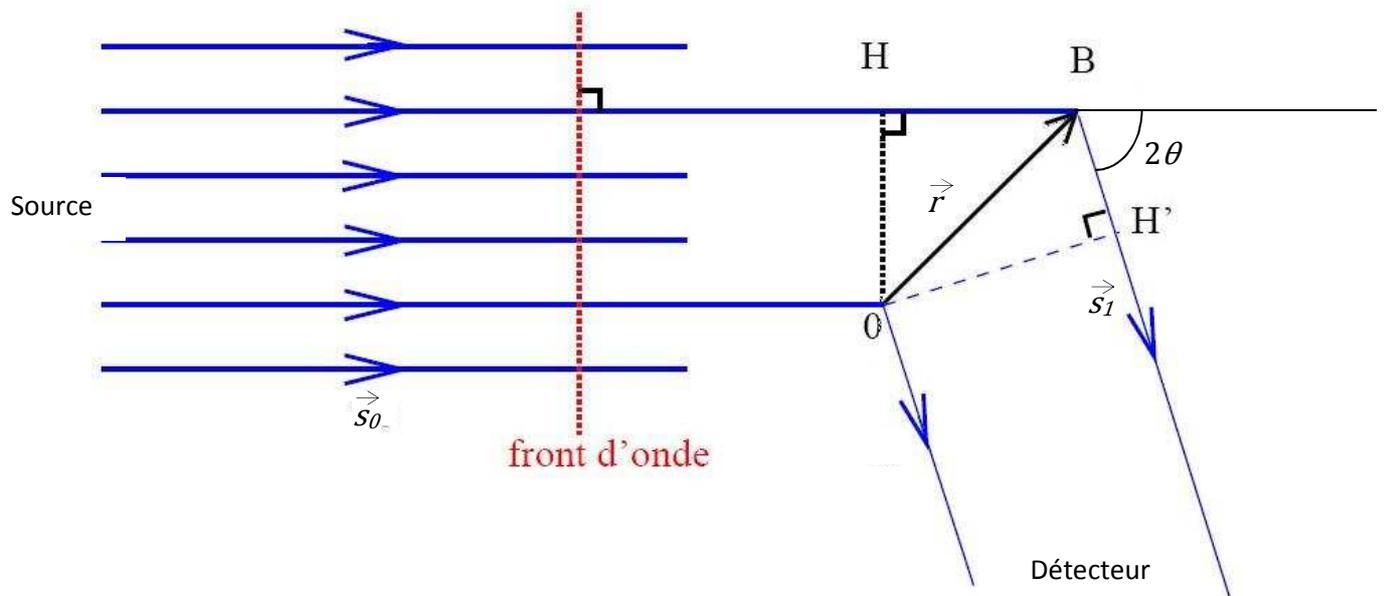
## 1-2 Théorie de la diffusion des rayons X aux petits angles

Le principe du SAXS repose sur l'interaction entre un faisceau de rayons X ayant une longueur d'onde de l'ordre de l'ångström et les électrons composant l'objet étudié.

### 1-2-1 La diffusion par deux électrons

Prenons le cas de deux électrons, respectivement placés en O (origine) et en B (coordonnées  $\vec{r}$ ) (Figure 21). La source de rayons X est supposée très lointaine, et l'observation se fait à l'infini, donc les faisceaux incidents et diffusés sont parallèles. Le vecteur d'onde de l'onde incidente vaut  $\vec{k}_0$ , celui de l'onde diffusée suivant l'angle  $2\theta$  vaut  $\vec{k}_1$ . La diffusion étant élastique, la norme du vecteur d'onde de l'onde incidente est égale à celle du vecteur d'onde de l'onde diffusée :

$$|\vec{k}_0| = |\vec{k}_1| = \frac{2\pi}{\lambda}$$

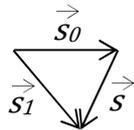


**Figure 21. Trajet de deux rayons diffusés par deux électrons situés en O et en B.**

La différence de marche optique vaut :

$$\Delta l = HB + BH' = \lambda \vec{r} \cdot \vec{s}_0 - \lambda \vec{r} \cdot \vec{s}_1 = -\lambda \vec{r} \cdot \vec{s}$$

avec  $\vec{s}_0 = \frac{\vec{k}_0}{2\pi}$ ,  $\vec{s}_1 = \frac{\vec{k}_1}{2\pi}$  et  $\vec{s} = \vec{s}_1 - \vec{s}_0$



Le déphasage entre les deux ondes est donc :

$$\Delta\varphi = -2\pi \vec{r} \cdot \vec{s} = -\vec{q} \cdot \vec{r}$$

avec  $\vec{q} = 2\pi \vec{s}$

L'amplitude de diffusion par un électron de coordonnées  $\vec{r}$  vaut alors :

$$F(\vec{q}) = A_0 f e^{-i \vec{q} \cdot \vec{r}}$$

où  $f$  est le facteur de diffusion de l'électron et  $A_0$  l'amplitude de l'onde incidente. Nous considérerons dans la suite que  $A_0 f = 1$ .

## 1-2-2 La diffusion par N électrons

On considère maintenant la diffusion par une particule constituée de  $N$  électrons de coordonnées respectives  $\vec{r}_j$  (figure 22). L'amplitude de diffusion de la superposition de toutes les ondes diffusées est:

$$F(\vec{q}) = \sum_{j=1}^N e^{-i\vec{q}\cdot\vec{r}_j}$$

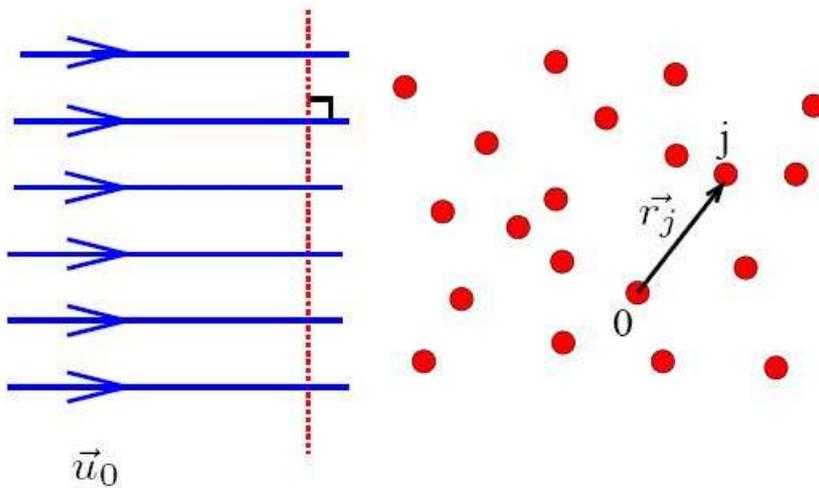


Figure 22. Schéma de rayons X arrivant sur une structure de  $N$  atomes.

En réalité, les électrons sont délocalisés autour des noyaux atomiques. On considère donc la densité électronique de l'objet irradié  $\rho(\vec{r})$ , égale au nombre d'électrons par unité de volume. Le grand nombre d'électrons de la particule impose le passage au continu, la somme se transformant en intégrale, on obtient donc:

$$F(\vec{q}) = \int_V \rho(\vec{r}) e^{-i\vec{q}\cdot\vec{r}} dV$$

Cette amplitude est donc la transformée de Fourier de la densité électronique.

### 1-2-3 Diffusion d'une particule en solution

En pratique les particules étudiées sont en solution. Ce qui nous intéresse est la diffusion des rayons X liée aux particules d'intérêt et non la diffusion liée aux molécules du solvant :

$$F(\vec{q}) = \int_V (\rho(\vec{r}) - \rho_s) e^{-i\vec{q}\cdot\vec{r}} dV \quad (1)$$

où  $\rho_s$  est la densité électronique du solvant, estimée uniforme.

Les particules étudiées sont soumises à l'agitation thermique ; durant le temps d'exposition aux rayons X (environ une seconde), ces particules possèdent toutes les orientations par rapport au faisceau de rayons X. De plus toutes les molécules étudiées ont des orientations différentes par rapport à ce faisceau. Ainsi seule la moyenne sphérique de l'intensité diffusée est accessible expérimentalement :

$$I(q) = \langle I(\vec{q}) \rangle = \langle F(\vec{q})F^*(\vec{q}) \rangle \quad (2)$$

où  $F^*(\vec{q})$  est le complexe conjugué de l'amplitude de diffusion et  $q = |\vec{q}| = \frac{4\pi\sin(\theta)}{\lambda}$ . La longueur d'onde  $\lambda$  est constante lors d'une expérience et  $\sin(\theta) \approx \theta$  aux petits angles,  $q$  est donc proportionnel au demi-angle de diffusion  $\theta$ .

Deux conditions sont requises pour une analyse correcte des courbes de diffusion : l'idéalité et la monodispersité de la solution. Une solution est dite idéale s'il n'y a pas d'interaction entre ces particules. Pour une solution idéale composée de  $N$  particules ayant chacune une diffusion  $i_j(q)$ , l'intensité totale diffusée vaut :

$$I(q) = \sum_{j=1}^N i_j(q) \quad (3)$$

Pour une solution monodisperse l'intensité diffusée par chacune des particules est identique, donc l'intensité totale diffusée vaut :

$$I(q) = Ni(q) \quad (4)$$

Pour une solution idéale et monodisperse, la mesure de l'intensité diffusée totale, accessible expérimentalement, permet donc de déterminer l'intensité de diffusion d'une particule, qui renseigne sur sa masse molaire, sa taille et sa forme. Ces deux conditions doivent donc être vérifiées lors des expérimentations. La monodispersité dépend du protocole de purification et peut notamment être estimée par gel SDS-PAGE et par diffusion dynamique de la lumière (DLS). L'idéalité est théoriquement atteinte si les particules sont sans interaction. En pratique elle est évaluée en effectuant des mesures à différentes concentrations pour vérifier que les courbes de diffusions sont indépendantes de la concentration.

#### 1-2-4 La loi de Guinier

La loi de Guinier stipule que la courbe de diffusion d'une particule peut être approximée par une gaussienne au voisinage de l'origine :

$$I(q) \cong I(0)e^{-\left(\frac{q^2 R_g^2}{3}\right)}$$

avec  $R_g$  le rayon de giration, pouvant s'exprimer :

$$R_g^2 = \frac{\int_V r^2 (\rho(\vec{r}) - \rho_s) dV}{\int_V (\rho(\vec{r}) - \rho_s) dV}$$

Ce rayon de giration correspond donc à la moyenne, pondérée par le contraste de densité électronique, du carré de la distance au centre de masse.

La loi de Guinier peut s'écrire sous sa forme logarithmique :

$$\ln[I(q)] \cong \ln[I(0)] - \frac{R_g^2}{3} q^2$$

Ainsi en traçant le logarithme népérien de l'intensité diffusée  $\ln[I(q)]$  en fonction de  $q^2$ , la pente de cette courbe vaut  $\frac{R_g^2}{3}$ , ce qui permet de déterminer le rayon de giration. Pour un objet « plein », cette grandeur est un « indice » de non-sphéricité. En effet, dans le cas où la densité électronique est uniforme dans la particule considérée, pour un volume donné le plus petit rayon de giration est celui d'une sphère pleine :  $R_g = \sqrt{\frac{3}{5}} R$ . A volume égal, plus l'objet est allongé plus le rayon de giration est important. Un objet creux, et à plus forte raison un objet en forme de couche sphérique, c'est-à-dire d'« écorce d'orange », a aussi un rayon de giration plus grand que celui d'une sphère pleine.

Typiquement, un photon sur  $10^5$  à  $10^7$  photons incidents est diffusé (Winick and Doniach 1980). Les rayons non déviés sont arrêtés par un cache appelé *beamstop* (figure 23), l'intensité diffusée à l'origine  $I(0)$  est donc inaccessible expérimentalement, mais la loi de Guinier en permet néanmoins une bonne approximation. D'après les équations (1), (2) et (4), l'intensité diffusée peut être exprimée comme :

$$I(q) = N \left\langle \left| \int_V (\rho(\vec{r}) - \rho_s) e^{-i \vec{q} \cdot \vec{r}} dV \right|^2 \right\rangle$$

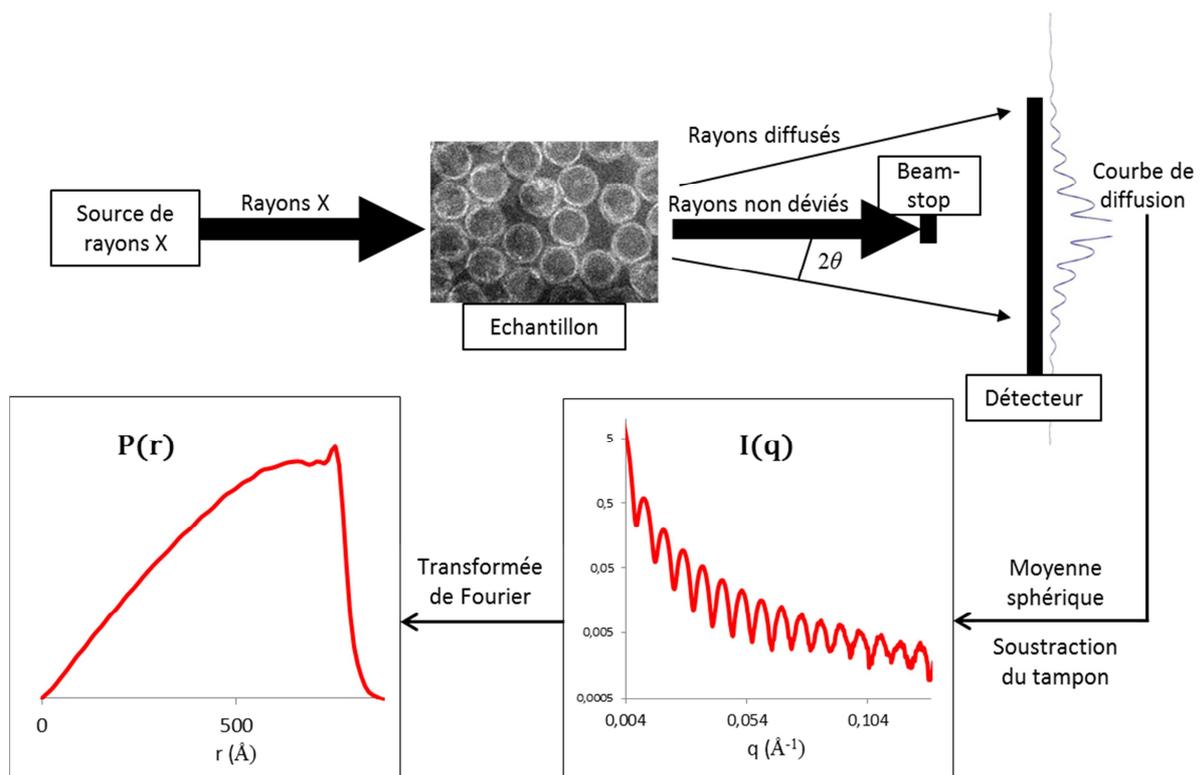
On en déduit :

$$I(0) = N \left\langle \left| \int_V (\rho(\vec{r}) - \rho_s) dV \right|^2 \right\rangle$$

On peut en déduire une expression de l'intensité diffusée à l'origine :

$$I(0) = \frac{V_{sol} c M}{N_A} \left( \frac{m_e N_A}{M} - \rho_s \bar{v}_p \right)^2$$

avec  $N_A$  la constante d'Avogadro,  $V_{sol}$  le volume de la solution sous faisceau,  $c$  la concentration massique de la protéine,  $M$  sa masse moléculaire,  $m_e$  son nombre d'électron,  $\bar{v}_p$  son volume partiel spécifique et  $\rho_s$  la densité électronique du solvant. Le volume partiel spécifique peut être estimé d'après la séquence de la protéine. La connaissance de l'intensité diffusée à l'origine permet donc de déterminer de manière approchée la masse molaire, ce qui permet de connaître l'état d'oligomérisation de la particule en solution.



**Figure 23. Vue d'ensemble d'une expérience de SAXS et du traitement des résultats.** En haut : schéma du montage expérimental d'une expérience de diffusion des rayons X aux petits angles. La partie « source de rayons X » comprend, outre la source lumineuse proprement dite (anode tournante ou synchrotron), le système optique permettant de sélectionner une longueur d'onde précise et de collimater le faisceau.

D'autre part, pour chaque expérience de SAXS il convient de vérifier la linéarité de l'intensité diffusée  $\ln[I(q)]$  en fonction de  $q^2$ , et ceci pour différentes concentrations. En effet, une diminution de  $I(0)$  quand la concentration augmente est une indication d'interactions répulsives entre les particules diffusantes, le rayon de giration aura alors

tendance à être sous-estimé. A l'inverse, une augmentation de  $I(0)$  quand la concentration augmente est une indication d'interactions attractive ou d'agréations ; dans ce cas le rayon de giration sera surestimé.

## 1-2-5 La fonction de distribution des distances

Bien que la courbe de l'intensité diffusée en fonction de l'angle  $I(q)$  soit reliée à la forme de la particule en solution, la fonction de distribution des distances  $P(r)$  en permet une représentation structurale plus aisée (figure 24). La  $P(r)$  représente l'histogramme des distances entre chaque paire d'éléments diffusants, c'est-à-dire d'électrons, de la particule étudiée. Mathématiquement, c'est la transformée de Fourier inverse de l'intensité diffusée :

$$P(r) = \frac{r^2}{2\pi} \int_0^{\infty} q^2 I(q) \frac{\sin(qr)}{qr} dq$$

Cependant l'intensité diffusée n'étant connue que sur une gamme finie  $[q_{min}, q_{max}]$ , le calcul de la  $P(r)$  ne peut se faire en utilisant directement cette formule. Alternativement, ce calcul peut être effectué par la méthode de la transformée de Fourier indirecte (Glatter 1977) qui en pratique, peut être exécutée par les programmes GNOM (Svergun, Semenyuk et al. 1988; Svergun 1992) ou GIFT (Bergmann, Fritz et al. 2000).

A noter que le diamètre maximal  $D_{max}$  de la particule s'obtient avec la  $P(r)$  : il s'agit de la valeur de  $r$  pour laquelle cette fonction s'annule. Pour un calcul correct de la fonction de distribution des distances il est nécessaire de collecter les données dans une gamme angulaire dont le plus petit angle est suffisamment petit (Moore 1980) :

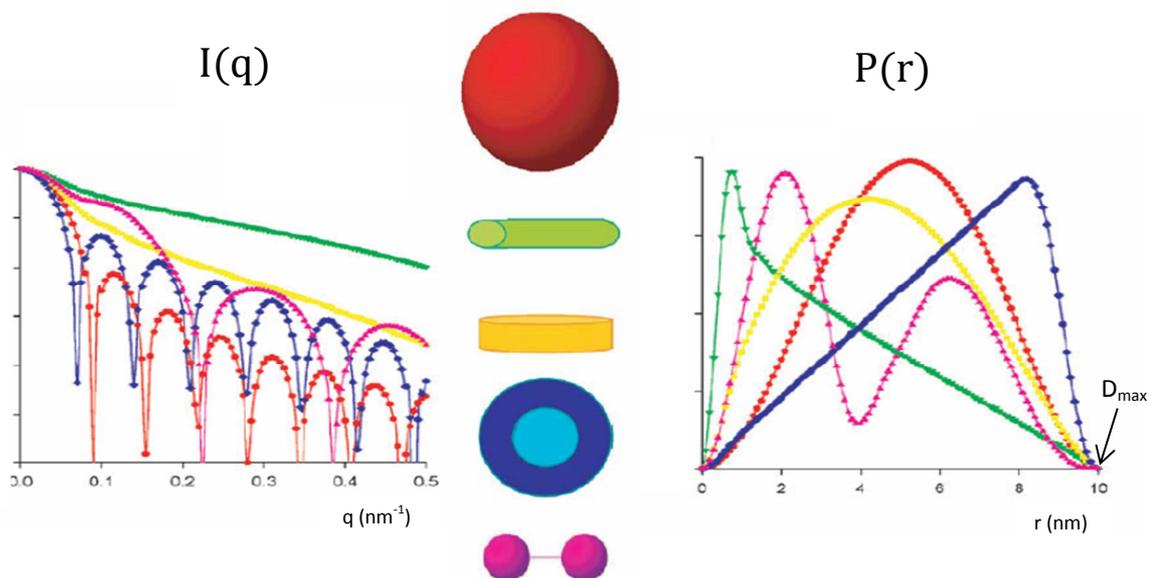
$$q_{min} < \frac{\pi}{D_{max}}$$

D'autre part, le rayon de giration ainsi que l'intensité diffusée à l'origine peuvent s'obtenir à l'aide de la fonction de distribution des distances en utilisant les formules suivantes :

$$R_g^2 = \frac{\int_0^{D_{max}} r^2 p(r) dr}{2 \int_0^{D_{max}} p(r) dr}$$

$$I(0) = 4\pi \int_0^{D_{max}} p(r) dr$$

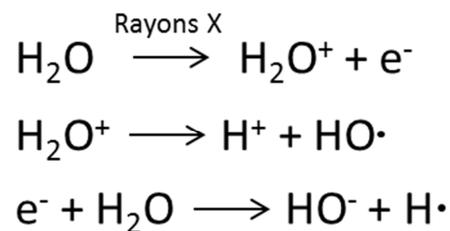
Alors que la loi de Guinier ne s'applique que dans la région des petits  $q$ , cette méthode alternative utilise l'intégralité de la courbe de diffusion. Elle est donc moins sensible aux interactions répulsives ou attractives entre les particules ou à la présence d'oligomères. La comparaison des résultats issus des deux méthodes est également un bon contrôle de la validité de ces résultats.



**Figure 24. Exemples de courbes d'intensité de diffusion  $I(q)$  et de distribution de distance  $P(r)$ .** Courbes théoriques d'une boule (rouge), d'un cylindre de grande hauteur ( $h \gg R$ ) (vert), d'un cylindre de petite hauteur  $h \ll R$  (jaune), d'une couche sphérique (bleu) et d'un solide composé de deux boules reliées par une tige infiniment fine (magenta). Figure adaptée de (Svergun and Koch 2003).

### 1-2-6 Le problème des radicaux libres

Les rayons X interagissent avec l'eau en provoquant son ionisation (figure 25). Les molécules d'eau ionisées perdent un électron, engendrant une molécule d'eau cationique  $\text{H}_2\text{O}^+$ . L'eau cationique est un acide fort qui perd rapidement un proton au profit de l'eau environnante, formant un radical  $\text{OH}\cdot$ . Les électrons perdus réagissent avec des molécules d'eau pour donner des radicaux  $\text{H}\cdot$  (Bensasson 1993).



**Figure 25. Equations chimiques de la formation des radicaux libres pendant l'exposition à un faisceau de rayons X.**

Ces radicaux libres sont très réactifs. Ils oxydent les protéines ce qui provoque des modifications structurales (pontage intra-protéine) et l'agrégation de ces protéines (pontage inter-protéines) (Kuwamoto, Akiyama et al. 2004). Cela peut altérer grandement la qualité des mesures de SAXS et compromettre leur interprétation, notamment la valeur du rayon de giration et de l'intensité diffusée à l'origine, mais également de la  $P(r)$ . L'identification de la présence de tels dégâts d'irradiation se fait à l'aide d'une comparaison des spectres obtenus après différents temps d'exposition. L'utilisation d'agent réducteur tel le DTT permet de limiter ces dégâts d'irradiation, cependant le DTT ne peut pas être utilisé en condition fortement acide ( $\text{pH} \leq 4$ ).

## 1-2-7 Reconstitution de l'enveloppe moléculaire à l'aide du programme

### GASBOR

Outre les données précédemment citées, la technique de SAXS permet de reconstituer *ab initio* et à basse résolution l'enveloppe moléculaire de la protéine étudiée. Je décris ici le programme GASBOR (Svergun, Petoukhov et al. 2001), qui a été utilisé pour déterminer l'enveloppe de la protéine pb10, pour laquelle nous ne disposons d'aucune autre information structurale. Ce programme est basé sur la modélisation de la protéine étudiée par une chaîne continue de pseudo-résidus, chacun étant représenté par une sphère de diamètre 3,8 Å centrée sur la position du carbone  $\alpha$  et ayant une densité électronique égale à la densité électronique moyenne d'un acide aminé. L'algorithme, qui est basé sur le principe du recuit simulé, permet de minimiser la fonction :

$$E(\mathbf{r}) = \chi^2 + \alpha P(\mathbf{r})$$

où  $\mathbf{r}$  représente les coordonnées des pseudo atomes,  $\chi^2$  correspond à la différence entre les données expérimentales et les données calculées à partir du modèle, divisé par la barre d'erreur :

$$\chi^2 = \frac{1}{n-1} \sum_{j=1}^n \left[ \frac{\langle c(s) \rangle I_{DR}(s_j) - c I_{exp}(s_j)}{\sigma(s_j)} \right]^2,$$

$P(\mathbf{r})$  est une fonction de pénalité permettant d'éliminer les modèles ayant une forme non compatible avec celle d'une protéine, notamment les modèles discontinus, et  $\alpha$  est choisi pour permettre au terme  $\alpha P(\mathbf{r})$  de contribuer pour 10 à 50 % à la fonction  $E(\mathbf{r})$ .

La non-unicité de la solution de reconstruction d'enveloppe moléculaire par SAXS impose de calculer plusieurs fois (au minimum 10) l'enveloppe à partir des mêmes données de diffusion, et de comparer les résultats obtenus en calculant la différence spatiale normalisée (NSD) (Kozin and Svergun 2001). Cette grandeur est une mesure de la similarité

entre des ensembles de points dans un espace tridimensionnel ; plus elle est petite et plus les ensembles de points sont similaires. Pour des objets identiques la valeur de la NSD est de 0, dans le cas du programme GASBOR une valeur de l'ordre de 1 permet de conclure à une bonne similitude des enveloppes.

## 2- La décomposition en valeurs singulières (SVD)

### 2-1 Définition

La décomposition en valeurs singulières est un outil mathématique permettant d'analyser des jeux de données en mettant en évidence des caractéristiques de ces données. Elle est basée sur la factorisation suivante d'une matrice  $m$  à coefficients réels ou complexes :

$$m = usv^T \quad (5)$$

où  $s$  est une matrice rectangulaire diagonale et  $u$  et  $v$  sont deux matrices unitaires, c'est-à-dire que  $uu^T = u^T u = I$  et  $vv^T = v^T v = I$ ,  $u^T$  et  $v^T$  étant les matrices transposées de  $u$  et  $v$ , et  $I$  étant la matrice unité. La matrice  $s$  contient les valeurs singulières de la matrice  $m$ , qui sont par convention rangées par ordre décroissant. Cette matrice est alors déterminée de façon unique.

### 2-2 Exemple de la décomposition d'une matrice à coefficients réels

On considère la matrice suivante :

$$m = \begin{bmatrix} 3 & 1 & 1 \\ -1 & 3 & 1 \end{bmatrix}$$

La matrice  $u$  s'obtient à l'aide des vecteurs propres de la matrice  $mm^T$ .

$$mm^T = \begin{bmatrix} 3 & 1 & 1 \\ -1 & 3 & 1 \end{bmatrix} \begin{bmatrix} 3 & -1 \\ 1 & 3 \\ 1 & 1 \end{bmatrix} = \begin{bmatrix} 11 & 1 \\ 1 & 11 \end{bmatrix}$$

Les valeurs propres de cette matrice sont  $\lambda_1 = 12$  et  $\lambda_2 = 10$  avec pour vecteurs propres associés  $\begin{bmatrix} 1 \\ 1 \end{bmatrix}$  et  $\begin{bmatrix} 1 \\ -1 \end{bmatrix}$ . La matrice  $\mathbf{u}$  s'obtient après ortho-normalisation de la matrice dont les colonnes sont composées des vecteurs propres de  $mm^T$ :

$$\mathbf{u} = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}$$

La matrice  $\mathbf{v}$  s'obtient de la même manière à partir de  $m^T m$ .

$$\mathbf{v}^T = \frac{1}{\sqrt{30}} \begin{bmatrix} \sqrt{5} & 2\sqrt{5} & \sqrt{5} \\ 2\sqrt{6} & -\sqrt{6} & 0 \\ 1 & 2 & -5 \end{bmatrix}$$

Les coefficients de la matrice  $\mathbf{s}$  sont nuls sauf sur la diagonale qui est composée de valeurs propres non nulles de  $mm^T$  ou de  $m^T m$  (qui sont identiques) :

$$\mathbf{s} = \begin{bmatrix} \sqrt{12} & 0 & 0 \\ 0 & \sqrt{10} & 0 \end{bmatrix}$$

Finalement, nous pouvons effectuer la décomposition de la matrice  $\mathbf{m}$  en valeurs singulières :

$$\mathbf{m} = \begin{bmatrix} 3 & 1 & 1 \\ -1 & 3 & 1 \end{bmatrix} = \frac{1}{\sqrt{60}} \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} \begin{bmatrix} \sqrt{12} & 0 & 0 \\ 0 & \sqrt{10} & 0 \end{bmatrix} \begin{bmatrix} \sqrt{5} & 2\sqrt{5} & \sqrt{5} \\ 2\sqrt{6} & -\sqrt{6} & 0 \\ 1 & 2 & -5 \end{bmatrix}$$

### 2-3 Application à l'étude d'un ensemble de courbes

La technique de SVD peut être appliquée sur un ensemble de courbe pour en déterminer certaines propriétés. Elle permet notamment de déterminer le nombre minimal de courbes nécessaires pour générer, par combinaison linéaire, chaque courbe.

### 2-3-1 Courbes théoriques sans bruit de fond expérimental

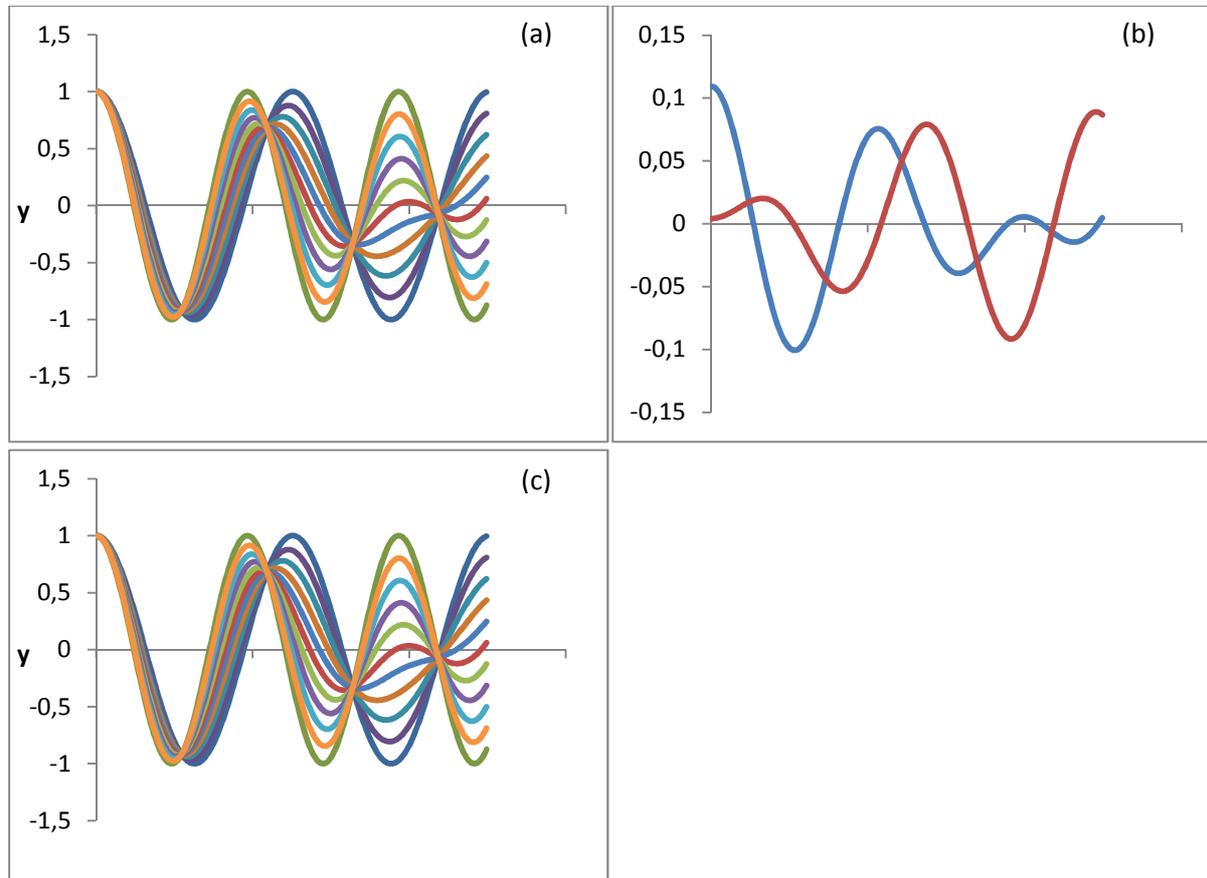
En pratique, chaque colonne de la matrice  $m$  est associée à une courbe ; la valeur de l'ordonnée de chaque point d'une courbe forme un coefficient de cette colonne. Si l'abscisse comporte  $a$  points et que l'ensemble est composée de  $b$  courbes, alors la matrice  $m$  sera de dimension  $b \times a$ . La décomposition en valeurs singulières de  $m$  permet de déterminer les matrices  $u$ ,  $s$  et  $t$  qui vérifient l'équation (5). L'information contenue dans la matrice  $s$  est particulièrement importante. Le nombre de ses coefficients non nuls, c'est-à-dire le nombre de valeurs singulières, est égal au rang de la matrice  $m$ , c'est-à-dire au le nombre minimal de courbes nécessaires pour générer, par combinaison linéaire, chaque courbe de l'ensemble. La matrice  $u$  contient les vecteurs de bases, c'est-à-dire les courbes à partir desquelles il est possible de reconstituer l'ensemble des courbes. Le premier vecteur de base correspond à la moyenne de toutes les courbes de l'ensemble ; il est normalisé : l'intégrale de la valeur absolue de ce vecteur vaut donc 1. Le second vecteur est le vecteur orthogonal au premier vecteur qui minimise la différence entre chacune des courbes de l'ensemble et sa reconstitution par les 2 premiers vecteurs propres ; les vecteurs suivants sont définis de la même manière, ils sont orthogonaux les uns aux autres. La matrice  $v$  contient les projections de chaque courbe sur chacun des vecteurs propres.

L'application de la décomposition en valeurs singulières à l'étude d'un ensemble de courbes est illustrée par les exemples suivants.

#### Exemple 1

Considérons deux courbes, d'équation  $y_A = \cos x$  et  $y_B = \cos 1.3x$ , définies pour  $x \in [0, 4\pi]$ . Soit  $\mathcal{C}$  un ensemble de courbes générées par combinaison linéaire de ces deux courbes, défini par  $y_n = 0.1n \times y_A + (1 - 0.1n) \times y_B$ , avec  $n$  compris entre 0 et 10 (figure 26). Toutes ces courbes étant combinaison linéaire de deux courbes, la décomposition en

valeurs singulières de la matrice représentant l'ensemble  $\mathfrak{C}$  ne fait apparaître que 2 valeurs singulières non nulles (table 1). De plus, il est possible de reconstituer parfaitement chaque courbe à partir des deux premiers vecteurs propres (figure 26 (c)).



**Figure 26. Analyse SVD de  $\mathfrak{C}$ , un ensemble de courbes qui sont combinaison linéaire de 2 courbes.** Visualisation des 11 courbes de l'ensemble  $\mathfrak{C}$  (a), des deux vecteurs propres associés aux deux premières valeurs singulières (b), ainsi que de la reconstruction des courbes de l'ensemble  $\mathfrak{C}$  à partir de ces deux vecteurs (c). La courbe du premier vecteur propre est tracée en bleu et celle du deuxième vecteur propre en rouge.

Valeurs singulières	$\mathfrak{E}$	$\mathfrak{E}_1$	$\mathfrak{E}_2$
1 <sup>ère</sup>	30.28	30.28	30.30
2 <sup>ème</sup>	22.68	23.00	22.68
3 <sup>ème</sup>	0	13.7	0.8
4 <sup>ème</sup>	0	0	0.6
...	0	0	~ 0.6
11 <sup>ème</sup>	0	0	0.5

**Table 1. Valeurs singulières issue de l'analyse SVD des ensembles  $\mathfrak{E}$ ,  $\mathfrak{E}_1$  et  $\mathfrak{E}_2$ .**

### Exemple 2

Ajoutons maintenant la courbe d'équation  $y = \cos 1.6x$ , qui n'est donc pas une combinaison linéaire de  $y_A$  ou de  $y_B$ , à l'ensemble  $\mathfrak{E}$  et appelons ce nouvel ensemble  $\mathfrak{E}_1$  (figure 27 (a)). La décomposition en valeurs singulières de la matrice représentant cet ensemble fait maintenant apparaître 3 valeurs singulières non nulles (table 1). L'utilisation des 2 premiers vecteurs ne permet pas de reconstituer les courbes de l'ensemble  $\mathfrak{E}_1$ , la courbe d'équation  $y = \cos 1.6x$  notamment est très mal reconstruite, alors que l'utilisation des 3 premiers vecteurs le permet parfaitement (figure 27 (c) et (d)).

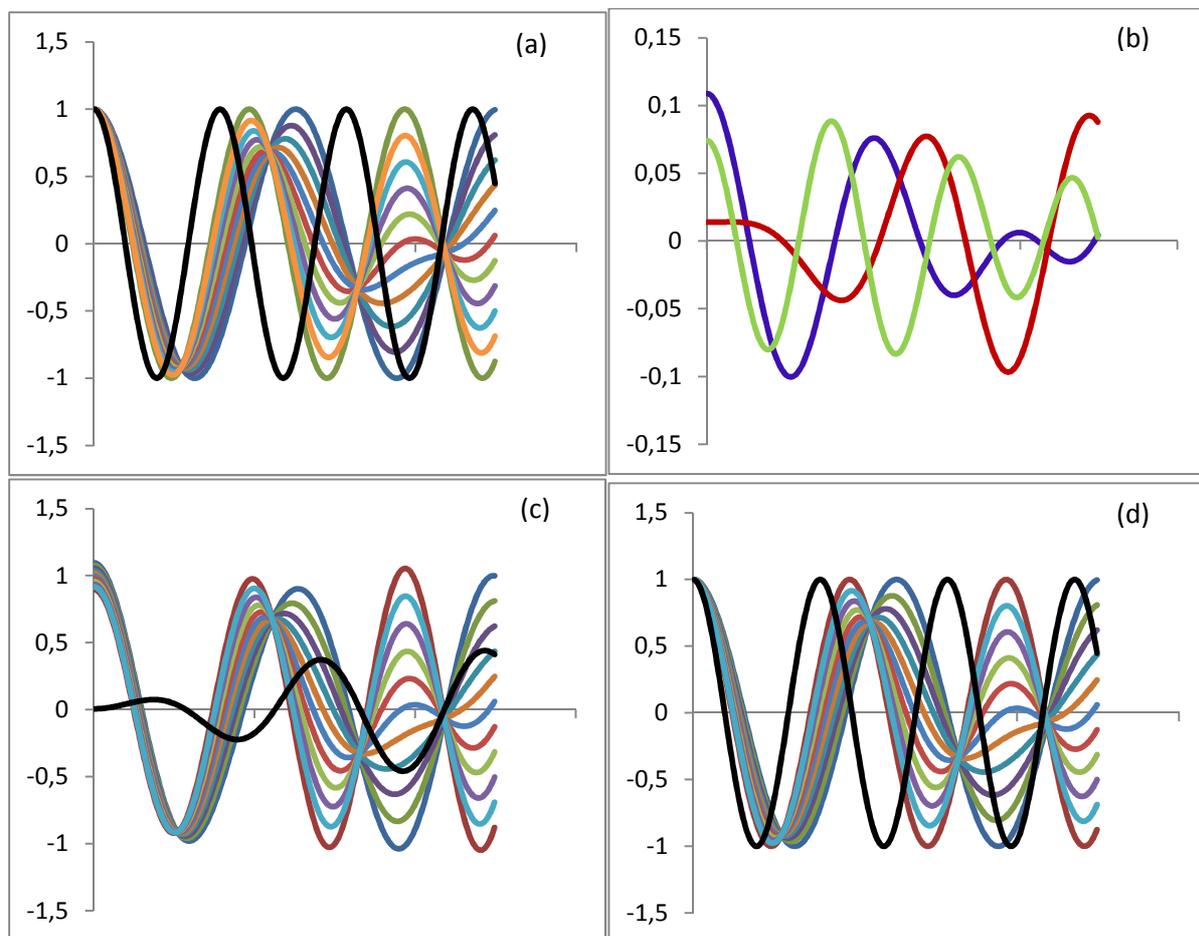
## **2-3-1 Courbes expérimentales**

En pratique les courbes expérimentales sont nécessairement bruitées, intéressons-nous donc à l'effet de ce bruit expérimental sur la décomposition en valeurs singulières avec l'exemple suivant.

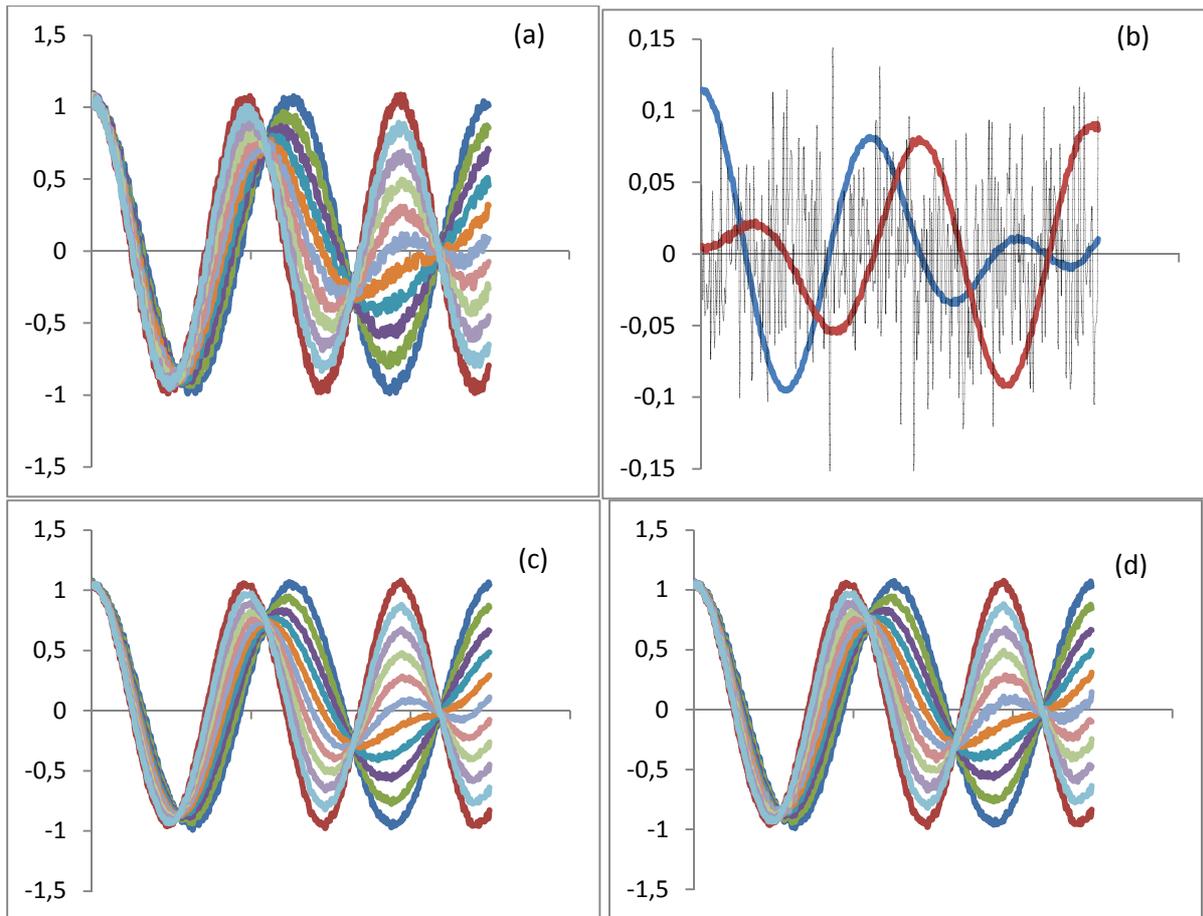
### Exemple 3

L'ensemble  $\mathfrak{E}_2$  est composé de toutes les courbes de l'ensemble  $\mathfrak{E}$  auxquelles on a ajouté une fonction aléatoire d'amplitude  $\pm 0.01$  (figure 28 (a)). Les 11 valeurs singulières ne sont pas nulles, mais les deux premières sont bien plus grandes que les suivantes. Seuls les

vecteurs propres associés à ces 2 valeurs singulières ne sont pas distribués aléatoirement autour de zéro, les vecteurs propres suivant ne rendent compte que du bruit expérimental (figure 28 (b)). L'utilisation des 2 premiers vecteurs permet la reconstitution des données, au bruit expérimental près, tandis que l'utilisation du 3<sup>ème</sup> vecteur ne permet que d'ajouter du bruit de fond aux données reconstituées (figure 28 (c) et (d)).

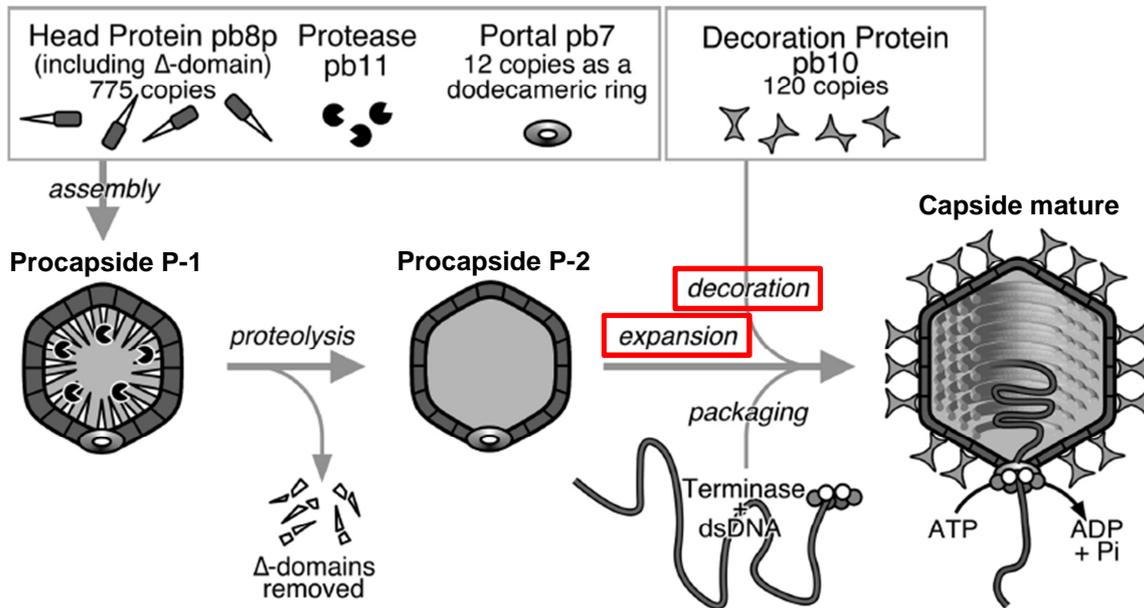


**Figure 27. Analyse SVD de  $\mathfrak{C}_1$ , un ensemble de courbes qui sont combinaison linéaire de 3 courbes.** Visualisation des 12 courbes de l'ensemble  $\mathfrak{C}_1$  (a), des trois vecteurs propres associés aux trois premières valeurs singulières (b) ainsi que de la reconstruction des courbes de l'ensemble  $\mathfrak{C}_1$  à partir des deux (c) et trois (d) premiers vecteurs propres. La courbe d'équation  $y = \cos 1.6x$  est tracée en noir. La courbe du premier vecteur propre est tracée en bleu, celle du deuxième vecteur propre en rouge et celle du troisième vecteur en vert. On peut constater le fait qu'elle est très mal reconstituée avec les deux premiers vecteurs propres uniquement mais très bien reconstituée en utilisant les trois premiers vecteurs propres.



**Figure 28. Analyse SVD de  $\mathfrak{C}_2$ , un ensemble de courbes qui sont combinaison linéaire de 2 courbes.** Visualisation des 11 courbes de l'ensemble  $\mathfrak{C}_2$  (a), des trois vecteurs propres associés aux trois premières valeurs singulières (b) ainsi que de la reconstruction des courbes de l'ensemble  $\mathfrak{C}_2$  à partir des deux (c) et trois (d) premiers vecteurs propres. La courbe du premier vecteur propre est tracée en bleu, celle du deuxième vecteur propre en rouge et celle du troisième vecteur en gris.

### III- Présentation de la thématique de recherche



**Figure 29. Schéma de l'assemblage et de la maturation de la capsid du phage T5.** En rouge sont encadrées les deux étapes de cette maturation sur lesquelles portent mes travaux de recherche : les étapes d'expansion et de décoration de la capsid. Figure adaptée de (Huet, 2010).

Si le phage T5 a été abondamment étudié, notamment les processus de fixation au récepteur membranaire et d'éjection de l'ADN (Letellier, Boulanger et al. 2004), l'étude de sa structure et de son assemblage est plus récente. Les premières étapes de l'assemblage conduisant à une procapsid compétente pour l'encapsidation et impliquant la protéase de maturation pb11 viennent d'être décrites (Huet, Conway et al. 2010) (Zivanovic et al. *en préparation*) (figure 29), cependant les étapes de maturation de la capsid n'ont pas été décrites de manière détaillée comme chez le phage HK97. L'intérêt d'étudier cette étape chez T5 est multiple. D'une part, la capsid a un nombre de triangulation plus important (respectivement  $T = 7$  et  $T = 13$ ), ce qui implique un nombre d'interactions entre sous-unités significativement supérieur. D'autre part, nous disposons d'un mutant de T5 qui permet la

production de procapside P-2, intermédiaire physiologique de l'assemblage dont on peut étudier la maturation, alors que les travaux effectués sur la capsid de HK97, ont été menés sur une capsid reconstituée à partir de la protéine majeure surexprimée chez *E. Coli*, et qui ne contient donc pas la portale.

L'objectif de mon travail de thèse était d'apporter des données concernant :

- l'expansion de la capsid, notamment la présence éventuelle d'états intermédiaires, la coopérativité des changements structuraux ainsi que l'impact de ce processus sur la stabilité de la capsid
- La décoration de la capsid. Il s'agissait de caractériser la protéine de décoration sur le plan structural et d'étudier sa fixation sur la capsid.



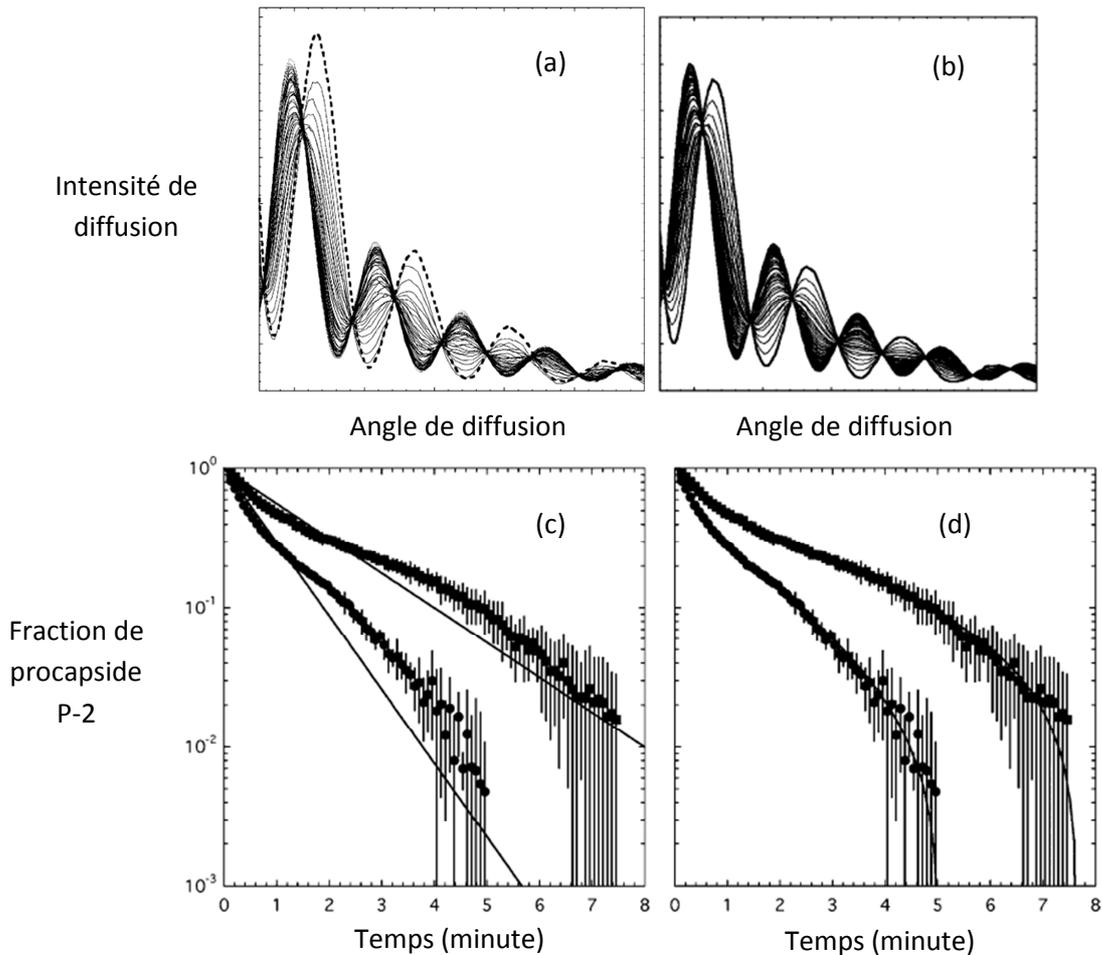
## **Etude cinétique du processus d'expansion**

# I- Introduction

## 1- Etudes cinétiques de l'expansion du phage HK97

### 1-1 Etude de la première phase de l'expansion

La transition entre l'état procapside P-2 et le premier état intermédiaire d'expansion EI-1 peut être induite par acidification à pH~4. Elle a été étudiée par des expériences de SAXS résolues dans le temps (Lee, Tsuruta et al. 2005). Les courbes SAXS, enregistrées toutes les 5 secondes pendant les 5 minutes suivant l'acidification, sont présentées dans la figure 30. Les points isobestiques suggèrent la présence de seulement deux états lors de cette transition. L'analyse SVD effectuée sur l'ensemble des courbes SAXS permet de déterminer le nombre d'espèces de capsides qui contribuent au signal de diffusion des rayons X. Cette analyse indique que seules les deux premières valeurs singulières sont significativement supérieures aux 58 autres et que les deux premiers vecteurs propres seuls sont suffisants pour reconstruire l'ensemble des courbes SAXS. Cela confirme la présence de seulement deux états qui peuvent être distingués par SAXS, c'est-à-dire qu'il n'y a pas d'état intermédiaire dont la forme ou la taille soit significativement différente des états P-2 et EI-1. Ceci montre que la transition est coopérative à l'intérieur d'une capsid : tous les capsomères sont dans le même état (P-2 ou EI-1) (figure 31).



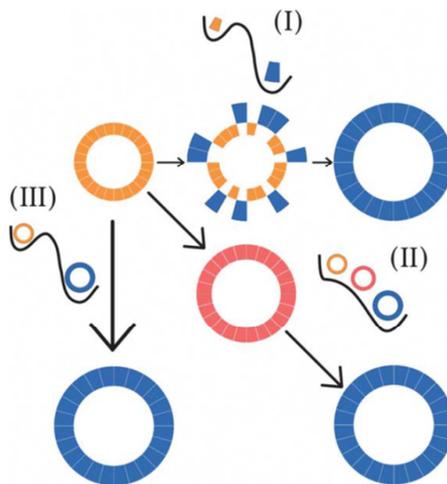
**Figure 30. Analyse par SAXS résolue dans le temps de la première étape de l'expansion de la capsid du phage HK97.** (a) Ensemble de courbes SAXS au cours de l'expansion des procapsides P-2 après acidification à pH 3.8. Ces courbes ont été obtenues toutes les 5 secondes pendant 5 minutes. La courbe en pointillée correspond à l'état P-2 non acidifié. (b) Reconstitution de l'ensemble des courbes expérimentales par les deux premiers vecteurs issus de l'analyse SVD. (c) et (d) Fraction de procapside P-2 (échelle logarithmique) en fonction du temps après acidification à pH 3.8 (évolution la plus rapide) et pH 4.1 (évolution la plus lente). Les lignes continues correspondent aux « fits » obtenus avec le modèle de simple exponentielle (c) (équation 6) et avec une exponentielle double (d). Figure adaptée de (Lee, Tsuruta et al. 2005).

Cependant une transition irréversible sans état intermédiaire donne lieu à une cinétique du premier ordre, c'est-à-dire dont les concentrations  $[E_i]$  et  $[E_F]$  en espèces initiale et finale varient suivant la formule suivante (cf. démonstration en annexe) :

$$[E_i] = C_0 e^{-kt} \quad (6)$$

$$[E_F] = C_0(1 - e^{-kt})$$

où  $C_0$  est la concentration initiale en espèce initiale et  $k$  la constante de vitesse de la transition. Or la variation de la concentration en procapside P-2 en fonction du temps déterminé expérimentalement par SAXS diffère légèrement mais significativement des résultats attendus par ce modèle (figure 30 (c) et (d)), ce qui peut être interprété par l'existence de plusieurs états conformationnels de la procapside P-2, états qui ne peuvent pas être distingués par SAXS. Cependant, les données n'ayant pas été comparées avec un modèle de transition impliquant l'existence d'un état intermédiaire « invisible » par SAXS, l'analyse effectuée ne permet pas de conclure quant à l'existence d'un tel état intermédiaire.

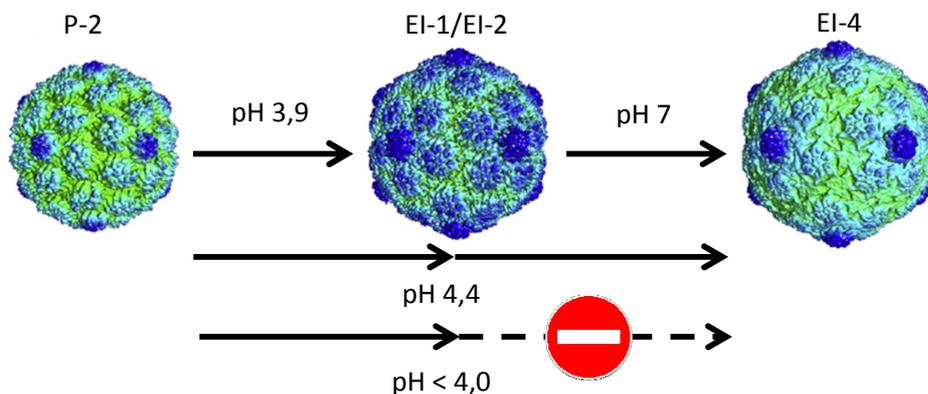


**Figure 31. Schéma des différents chemins réactionnels possibles de la transition de l'état procapside à l'état expansé.** Figure adaptée de (Lee, Tsuruta et al. 2005).

## 1-2 Etude de la seconde phase de l'expansion

La cinétique de l'étape de transition entre les états EI-1 et EI-4 a également été étudiée (Lee, Gan et al. 2008). L'acidification de la procapside P-2 à un pH inférieur à 4 permet de produire des capsides à l'état EI-1 et EI-2 (Ross, Cheng et al. 2005) (figure 32). L'état EI-2

fait référence à des particules de forme et de taille similaires à EI-1 (les états EI-1 et EI-2 ne peuvent donc pas être distingués par SAXS) mais dont le pontage covalent entre leurs capsomères a débuté. La neutralisation à pH 7 des capsides dans l'état EI-1/2 provoque leur passage vers l'état EI-4. Comme pour la transition entre l'état P-2 et l'état EI-1, l'analyse SVD de l'étude par SAXS résolue dans le temps de cette transition indique qu'il n'y a pas d'état intermédiaire suffisamment stable et de taille ou de forme suffisamment différente de l'état initial et de l'état final pour pouvoir être distingué par cette méthode.



**Figure 32. Conditions d'expansion in vitro de la capside du phage HK97.** L'acidification à pH 4,4 des procapsides provoque leur expansion vers l'état EI-1 puis vers l'état EI-4. Un pH inférieur à 4 inhibe la transition de l'état EI-1- ou EI-2 vers l'état EI-4. La stratégie expérimentale permettant d'étudier cette transition consiste à acidifier à pH 3.9 un échantillon de procapsides puis à le neutraliser à pH 7. Figure adaptée de (Veesler, Quispe et al. 2012).

## 2- Présentation de la thématique de recherche

Les étapes de l'expansion de la capside du phage HK97 sont aujourd'hui bien caractérisées d'un point de vue cinétique, mais cette caractérisation n'a été effectuée pour aucune autre capside de phage. L'intérêt d'acquérir des données sur l'expansion des capsides

d'autres phages réside dans le fait que, malgré une très grande conservation de la structure tridimensionnelle des protéines majeures de capsides chez les phages caudés, le pontage inter-capsomères spécifique au phage HK97 pourrait faire de ce phage un modèle non représentatif des autres phages caudés. L'accumulation de données cinétiques permettrait ainsi de mieux caractériser les mécanismes de cette étape indispensable à la formation de tous les virions des *Caudovirales*. La production de la procapside P-2 nous permet d'étudier la cinétique de cette étape en vue d'obtenir des informations sur son mécanisme, notamment sur la coopérativité ou sur l'existence éventuelle d'état(s) intermédiaire(s). Cette procapside est produite par infection avec un mutant de T5, il s'agit donc d'un véritable état intermédiaire physiologique de l'assemblage du virion, composé de la protéine majeure de capside maturée et de la portale, alors que les travaux effectués sur la capside de HK97 ont été menés sur une procapside reconstituée à partir de la protéine majeure surexprimée chez *E. Coli*, qui ne contient pas la portale.

## **II- Etude cinétique de l'expansion**

### **1- Résumé de l'article**

Après avoir purifié la procapside P-2 produite à l'aide d'un mutant de T5 dont la terminase est inactivée, nous avons caractérisé sa taille et sa forme par SAXS et par cryo-microscopie électronique. Nous avons ensuite déterminé des conditions permettant son expansion *in vitro*, par acidification et baisse de la concentration en chlorure de sodium. Nous avons vérifié par SAXS et par cryo-microscopie électronique que la forme et la taille des capsides après expansion étaient celles des capsides des virions matures. L'analyse par

calorimétrie différentielle (DSC) des états compact et expansé de la capsid e a montré une très importante stabilisation de la capsid e. Il est à noter que cette stabilisation, du même ordre que celle observée lors de l'expansion de la capsid e du phage HK97, est ici obtenue sans pontage covalent entre les capsomères, par un « simple » réarrangement structural des protéines majeures de capsid e.

L'étude cinétique de ce processus par diffusion dynamique de la lumière (DLS) nous a ensuite permis de déterminer l'ordre de grandeur du temps caractéristique de cette étape. L'étude de la cinétique par SAXS résolue dans le temps (avec un stopped-flow) nous a permis de montrer la grande sensibilité de ce temps caractéristique vis-à-vis du pH. L'analyse par décomposition en valeur propre des séries de courbes SAXS au cours du temps montre que l'étape d'expansion est hautement coopérative et qu'il n'y a pas d'état intermédiaire dont la forme ou la taille diffère significativement des états compact et gonflé.

## **2- Article : two-state cooperative expansion converts the procapsid shell of bacteriophage T5 into a highly stable capsid isomorphous to the final virion head**

Olivier Preux<sup>1</sup>, Dominique Durand<sup>1</sup>, Alexis Huet<sup>1,2</sup>, Aurélie Bertin<sup>1</sup>, Claire Boulogne<sup>1</sup>,  
Jeannine Drouin-Wahbi<sup>1</sup>, Didier Trévarin<sup>3</sup>, Javier Pérez<sup>3</sup>, Patrice Vachette<sup>1</sup>, Pascale  
Boulanger<sup>1\*</sup>

<sup>1</sup> Institut de Biochimie et Biophysique Moléculaire et Cellulaire, Université Paris-Sud, UMR  
CNRS 8619, F- 91405 Orsay, France

<sup>2</sup> Department of Structural Biology, University of Pittsburgh School of Medicine,  
Pittsburgh, PA 15260, USA

<sup>3</sup> Beamline SWING, Synchrotron SOLEIL, F-91192 Gif sur Yvette, France

**\* Corresponding author:**

Pascale Boulanger

[pascale.boulanger@u-psud.fr](mailto:pascale.boulanger@u-psud.fr)

phone : +33 1 69 15 63 40

fax : +33 1 69 85 37 15

**Keywords:** bacteriophage T5 / capsid expansion / two-state transition / cryo-electron  
microscopy / Small Angle X-Ray Scattering (SAXS)

## Abstract

Capsids of dsDNA bacteriophages initially assemble into compact procapsids, which undergo expansion upon the genome packaging. This shell remodeling results from a structural rearrangement of head protein subunits. It is a critical step in the capsid maturation pathway that yields final particles capable to withstand the huge internal pressure generated by the packed DNA. Here we report the study of the expansion process of the large capsid (T=13) of bacteriophage T5. We purified the intermediate form prohead II, which is competent for packaging the 121 kbp dsDNA genome and we investigated its morphology and dimensions using cryo-electron microscopy and Small Angle X-Ray Scattering. Decreasing the pH or the ionic strength of prohead II solution triggers their expansion, converting them into thinner and more faceted capsids isomorphous to the mature virion particles. At low pH, prohead II expansion is a highly cooperative process, without any detectable intermediate. This two-state reorganization of the capsid lattice *per se* leads to a remarkable stabilization of the particle. The melting temperature of expanded T5 capsid is virtually identical to that of more complex shells that are reinforced by inter-subunit crosslinking (HK97) or by additional cementing proteins (T4). T5 capsid with its “simple” two-state conversion thus appears to be a very attractive model for investigating the energetics of the large-scale allosteric transition that takes place upon the genome packaging of dsDNA bacteriophages.

## Introduction

Tailed bacteriophages, which represent more than 95% of bacterial viruses<sup>1</sup>, have an icosahedral capsid containing a densely packed double-stranded DNA and a tail that serves to deliver the genome into the host cell. Their assembly into structurally well-defined and highly stable infectious particles proceeds through the assembly of the DNA filled capsid that is completed by addition of the tail. The long tail of *Myoviridae* and *Siphoviridae* is assembled in a separate pathway and connected to the mature capsid, while the tail proteins of short tailed *Podoviridae* are sequentially added to the capsid. Basic principles in assembly and maturation of icosahedral capsids of dsDNA bacteriophages have been determined from studies of small capsids (Triangulation number T=7) of phages HK97, P22,  $\lambda$ , SPP1, T7 and of the larger capsid of phage T4 by combining biological, biochemical and structural data (reviewed in Refs. 2, 3,4). Viral DNA is transported into a preformed procapsid. A scaffolding protein guides the correct assembly of this initial shell, which is built from several hundred copies of a major coat protein (two in the case of T4) arranged as pentamers at the vertices and hexamers at the faces of the icosahedron. A dodecameric portal protein forms the gate, at a unique vertex, through which DNA is packaged and released at the onset of infection. DNA transport into the procapsid is driven by the terminase, a two-subunit enzymatic complex that, together with the portal protein, constitutes the packaging machinery<sup>5,6</sup>. Structural and functional characterization of the terminase<sup>7; 8; 9</sup> and portal proteins<sup>10; 11; 12</sup> as well as *in vitro* studies of DNA packaging<sup>13; 14; 15</sup> is starting to shed light on the molecular mechanism of DNA translocation. DNA packaging is coupled to major structural rearrangements of the coat protein subunits leading to a two-fold volume expansion of the capsid and allowing it to accommodate the full genome. Head completion proteins then seal the portal and form the fastening point for the tail<sup>16</sup>. It is now becoming established that the capsid proteins of most tailed dsDNA bacteriophages (coat, portal, head completion) share

common folds despite a remarkable divergence in their sequences<sup>17</sup>, which basically explains the similarity in the mechanisms that govern capsid assembly.

*In vitro* studies combining cryo-EM and X-ray crystallography methods have provided a high-resolution description of the structural rearrangements of the capsid protein subunits associated to the expansion. The most complete data have been obtained with the portal-less procapsids of the lambdoid phage HK97, which can be assembled in *E.coli* cells from overexpression of the coat protein gp5 and the head maturation protease gp4. Gp4 specifically cleaves the N-terminal  $\Delta$ -domain of gp5 that plays the role of the scaffolding protein, yielding empty procapsids that can be matured *in vitro* by various chemical treatments<sup>18</sup>. HK97 procapsids swell in a two-state transition with no populated intermediate, as shown by time-resolved SAXS experiments.<sup>19</sup> This expansion is accompanied by thinning of the shell and involves twisting as well bending motions of specific helical and  $\beta$ -sheet regions of the gp5 subunits.<sup>20; 21</sup> Such a large-scale cooperative transition with partial refolding of the coat protein occurs in a similar way in phages T7<sup>22</sup> and P22<sup>23</sup>, which share with HK97 the same overall fold of gp5 protein. Expanded capsids are highly stable particles that sustain a pressure of ca. 60 atmospheres generated by the densely packed DNA.<sup>24</sup> Some mature capsids are strengthened by ultimate modifications, such as inter-subunit auto-catalytic crosslinking of the coat protein in HK97<sup>25</sup> or the addition of cementing outer capsid proteins in the case of Lambda<sup>26</sup> and T4<sup>27</sup>.

The in depth understanding of the structural and energetic determinants of the capsid maturation requires investigating diverse phage models, and in this respect, the large size capsid of bacteriophage T5 is an advantageous system. T5 icosahedral capsid contains a single chromosome of 121750 bp dsDNA. Its structure was solved at 20 Å resolution by cryo-EM and image reconstruction<sup>28</sup> (geometry T=13, 900 Å diameter) and a schematic of the T5 head assembly pathway has been recently established<sup>29</sup>. The first procapsid form, Prohead I,

is built from 775 copies of the head protein pb8 arranged as 11 pentamers and 120 hexamers, from the portal pb7 lodged at the twelfth vertex and from the head maturation protease pb11, which subsequently cleaves all protein subunits. The 458 residue-long capsid protein pb8 sheds its N-terminal scaffolding domain (159 residues), yielding pb8m (MM 32.9 kDa), while the 10 N-terminal residues from pb7 (43.8 kDa) as well as both N- and C-termini of pb11 are released (manuscript in preparation). This general protein trimming converts prohead I into prohead II with a molecular mass of 26.0 MDa. Packaging of the T5 genome into prohead II yields the expanded mature capsid that is then sealed by one head completion protein and decorated by the dispensable protein pb10<sup>30</sup>, which binds to the center of pb8 hexamers<sup>28</sup>.

Here we produced and purified T5 prohead II by using a mutant deficient in DNA packaging. The prohead II particle is free of any scaffolding domain and constitutes the assembly intermediate form competent for DNA packaging. We triggered its expansion *in vitro* by acidification or decrease of the salt concentration. We characterized the morphology and dimensions of prohead II and expanded capsid by Cryo-EM imaging and static SAXS measurements and we analyzed the swelling transition using time-resolved SAXS experiments. T5 prohead II converts in a highly cooperative way into an expanded capsid with the same size as the mature capsid resulting from *in vivo* DNA packaging. Expansion of the T5 large size capsid illustrates the all-or-nothing transition that has been previously proposed to be a ubiquitous structural reorganization of the capsid shell induced by DNA packaging.

## Results

### Characterization of compact prohead II and expanded capsid by cryo electron microscopy and static SAXS measurements

To investigate the expansion of T5 capsid we produced and purified prohead II as indicated in Materials and Methods. The protein composition of prohead II includes the processed head proteins pb8, portal pb7 and pb11 as attested by SDS-PAGE (Fig.S2). The change in capsid radius upon expansion is associated with an increase in hydrodynamic radius large enough to be easily measured by DLS. In this way, we determined that decreasing the concentration of NaCl below 100mM at neutral pH caused capsid expansion. Native proheads II stored in 20 mM Hepes buffer pH 7.6 containing 1M NaCl and the same sample dialyzed against 20 mM Hepes buffer pH 7.6, without NaCl were imaged by cryo-EM. The prohead II population contained homogeneous round-shaped particles that did not display any internal structure (Fig. 1a). They exhibited a shell thickness of  $\sim 50 \text{ \AA}$ , with inner and outer diameters of  $590 \text{ \AA}$  and  $690 \text{ \AA}$  ( $\pm 10$ )  $\text{ \AA}$  respectively. The dialyzed sample contained larger particles with a thinner wall, exhibiting the faceted morphology typical of mature capsid shells. These had a maximum diameter of  $930 (\pm 10) \text{ \AA}$  as measured between 2 vertices along the icosahedral 5-fold axis. The inner and outer dimensions measured between two faces along the 3-fold axis were  $790 \text{ \AA}$  and  $840 \text{ \AA}$  ( $\pm 10$ ) respectively, giving a shell thickness of  $\sim 25 \text{ \AA}$  (Fig. 1b). This full and homogeneous change in the capsid morphology attested that T5 prohead II could be converted into stable expanded particles by decreasing ionic strength. A small fraction of the capsid population in the starting sample (5%-8%) were expanded capsids of the same shape and size as the dialyzed ones, indicating that spontaneous expansion of a few prohead II occurred throughout the purification, storage or cryo-EM steps. Noteworthy, is the presence

in prohead II samples of bilayer vesicles (~ 15% of the population) with a diameter ranging from 200 Å up to about 1000 Å. These contaminating membrane vesicles arose from lysis of the host bacteria upon infection by T5AmN5 and could not be totally separated from proheads II during purification. Vesicles were still present in the dialyzed sample.

We used Small Angle X-ray Scattering (SAXS) to characterize bulk populations of prohead II and expanded capsids in solution. The SAXS pattern of a spherical shell is composed of a series of minima and maxima, the position of which is characteristic of the overall dimension of the particle, while the height of maxima and the depth of minima are characteristic of the sphericity of the particles and the homogeneity of the population. For both prohead II and expanded capsid the SAXS patterns are typical of a spherical shell (Fig. 2a); over the angular range of our study [ $0.0046 \text{ \AA}^{-1}$ ,  $0.13 \text{ \AA}^{-1}$ ] 9 and 15 minima are visible in the prohead II and expanded prohead patterns respectively. The first minimum is observed at  $q = 0.0101 \text{ \AA}^{-1}$  in the prohead II pattern and at  $q = 0.0085 \text{ \AA}^{-1}$  in the expanded capsid pattern; this shift is characteristic of an increase in the size of the particle.

The two distance distribution functions,  $p(r)$ , are shown in Fig. 2b. The regular increase in  $p(r)$  values with  $r$  followed by a sharp drop to zero is typical of hollow spheres with a diameter equal to  $D_{\max}$ . The  $p(r)$  of prohead II exhibits a low but clear tail extending out to 860 Å (Fig. 2b), betraying the presence in the sample of a small fraction (~ 6%) of expanded proheads already observed on electron micrographs. After subtracting the corresponding percentage of the expanded prohead  $p(r)$  profile the maximum diameter of prohead II particles could be estimated to be 720 Å, only slightly larger than the 690 Å estimate of cryo-EM. The corresponding value for expanded capsids was about 860 Å ( $\pm 10$ ) Å, significantly smaller than the value of 930 ( $\pm 10$ ) Å derived from cryo-EM for the diameter of *in vitro* expanded capsids. This significant discrepancy, not observed in the case of prohead II, is due to the faceted structure of the expanded capsid. The cryo-EM estimate was measured

across the 5-fold axis. In SAXS, the maximum distance measured is close to the distance between two opposite faces of the icosahedron while the larger distances between the tips of pentons around 5-fold vertices do not contribute enough to be clearly detected.

### **Cryo-EM imaging of the acid-induced expansion of prohead II**

To investigate the content of T5 capsid solution during maturation we used the ability of acidic pH to induce *in vitro* expansion of prohead II. Preliminary DLS measurements allowed us to determine the pH range that triggered complete expansion of prohead II in less than 30 minutes without causing significant damage to capsids. Complete expansion was obtained by reducing the ionic strength 2.5 fold and subsequent acidification by mixing with sodium acetate buffer as described in Materials and Methods. Thus expansion resulted from lowering the ionic strength and the pH value. Analysis of capsid samples expanded at low ionic strength (dialysis) or at low pH (mixing) gave similar values of hydrodynamic radius and polydispersity.

Prohead II samples were acidified at pH 4.30 and samples for cryo-electron microscopy were prepared 1, 5, 8 and 30 minutes after acidification. The proportion of expanded capsids increased with time as shown in Fig.1c, 1d and 1e. At 8 minutes, the majority of the 554 selected particles were compact proheads (69%) with only 26% of expanded capsids while, at 30 minutes, all of the 202 counted particles were expanded capsids. These two populations of particles displayed the shape, thickness and diameter characteristic of the native prohead II or of the expanded capsid observed at neutral pH (Fig.1a and 1b), with no detectable third state. Some particles with the same thickness (~25 Å) and angular shape as the expanded capsid but displaying at least one broken side (Fig. S3) can be noticed. This damage to the thin wall of expanded capsids may be due to the rapid pH

drop from pH 7.6 to ~ pH 4 used to trigger expansion and/or to the effect on capsids of sample preparation by cryo-electron microscopy imaging (disruptions induced by the surface tension of the thin film of ice). Time-resolved SAXS experiments (see next section) indicated that disruption of the capsid at acidic pH was a rare event in solution and was very likely an electron microscopy artefact. Interestingly, the population of membrane vesicles in the sample of prohead II decreased after the pH drop from ~ 15% to ~1% of the total number of particles (Fig. S3).

### **Time-resolved SAXS study of acid-induced prohead II expansion**

We performed time-resolved SAXS experiments to monitor the course of prohead II expansion and to check for possible intermediate states that could have gone unnoticed in electron micrographs. Using a stopped flow device (as detailed in Materials and Methods) we induced capsid expansion by acidifying prohead II samples at the pH values of 4.30, 4.26, 4.23 and 4.17. Eighteen SAXS patterns were collected for 1s at 60s intervals, starting 4 seconds after acidification. The first pattern was almost identical to the reference SAXS pattern of prohead II measured at neutral pH (Fig. S5). From the  $p(r)$  profile, we determined that the fraction of expanded capsid in prohead II, prior to acidification and at  $t=4s$  seconds, was  $\approx 6\%$ . The absence of vesicles in acidified prohead sample may explain the very small shift of minima and maxima to larger angles (Fig. S5).

The distance distribution functions  $p(r)$  computed from the first and the last frames yielded estimates of the outer diameters of 720 Å and 860 Å ( $\pm 10$ ) Å respectively, which are identical to the dimensions of prohead II and expanded capsids derived from the static measurements. For each set of spectra, the minima and maxima of the scattering curves exhibited a progressive shift to smaller angle over the course of the transition, indicating an

average increase of the particle size (Fig. 3). This shift was faster at pH 4.26 than at pH 4.30, suggesting a pH-dependence of swelling kinetics discussed below.

For all data sets at all pH values we observed isoscattering points (Fig. 3a-3b). This suggests that all curves are linear combinations of only 2 curves: the scattering patterns of prohead II and of the expanded capsid. Distance distribution functions  $p(r)$  were computed and scaled to the same integral value as explained in Material and methods. In all profile sets we observed a crossing point common to all curves (Fig. 3c-3d).

### **SVD analysis**

To confirm the two state transition hypothesis suggested by our cryo-EM analysis and our observation of isoscattering points in TR-SAXS data sets, we performed a Singular Value Decomposition (SVD) analysis of SAXS data. This mathematical analysis is based on the decomposition of the set of experimental patterns in a set of orthogonal basis vectors and yields the minimum number of vectors necessary to account for the entire data set within experimental uncertainties, *i.e.* the dimension of the data set (see the Material and Methods section for details). This number is equal to the number of different scattering species in solution.

We present in some detail the analysis of a data set recorded at pH 4.26. The mere examination of the singular values is informative: the first two singular values are 25.18 and 2.51; whereas the third and fourth are 0.12 and 0.066, and the last fourteen singular values lie in between 0.057 and 0.014. The third singular value thus appears to represent less than 0.5% of the first one. While the first and second eigenvectors exhibit a smooth dependence on  $q$  (Fig. 4a), as befits a scattering pattern or a linear combination thereof, the third and fourth eigenvectors exhibit sharp, seemingly random variations around zero (Fig. 4b), likely to

represent experimental noise. Their much higher intensity at lower angle is due to the much higher experimental intensity - and associated uncertainty - scattered at these angles. This interpretation is supported by the study of the time dependence of the projection of the experimental curves onto these basis vectors (Fig. 4c and 4d). The projections of the experimental curves onto the first and second vectors are monotonic functions of time and are smooth too. In contrast, the projections of the experimental curves onto the third and fourth basis vectors exhibit a random evolution with time.

Finally, if an ensemble of curves has a dimension  $L$ , the reconstruction of each curve as a linear combination of the first  $L$  eigenvectors using equation 2 with  $N=L$  should be identical to the experimental curve within experimental uncertainties. Fig. 5a displays one experimental curve recorded at pH 4.26 together with its reconstruction for  $L=2$ . The two curves appear to be indistinguishable. The agreement is most convincingly represented by the distribution of reduced residuals shown in Fig. 5b, which are randomly distributed between +2 and -2.

In conclusion, all mentioned criteria allow us to conclude that all curves recorded at pH 4.26 are linear combination of only the first two eigenvectors. The swelling transition is concerted, all-or-none, and only involves initial compact prohead II and expanded capsid, the fractional concentration of which increases with time. The same analysis was performed on data sets recorded at pH 4.23 and 4.30 and led us to the same conclusion of a concerted expansion.

In the case of data at pH 4.17, the third singular value is much larger with respect to the first one (4% vs 0.5%) and the time dependence of the projection of experimental curve on the third basis vector is not clearly random. This may be attributed to the occurrence of capsid breakage due to the very low pH, as suggested by the filling up of minima in the last frames

recorded at this pH. We may have reached here the lower limit of the pH-range compatible with capsid integrity.

### **Expansion kinetics**

The fraction of expanded prohead as a function of the time after acidification at pH 4.30, 4.26, 4.23 and 4.17 was determined from the TR-SAXS experiments in two different ways: using the pair distance distribution and the evolution of the projections of experimental curves on the second eigenvector in our SVD analysis (see Materials and Methods for details). These two methods gave identical results and the time evolutions derived from the SVD analysis are shown in Fig. S6. We believe these kinetic curves are meaningful over the first 600 seconds (ten first frames). We strongly suspect the later part is perturbed by artefacts due to the need to displace the sample in the beam after each frame, which caused the dilution mentioned in Materials and Methods. While we could correct this effect and while SVD analysis is independent of concentration, it appears that the effect on the kinetics are more deleterious. We therefore limit ourselves to semi-quantitative considerations. The most striking feature is the remarkable pH-dependence of the transition time scale. The transformation of half the capsid population is reached at around 550 s, 270 s, 180s and 110 s, at pH 4.30, 4.26, 4.23 and 4.17, respectively. The expansion phase is preceded by an apparent lag-phase with a pH-dependent characteristic time. The strong pH-dependence of the expansion rate may be the cause of our dilution artefact spoiling the later phase of the process. Indeed, despite the pushing buffer was supposed to have the same pH as the final buffer, this cannot be guaranteed within 0.01 pH unit. Therefore, no detailed kinetic analysis has been attempted.

## Stability of expanded T5 capsid

To address the question of whether *in vitro* transition of T5 prohead mimics the physiological process and yields the final conformation of T5 capsid, the thermal stability of prohead II and of the expanded capsid was investigated by differential scanning calorimetry. The resulting thermograms exhibited a unique transition at 77°C and 96°C for prohead II and expanded capsid respectively (Fig. 6). This spectacular increase in the melting temperature shows that expansion greatly stabilizes T5 capsid. Remarkably, the thermogram displayed no peak below 77°C, indicating that heating did not induce any significant conformational rearrangement of proheads II before denaturation and *a fortiori* did not trigger their expansion.

## Discussion

The comprehensive description of the assembly and maturation pathway of bacteriophage T5 capsid relies on the characterization of intermediate states that can be isolated *in vivo* or assembled from recombinant proteins. We have previously shown that assembly of T5 icosahedral T=13 proheads I can be carried out *in vitro*, starting from the major head protein with its scaffolding domain<sup>29</sup>. These structures display the radial organization of the scaffolding domains within the icosahedral structures and provide a morphologically faithful model of the physiological prohead I, even though they do not include the portal vertex. In this work, we have investigated the *in vitro* expansion of the physiological precursor prohead II obtained with the T5AmN5 mutant inactive in DNA packaging. Preparing the mg amount of highly purified prohead II required for TR-SAXS experiments was a challenge. Indeed, infection of bacterial cells with T5AmN5 mutant yielded fewer particles than overexpression of the capsid genes as reported for the portal-less

particles of phage HK97<sup>18</sup>. We thus used large volumes of bacterial lysates and carried out two chromatography steps to achieve a high degree of purification, while recovering substantial amounts of prohead II. Although some contaminating bacterial membrane vesicles were still present in prohead II samples, these did not hamper the characterization of the compact and expanded capsids by electron microscopy. Their contribution to the SAXS patterns obtained from static and time-resolved experiments had no significant impact either on the analysis of the expansion process (see below).

### ***In vitro* expansion of T5 prohead II yields the mature capsid shell**

The purified T5 prohead II samples included almost entirely unexpanded particles that could retain their compact state as long as they were kept in the presence of high salt concentrations (> 500 mM NaCl). The presence of a very minor fraction of expanded capsids (typically 3-5% of the population, as attested by DLS analyses) reflects the intrinsic instability of T5 proheads, which are metastable intermediates primed to gain a lower energy conformation. Most probably all proheads II were in the compact conformation in the cells and underwent expansion during *in vitro* handling. The same observation was done with the prolate prohead of bacteriophage T4 (T=13, Q=21), the only other large size procapsid that has been previously purified<sup>31</sup> as well as with the smaller size procapsid of phage  $\lambda$ <sup>32</sup> or portal-less prohead II of HK97<sup>33</sup>.

We induced total expansion of prohead II in a controlled manner by modifying the ionic environment of the capsid protein subunits. The two methods used – pH lowering or ionic strength decrease - yield expanded capsids whose faceted morphology and dimensions are very close to those of the capsids of T5 virions containing the genome<sup>28</sup>. The swelling transition involves an increase in the shell internal diameter from *ca* 590 to *ca* 790 Å, which

leads to  $\approx 2.4$ -fold increase of the capsid volume with a correlative thinning of the shell from 50 Å to 25 Å. This coordinated swelling, thinning and angularization of the capsid shell has been previously reported for several phages, among which HK97<sup>18</sup>, T7<sup>22</sup>, T4<sup>34</sup> and P22<sup>23</sup> that all built their capsid shell from the same canonical HK97 gp5-like protein. Such remodeling of the capsid architecture results from structural modifications of the capsid protein subunits and from their rearrangement within the shell, which take place *in vivo* upon DNA packaging. It is thought to be a general mechanism for capsid maturation of dsDNA bacteriophages. In view of the good agreement between the dimensions of our empty expanded particles and the T5 mature capsids we believe that the *in vitro* induced transition studied here reflects what happens upon the physiological expansion process. As sequence similarities and structure predictions based on cryo-EM and image reconstruction indicate that the T5 capsid protein pb8 also adopts the gp5-fold<sup>28</sup>, we propose that T5 prohead II undergoes expansion through similar movements of these common protein subunits.

### **Kinetics of the swelling transition**

We have seen that the later part of kinetics was altered by the consequence of the dilution effect caused by the requirement of bringing fresh solution into the SAXS measuring cell before each frame. This was not necessary in the case of HK97 where 60 or 90 frames were collected on the same sample with no mention of detectable radiation damage<sup>19</sup>. This seems to suggest that our T5 capsid exhibited a much higher sensitivity to radiation damage. However, using estimates for various experimental parameters given in the paper by Lee et al.<sup>19</sup> and in the instrumental paper reporting on the multi-layer monochromator<sup>35</sup> the number of photons per capsid appears to be of the same order of magnitude for one frame in our experiments and 90 frames in HK97 measurements. The apparent higher sensitivity in our case is essentially due to the much lower number of capsids and to the significantly smaller

beam-cross section at the sample level at the SWING beamline. A marked pH-dependence of the expansion rate is the most striking feature emerging from the very elementary analysis we deemed warranted. Indeed, we observe a 5-fold increase in speed for a 0.13 pH unit reduction in the pH of the solution, *i.e.* a 25% increase in proton concentration. This is similar to observations made for HK97 swelling transition from prohead II to EI-I and suggests a similar interpretation in terms of protonation of titrable residues, most likely aspartate and glutamate residues. A high cooperativity in proton binding is likely to underpin the strong pH-dependence. The second characteristic of the transition kinetics, different from HK97 observations, is the existence of a lag phase, the duration of which is also pH-dependent. This suggests that a preliminary phase, following proton titration takes place, possibly involving local conformational change of pb8, which makes capsids competent for expansion that subsequently occurs in an asynchronous way among the capsid population. As already mentioned in the study of HK97 kinetics <sup>19</sup>, little information can be retrieved from kinetic measurements in bulk solution, as compared with kinetic measurements on individual capsid particles. The complexity of the kinetics, unexpected in the case of a two-state concerted transition, probably reflects heterogeneity in either or both prohead II and expanded capsid populations associated with the large number of possible distributions of titrated groups over the several hundred molecules forming the particle. This heterogeneity can be energetic but also locally conformational without any difference in the overall size of the particle as monitored by both cryo-electron micrographs and SAXS patterns.

### **Expansion matures T5 prohead II into a highly stable capsid**

Bacteriophage capsids are extraordinarily robust macromolecular assemblies. Indeed, once the virion genome has been packaged by the ATP-powered terminase <sup>6</sup>, tightly packed DNA exerts a strong internal pressure on the capsid wall estimated to be *ca.* 60 atm.<sup>15</sup>.

Furthermore, capsids have to withstand possible external hostile conditions in the different ecosystems inhabited by their hosts, such as extreme pHs, presence of detergents or proteases, variable temperature. Virions have developed various strategies to produce capsids with the required properties. HK97, for which successive intermediate states and mature capsid are available at near atomic resolution, uses a unique strategy. gp5 protein is endowed with the autocatalytic capability of joining each subunit to two of its neighbors by a covalent link between the side-chains of Lys 169 and Asn356<sup>36</sup>. This ensures the high stability of the final product of a multi-step maturation process, as witnessed by DSC measurements showing that HK97 Prohead II and EI-I denature at very similar temperatures, 86 and 83 °C respectively, as compared to 96°C for the cross-linked Head II<sup>37</sup>. Other phages such as T4<sup>38</sup> and lambda<sup>26</sup> possess more complex shells that are reinforced by addition of cementing proteins at a late stage of their maturation process. For instance, phage T4 empty expanded capsids, in which gp23 capsomers are bonded by the Soc cementing protein, have a reported T<sub>m</sub> value of 88°C, 5°C higher than in the absence of soc<sup>31</sup>. Finally, phages T7 and P22 are well-studied representatives of another group of viruses that use simple capsid expansion to achieve their mature stable conformation. It has been recently shown using cryoEM and 3D image reconstruction that the molecular reorganization of T7 hexamers upon maturation involves refolding events of the T7gp10 protein similar to those of HK97gp5<sup>22</sup> (tilting, in-plane rotations and radial expansion of their subunit domains). The P22 capsid is assembled from a gp5-like protein that includes an extra Ig-like domain (Telokin domain) forming external protrusions at the capsid surface. This extra domain is required for procapsid assembly but not for stabilization of the expanded shell<sup>39; 40</sup>. Our results clearly show that T5 belongs to this latter group, with its apparently simple and direct maturation process: structurally, since no stable intermediate state can be detected between prohead II and the expanded capsid, and energetically, as the latter exhibits a great stability with a T<sub>m</sub> of 96°C identical to that of

HK97 Head II and 20°C higher than that of prohead II. Although mature T5 capsid is decorated by the accessory protein pb10, pb10 is dispensable for phage development<sup>30</sup>. This strongly suggests that T5 gains its final stability through the reorganization of pb8 subunits that takes place upon expansion with no need for additional chemical modification or accessory proteins. Whether the portal controls the prohead expansion and contributes to stabilize the expanded capsid remains an open question, which has been addressed for phages T4<sup>41</sup>, T7<sup>42</sup> and P22<sup>39</sup>. The expansion of the physiological T5 prohead II reflects the involvement of all the capsid components including the recently identified head protease. Further investigation using modified prohead II is required to determine the role of pb7 and pb11.

## **Conclusion**

It is now well established that all dsDNA bacteriophages use fundamentally common structural elements and similar assembly pathways to form highly stable infectious particles capable of replicating within their target host in a broad variety of ecosystems. However much remains to be done to unravel the detailed molecular mechanisms that control the maturation of bacteriophage capsids. Our results on T5 support the idea that the one-step expansion represents a widespread strategy to direct the maturation of a large, T=13 as well as of smaller T=7 capsids and to strengthen the capsid shell after DNA packaging. T5 is a model of choice for investigating how a very “simple” two-state cooperative expansion converts the prohead shell into a highly stable capsid isomorphous to the final virion head.

## **Materials and Methods**

### **Production and purification of prohead II**

Empty prohead II was produced from infection of *E. coli* F cells by T5stAmN5 amber mutant lacking functional terminase. Due to the absence of DNA packaging, capsid assembly stalls after proteolytic maturation of prohead I into prohead II by the protease pb11 (Ref. 29 and manuscript in preparation). Prohead II was purified by precipitation with PEG, followed by centrifugation in glycerol gradients and anion exchange chromatography essentially as previously described<sup>29</sup>. In order to ensure high purity of the large amounts of prohead II required for SAXS experiments, we modified the initial protocol and added a second step of anion exchange chromatography. Prohead II eluted from the final chromatography on HiTrapQ HP column (GE Healthcare) in 50 mM Tris Buffer pH 7.6, 400 mM NaCl. It is noteworthy that metastable T5 prohead II have a spontaneous propensity to expansion, which was circumvented by using freshly purified particles and a high salt concentration storage buffer. Prohead II were dialyzed against 50mM Hepes buffer (pH 7.6), 1M NaCl, and finally concentrated to 2.5 g/L by centrifugation through Amicon Ultra-4 30K filter units and stored at 4°C.

### **Expansion of prohead II**

Prohead II expansion was induced in two different ways. In static SAXS experiments, prohead II sample stored in 50mM Hepes buffer pH 7.6, 1M NaCl was dialyzed against a 50mM Hepes buffer pH 7.6 for 48-72 hours. In time-resolved experiments, prohead II expansion was triggered by a rapid drop in pH, with the sample previously diluted 2.5 fold into distilled water to yield a final salt concentration of 400 mM NaCl. We checked by static SAXS measurement that this limited desalting step did not induce significant expansion of the

capsids over a few hours range. The acid-triggered transition was initiated by mixing the diluted prohead II sample with an equal volume of 160 mM sodium acetate buffer at various pH values. Final buffer solutions were 80 mM sodium acetate, 10mM Hepes, 200mM NaCl at pH values, 4.30, 4.26, 4.23 or 4.17.

### **Cryo-electron microscopy**

4  $\mu$ L droplets of capsid samples were deposited on holey grids (Quantifoil R2/2) previously glow discharged for 30 seconds using a sputter coater (Quorum Technologies) and plunge-frozen in liquid ethane using a commercial apparatus (EMGP, Leica). Imaging was performed using a 200kV FEG (Field Emission Gun) microscope (JEOL, 2010F) at Genopole (Institut Pasteur). Images were collected using a low dose system (MDS, JEOL) with a CCD 4K X 4K camera (ultrascan, Gatan) at a magnification of 50,000 and about -1  $\mu$ m defocus. At such a magnification, the pixel size was 2.2 Å. The samples of prohead II (2.7 g/L in 50mM Hepes buffer pH 7.6, 1M NaCl) as well as dialyzed expanded capsid (1.9 g/L in 50mM Hepes buffer) were not diluted prior to vitrification. For imaging acidified prohead II over the course of expansion process, acidification of the sample was triggered as described above (final concentration 0.5 g/L, final pH=4.30). Aliquots were collected at given times (1, 5, 8 and 30 minutes) and the swelling transition was arrested by vitrification into liquid ethane.

### **Solution Small-Angle X-ray Scattering**

SAXS experiments were carried out on beamline SWING at the SOLEIL synchrotron-radiation facility (St-Aubin, France). The incident beam energy was 12.00 keV, and the sample to detector (Aviex CCD) was set to 4415 mm. The useful scattering vector range was

$0.0046 < q < 0.13 \text{ \AA}^{-1}$ , where  $q = \frac{4\pi \sin \theta}{\lambda}$ ,  $\lambda$  is the wavelength and  $2\theta$  the scattering angle.

During exposure, the solution was contained in a quartz capillary of 1.5 mm in diameter, within a vacuum path from the sample to the detector. For static measurements the solution was continuously circulated through the capillary using the local automated pump system. The maximum exposure time causing no detectable radiation damage was determined by comparison of successive frames in a preliminary measurement without circulation. Several frames (typically 30) of 1-s each separated by a 0.5-s pause were successively recorded for both sample and buffer.

SAXS data were normalized to the intensity of the transmitted beam, put on an absolute scale using the scattering by water<sup>43</sup> and background subtracted using the local dedicated program FoxTrot.

### **Time-resolved measurements**

Time-resolved SAXS experiments were carried out using a stopped-flow device (SFM-400 from BioLogic company) inserted into a locally developed circuit to meet the specific requirements (described below) of T5 large capsids as shown in Fig. S1 (Supplemental Material). T5 capsids are large size particles with icosahedral symmetry (diameter of  $\sim 690 \text{ \AA}$  for prohead II,  $\sim 840 \text{ \AA}$  for expanded prohead). Scattering patterns therefore exhibit series of deep minima and maxima, the position of which is defined by the particle size while the depth is strongly dependent on the size monodispersity of the solution. In order to record the first minima without instrumental distortion and also to be able to calculate the distance distribution function  $p(r)$ , data at very low scattering angle were required ( $q_{\min} D_{\max} \leq \pi$ , where  $D_{\max}$  is the particle diameter). Using the conventional

stopped-flow measuring capillary required the introduction of two windows in the beam upstream and downstream of the “in air” capillary that caused an unacceptably high background scattering level in the smallest angular range. We therefore connected the stopped-flow apparatus, installed sideways to the beamline, to the SAXS measuring cell contained in an evacuated vessel. This resulted in a dramatic reduction in background level at these scattering angles making the experiment feasible.

The solution of prohead-II (2.5 g/L in 50 mM Hepes (pH 7.6), 1 M NaCl) was first manually diluted 2.5 times with water and introduced in syringe 1. Particle expansion was initiated by rapid mixing of the prohead II solution with 160 mM sodium acetate buffer at the appropriate pH (syringe 2) to yield the desired final pH in the mixed solution. Preliminary Dynamic Light Scattering measurements had shown that the range of pH that induced the capsid expansion is narrow ( $4.1 \leq \text{pH} \leq 4.3$ ): using a lower pH-value causes capsid destruction while a higher pH-value does not induce expansion in a time scale compatible with SAXS measurements. Expansion of the majority of capsids within the above pH-range was observed after 2 to 15 minutes. TR-SAXS experiments were performed at four pH values: pH 4.17, pH 4.23, 4.26 and 4.30. The precision of pH determination was 0.02 pH unit but the accuracy of absolute pH values was 0.05 pH unit.

During each experiment we mixed 200  $\mu\text{L}$  of prohead II sample and 200  $\mu\text{L}$  of acetate buffer during 6 seconds. This slow mixing prevented building up of too high a pressure in the quartz capillary. A series of time-resolved measurements were made starting 4s after the end of mixing. For experiments at pH 4.30, 4.26 and 4.23, 18 1s-exposures were recorded every minute. At pH 4.17, 18 1s-exposures were recorded every 30s. Radiation damage tests had shown that capsids swollen at low pH cannot tolerate an X-ray exposure of more than 1 s. In order to prevent capsid degradation due to radiation damage 6  $\mu\text{L}$  of final Hepes/acetate buffer were pushed using syringe 4 between each measurement so as to bring fresh (*i.e.*

unirradiated) low pH sample into the beam. The number of possible successive shots before the mixed solution finishes was determined in preliminary measurements using a tryptophan solution and the UV-Vis Fiber Optics system which allows to measure the sample absorption directly on the quartz capillary<sup>44</sup>. After about ten moves leaving the absorbance unchanged a gradual decrease in absorbance was observed for each successive move due to dilution with the buffer used to displace the solution in the circuit.

### Data processing and internal pattern normalization

The distance distribution function  $p(r)$  is the distribution of the distances between any pair of scattering elements within the particle. It is the Fourier transform of the scattering curve  $I(q)$ :

$$p(r) = \frac{r^2}{2\pi^2} \int_0^\infty \frac{I(q) \sin(qr)}{qr} q^2 dq$$

In this study the  $p(r)$  function was calculated from the measured intensity  $I(q)$  in the  $q$ -range  $[0.0054 \text{ \AA}^{-1}, 0.1285 \text{ \AA}^{-1}]$  using the program GNOM<sup>45</sup>. The radius of gyration can be derived from  $p(r)$  function using the following expression:

$$R_g^2 = \frac{\int_0^{D_{\max}} r^2 p(r) dr}{2 \int_0^{D_{\max}} p(r) dr}$$

The scattered intensity extrapolated to the origin ( $q=0$ ) can also be expressed using the  $p(r)$  function:

$$I(0) = 4\pi \int_0^{D_{\max}} p(r) dr$$

$I(0)$  is proportional to the scattering particle molecular mass  $M$  and concentration  $c$  :

$$I(0) = \frac{cM}{N_A} \left( \frac{m_p N_A}{M} - \rho_s \bar{v}_p \right)^2$$

where  $N_A$  is the Avogadro number,  $c$  is the mass concentration of the particle,  $M$  its molecular mass,  $m_p$  its number of electrons,  $\bar{v}_p$  is the particle partial specific volume and  $\rho_s$  the electronic density of the solvent.  $I(0)$  was observed to be constant over ten frames before decreasing during the rest of the transition, in agreement with the dilution tests observed with the tryptophan solution. Since the swelling transition under study is a single particle process and, therefore, is independent of sample concentration, we corrected for the effect of dilution by rescaling the  $p(r)$  to a constant integral. Note that the solution was pushed using final buffer, therefore not changing the pH during dilution.

### Singular Value Decomposition (SVD) analysis

The SVD analysis is a very general method used to determine the number of independent species distinguishable by a given experimental technique during a transition process, without prior knowledge of the patterns of the individual species. First applied to the analysis of SAXS patterns by Fowler et al.<sup>46; 47</sup>, it has since been often used in SAXS-monitored conformational transitions of allosteric enzymes<sup>48</sup>, folding and unfolding processes<sup>49; 50; 51</sup> and viral structural transitions<sup>19; 52</sup>.

The set of  $N$  background corrected experimental curves, of  $M$  points each, are considered as the elements of a matrix  $\underline{\underline{\mathbf{A}}}(M \times N)$ . The  $k^{\text{th}}$  column of this matrix represents the  $k^{\text{th}}$  scattering frame  $I_k(q)$  normalized by the intensity at zero angle  $I_k(0)$ . This matrix can be decomposed as follows :

$$\underline{\underline{\mathbf{A}}}(M \times N) = \underline{\underline{\mathbf{U}}}(M \times N) \underline{\underline{\mathbf{w}}}(N \times N) \underline{\underline{\mathbf{B}}}(N \times N) \quad (1)$$

where the matrix  $\underline{\underline{\mathbf{U}}}$  consists of N columns  $\underline{u}_j(q)$  which are the orthonormal basis vectors, the diagonal non-negative matrix  $\underline{\underline{\mathbf{w}}}$  contains the singular values  $w_j$  of the decomposition, arranged in decreasing order, and  $\underline{\underline{\mathbf{B}}}$  contains the normalized projections  $b_j^k$  of the  $k^{\text{th}}$  experimental curve on the basis vector  $\underline{u}_j$ . Following Eq. (1), each experimental pattern can be expressed as :

$$I_{\text{exp}}^k(q_i) = \sum_{j=1}^N w_j b_j^k \underline{u}_j(q_i) \quad (2)$$

In most experimental situations, only the first L terms significantly contribute to the reconstruction of the experimental scattering profiles in Eq. (2). The  $w_j$  ( $j > L$ ), which have very small values while the corresponding  $b_j^k$  display random fluctuations around zero as a function of  $k$ , only account for measuring noise. L represents the number of independent scattering curves distinguishable by SAXS during the transition or, in other words, the minimum number of structural states involved in the swelling process.

The number L is determined by examining the evolution with time of the projections  $b_j^k$  of the experimental curves on the basis vector  $\underline{u}_j$ , the shape of the basis vectors  $\underline{u}_j$ , and checked by calculating the following chi-square value, normalized by the experimental error  $\sigma^k(q_i)$  :

$$\chi_p^2 = \frac{1}{(NM - 1)} \sum_{k=1}^N \sum_{i=1}^M \left[ \frac{(I_{\text{exp}}^k(q_i) - I_{\text{calc},p}^k(q_i))^2}{\sigma^k(q_i)^2} \right] \quad (3)$$

where  $I_{\text{calc},p}^k(q_i)$  is the best reconstruction of the  $k^{\text{th}}$  experimental curve  $I_{\text{exp}}^k(q_i)$  using the first p components in Eq.2. L is the minimum value of p for which  $\chi_p^2 \leq 1$ . The final and best control is provided by the examination of all reduced residual distributions obtained for  $p=L$ :

$$R_L^k(q_i) = \frac{I_{\text{exp}}^k(q_i) - I_{\text{calc},L}^k(q_i)}{\sigma^k(q_i)}$$

which should be randomly distributed between -2 and +2 for all scattering angles.

Let us underline that the basis vectors  $\underline{u}_j(q)$  are not the scattering curves of the L independent species or structural states.

The analysis was performed using a special version of the SVDPLOT program from the ATSAS suite<sup>53</sup> providing the reconstructions  $I_{\text{calc},p}^k(q_i)$  of all experimental curves for any p value (courtesy of V. Volkov). Analysis was restricted to the q-range [0.0046 Å<sup>-1</sup>, 0.0391 Å<sup>-1</sup>] so as not to introduce noisy intensities at larger scattering angle.

### **Quantification of prohead II and expanded prohead population during the expansion process**

The time dependent fractions of prohead II and expanded prohead after acid-induced maturation were determined in two different ways, using the pair distance distribution functions and the projections onto the second basis vector in SVD analysis respectively. The last frame of the TR-SAXS data set recorded at pH 4.26 is almost superimposable to the scattering pattern of the expanded prohead (Fig. S4). Therefore, we considered that the solution contained 100% expanded proheads and used the scattering pattern as a reference for scaling.

The p(r) function of a mixture is a linear combination of the p(r) functions of the individual components. Since only two components are involved in the expansion transition

of T5 capsid, a direct measurement of the expanded capsid proportion for the  $i^{\text{th}}$  frame  $F_{\text{exp},i}$  can be derived from the set of  $p(r)$  profiles. This takes the particularly simple expression:

$$F_{\text{exp},i} = \frac{\int_{r_1}^{r_2} p_i(r) dr}{\int_{r_1}^{r_2} p_{\text{end},\text{pH}4.26}(r) dr}$$

in the case where  $r_1$  is larger than the initial prohead II maximal diameter, so that both integrals have only contributions from expanded capsids. We used  $r_1=720 \text{ \AA}$ ,  $r_2=860 \text{ \AA}$  while  $p_{\text{end},\text{pH}4.26}(r)$  is the distance distribution function of the last frame of the experiment at pH 4.26.

The use of the time dependence of the projection onto the second basis vector is a model independent approach to quantify the time evolution of the fraction of expanded capsids. This fraction was computed using the projections  $b_i$ ,  $b_1$  and  $b_{\text{end}}$  of the scattering pattern onto the second basis vector at a given time  $i$ , at the beginning (first frame) and at the end of the transition:

$$F_{\text{exp},i} = \frac{(b_i - b_{\text{end}}) * (F_{\text{exp},\text{end}} - F_{\text{exp},1})}{(b_{\text{end}} - b_1)} + F_{\text{exp},\text{end}}$$

The fractions at the beginning  $F_{\text{exp},1}$  and at the end of the transition  $F_{\text{exp},\text{end}}$  were calibrated using the pair distance distribution approach.

## Differential Scanning Calorimetry

DSC was carried out using a MicroCal model VP-DSC. Scans were performed at a rate of  $1^\circ\text{C}/\text{minute}$  for temperatures ranging from  $20$  to  $100^\circ\text{C}$ . Prohead II ( $0.4 \text{ g/L}$ ) and capsid expanded by the dialysis protocol ( $0.4 \text{ g/L}$ ) (see above) were dialyzed in  $25 \text{ mM}$  sodium phosphate buffer, pH 7.6,  $200 \text{ mM}$  NaCl. All solutions were thoroughly degassed

before recording spectra. Each measurement was preceded by a baseline scan with the buffer. The heat capacity of the solvent was subtracted from that of the protein sample before analysis.

### **Acknowledgments**

We are very grateful to V. Volkov (Inst. of Cryst., Moscow) for the SVDPLOT developments and to M. Rocco (ISTGE, Genoa) for fruitful discussions on the stopped-flow use and the design of the complete mixing circuit. We are grateful to M. Aumont-Nicaise for the DSC facilities from the IBBMC/ImaGif platform. This work was supported by the French MENRT (Ph. D. fellowship to O.P.) and the French National Research Agency (post-doctoral fellowship to A.H.).

## Figure Captions

### **Figure 1. Cryo-electron microscopy analysis of prohead II and expanded capsids.**

(a) Cryo-electron micrograph of a prohead II sample (concentration: 2.5g/L), showing proheads II (1) and one membrane vesicle (2). Vesicles have diameters ranging from 200 to 1000 Å and are distinguishable from capsids by their non-regular shape and the presence of two membranes. (b) Cryo-electron micrograph of a prohead II sample (concentration: 1.9 g/L) dialyzed against 20 mM Hepes buffer (pH 7.6), without NaCl, showing expanded capsids (3) and vesicles (2). (c) and (d): Cryo-electron micrographs of a prohead II sample acidified to pH 4.30 in 200 mM NaCl, as described in Materials and Methods. The expansion process was arrested by vitrification of the samples after 8 (c) and 30 (d) minutes. Only prohead II (1) and expanded capsids (3) were observed without any detectable intermediate states. No vesicle is seen anymore. (e) Fractions of prohead II (blue) and expanded capsid (red) upon acidification. “Compact” and “expanded” correspond to the native proheads sample and dialyzed capsids sample respectively. The total number of capsids counted is approx. 2900.

### **Figure 2. SAXS patterns and distance distribution functions $p(r)$ of prohead II and expanded capsid.**

(a) SAXS patterns of prohead II (blue) and expanded prohead (red). Prohead II sample was diluted 2.5 folds into water at a final concentration of 1mg/l in 20 mM Hepes pH7.6, 400 mM NaCl. Expansion was triggered by dialysis of the prohead sample against a 20mM Hepes buffer pH 7.6, as described in Material and Methods. (b) distance distribution functions  $p(r)$

of prohead II (blue) and expanded capsid (red) computed from the corresponding scattering patterns using GNOM<sup>45</sup>.

**Figure 3. Time resolved SAXS patterns and distance distribution functions  $p(r)$  of prohead II during acid-induced expansion.**

Prohead II sample (concentration: 1g/l) was mixed with an equal volume of acetate buffer to a final concentration of 80mM acetate, 10 mM Hepes, and 200 mM NaCl at a final pH of 4.26 (a) or 4.30 (b) to initiate the transition. SAXS patterns were recorded during 1s at 60s intervals. In both graphs (a) and (b), the first curve (in blue) corresponds to the capsids acidified during 4s. The red curve corresponds to capsid acidified during 964s (a) and 1024s (b). The distance distribution functions  $p(r)$  are derived from the experimental scattering patterns at pH 4.26 (c) and 4.30 (d).

**Figure 4. Singular value decomposition of one experiment at pH 4.26.**

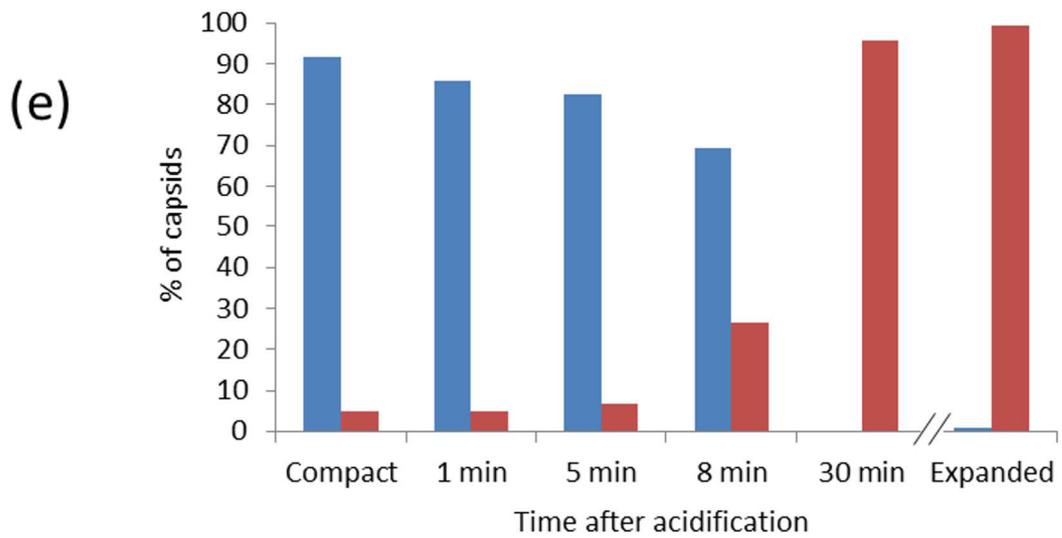
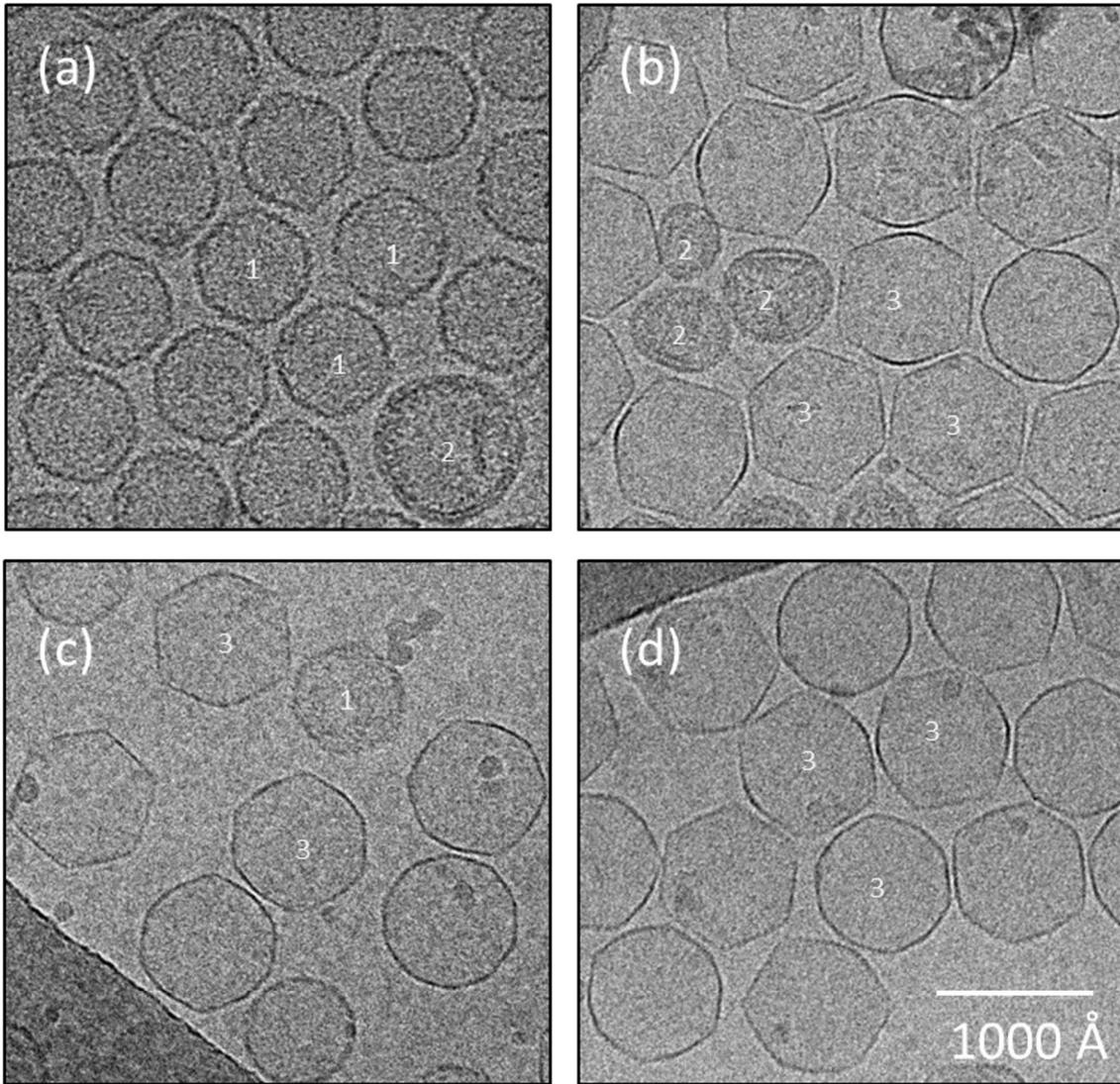
The first (continuous line) and second (broken line) basis vector are shown in (a), the third (continuous line) and fourth (broken line) basis vector are shown in (b). (c) Projections of the experimental curves onto the first (continuous line), second (broken line) basis vectors. (d) Projections of the experimental curves onto the third (continuous line), fourth (broken line) basis vectors.

**Figure 5. Restoration of time resolved patterns using only the first two eigenvectors from the SVD analysis.**

(a) The circles corresponds to the experimental SAXS scattering pattern recorded 8 minutes after the pH drop down to pH 4.26. The continuous line corresponds to the restored intensity using the first two basis vectors. (b) Curve of the restored intensity minus the experimental intensity divided by the experimental uncertainty (reduced residuals), as a function of the momentum transfer.

**Figure 6. Calorimetric recording of prohead II (continuous line) and expanded capsid (broken line).**

A sample of freshly purified proheads II was dialyzed against 20 mM Hepes buffer pH 7.6 during 24 hours to induce expansion. Both native proheads and expanded ones were then dialyzed against 25 mM sodium phosphate buffer pH 7.6, 200m M NaCl for calorimetric recording. The dramatic fall of the excess heat capacity after  $T_m$  is explained by protein aggregation following capsid dislocation. Because of this aggregation, thermodynamic parameters such as  $\Delta H_m$  and  $T_m$  cannot be determined. However  $T_m$  can be approximated by the temperature at which the maximum of each curve is reached. Both native and expanded prohead samples were checked by negative staining EM: the expanded sample included 3% of proheads II, which explains the small peak at around 77°C in the thermogram of expanded proheads.



**Figure 1. Cryo-electron microscopy analysis of prohead II and expanded capsids.**

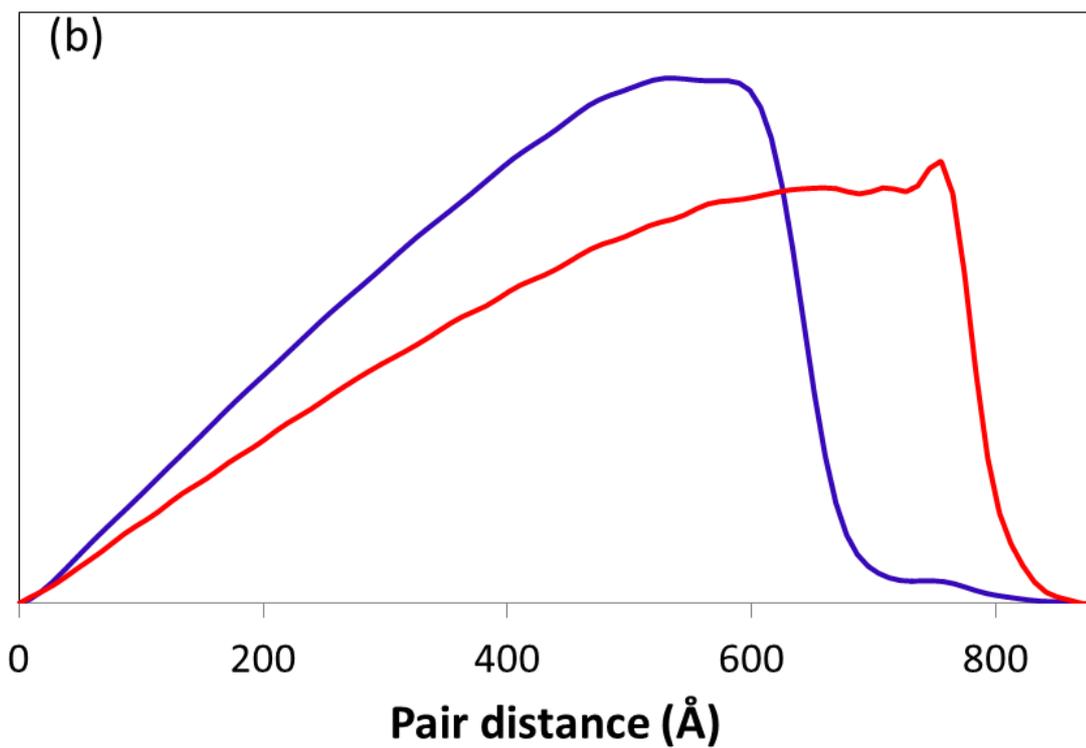
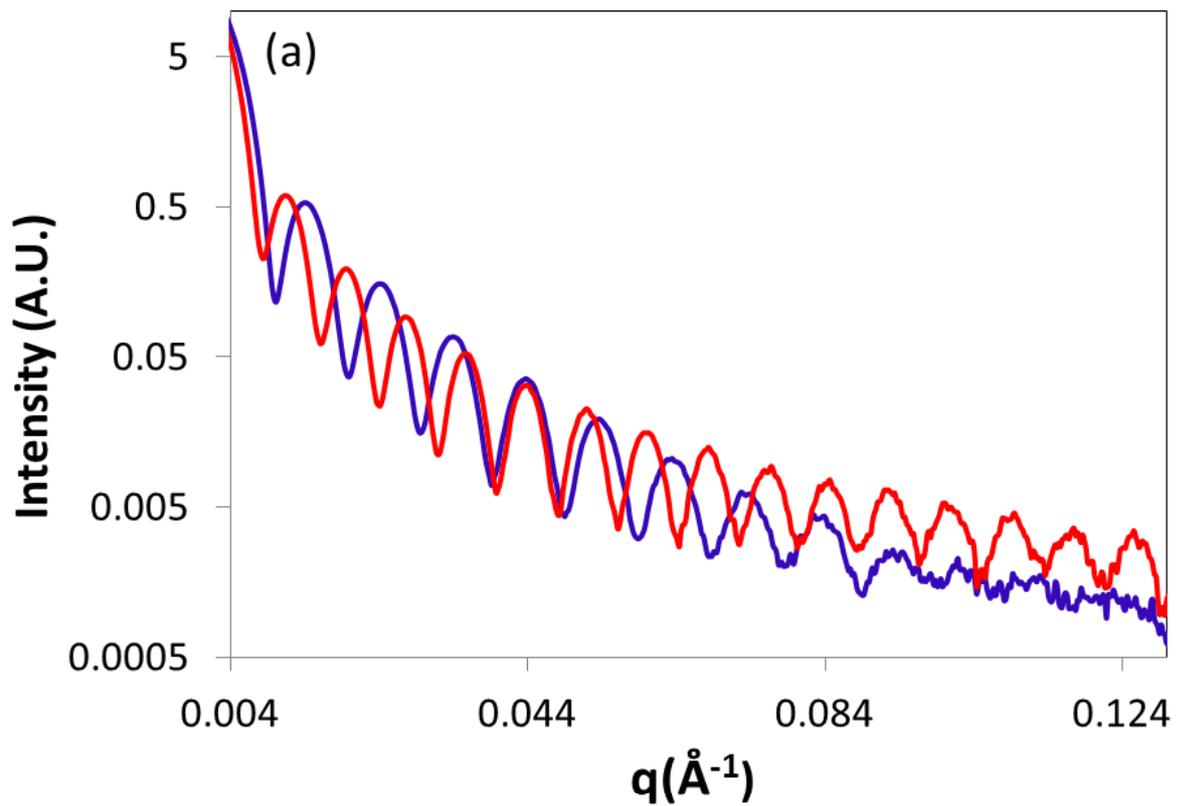
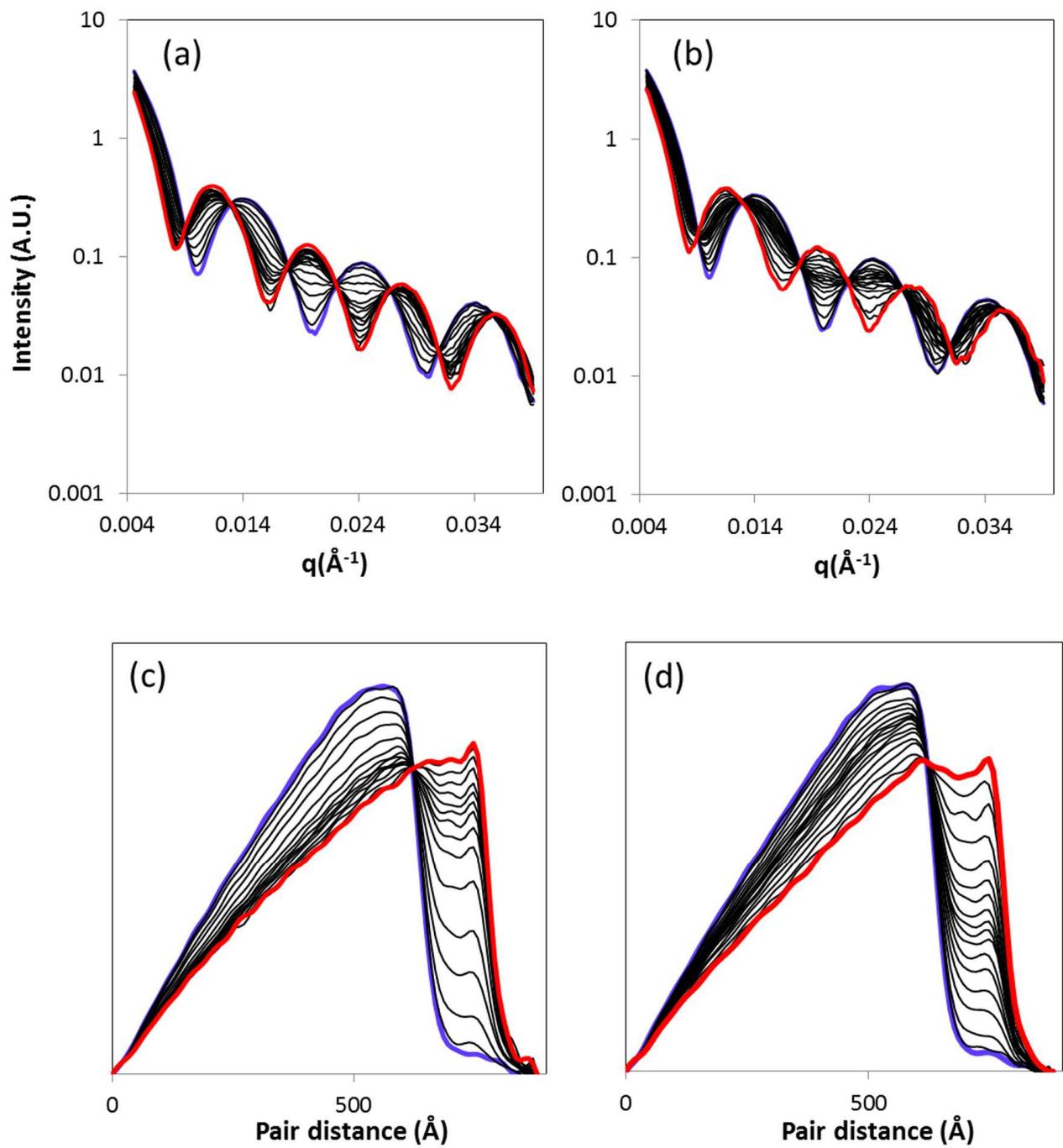
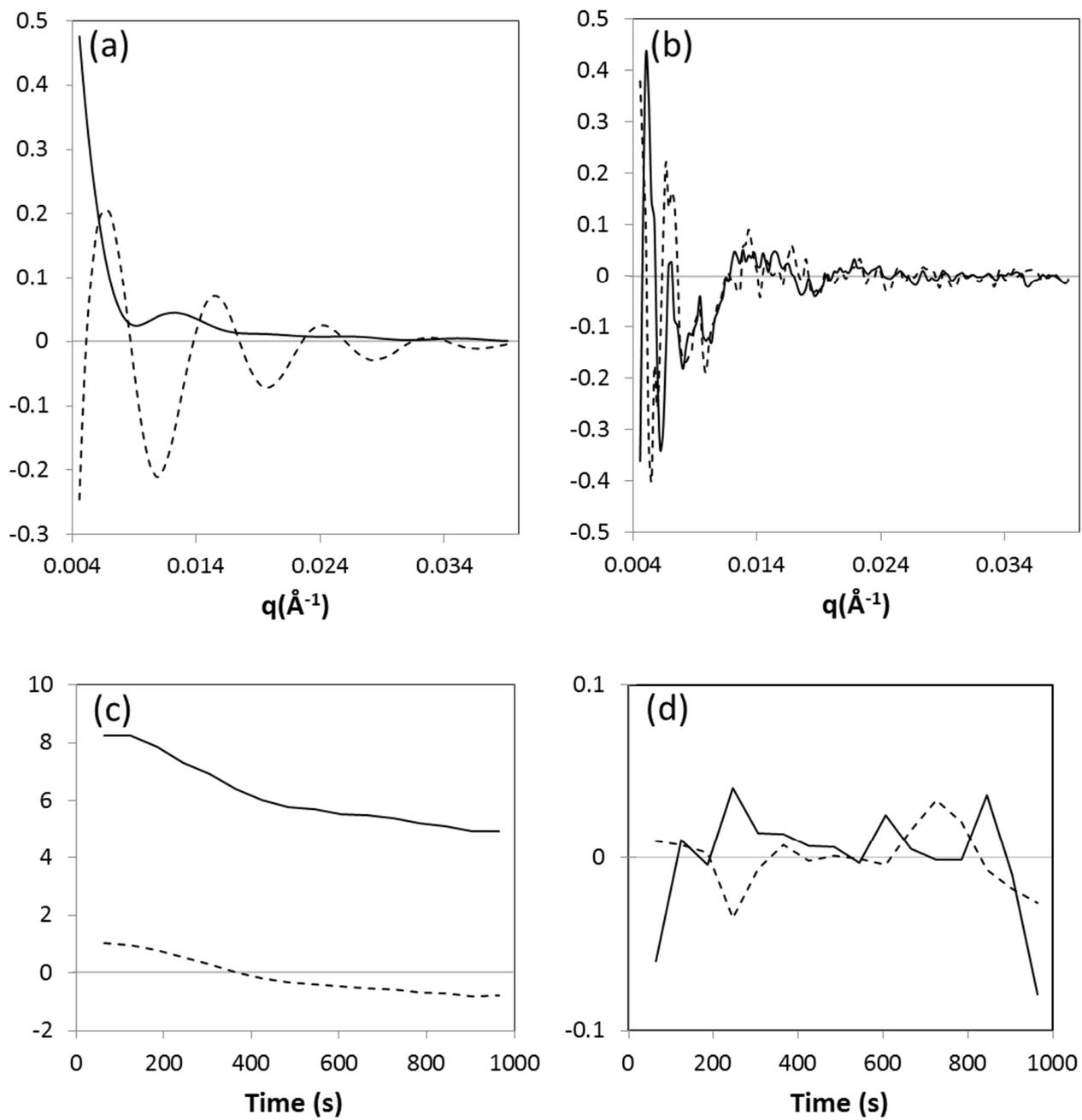


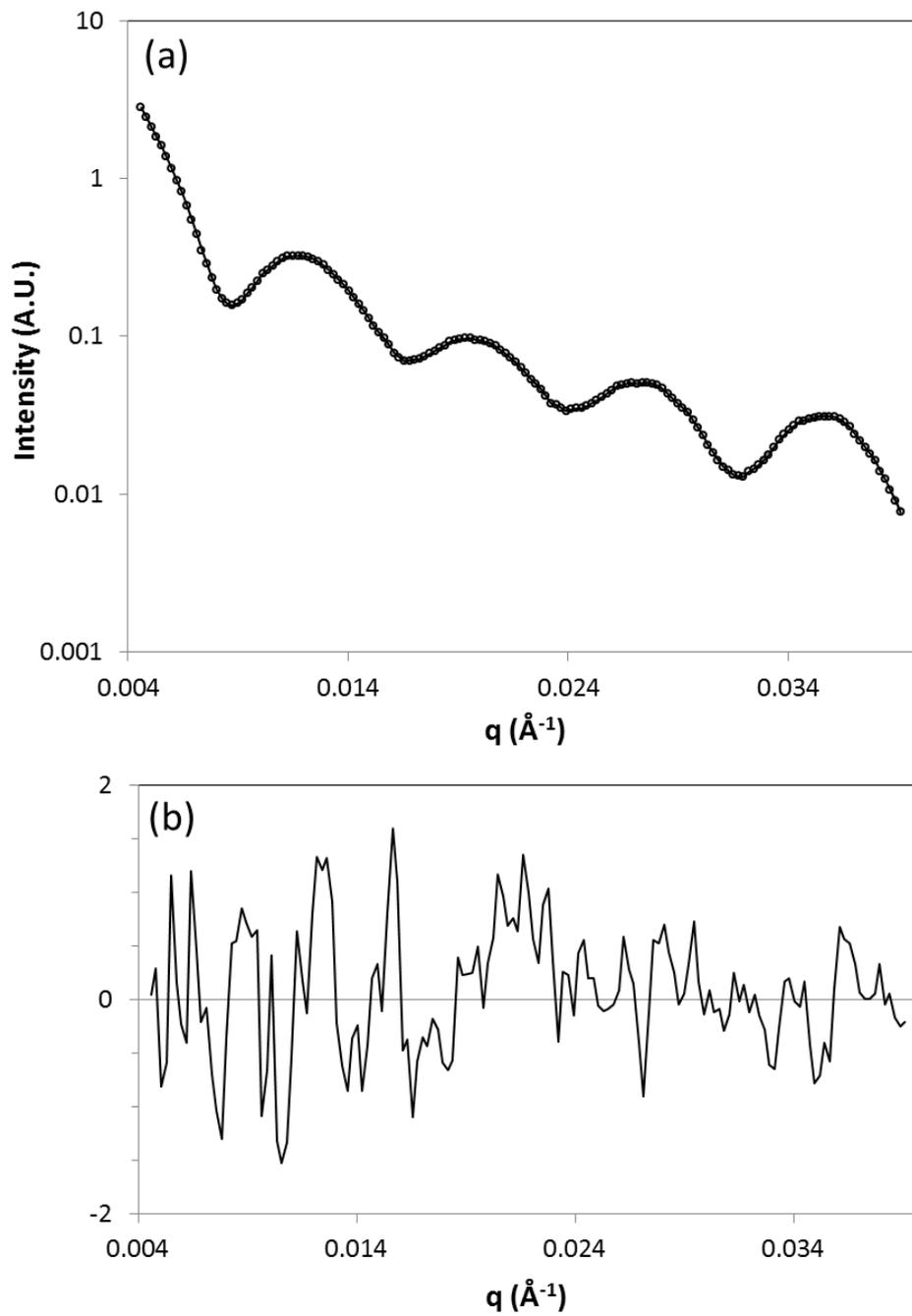
Figure 2. SAXS patterns and distance distribution functions  $p(r)$  of prohead II and expanded capsid.



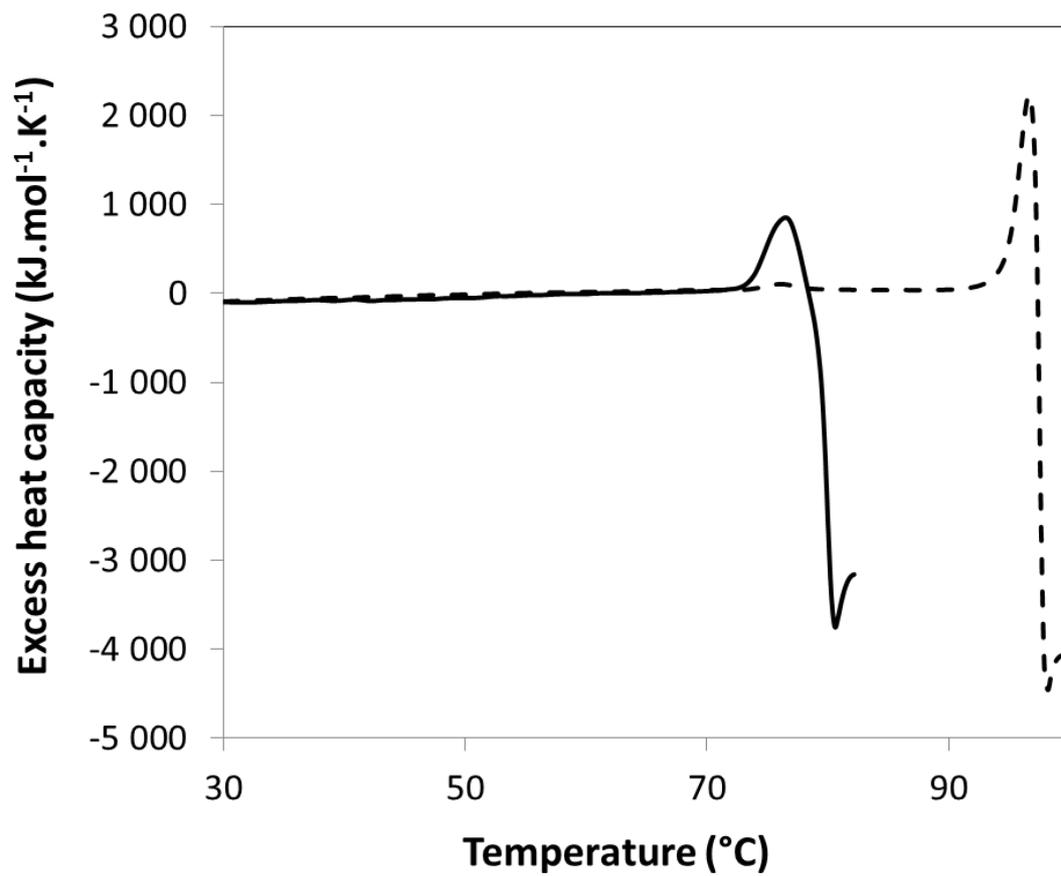
**Figure 3. Time resolved SAXS patterns and distance distribution functions  $p(r)$  of prohead II during acid-induced expansion.**



**Figure 4.** Singular value decomposition of one experiment at pH 4.26.



**Figure 5. Restoration of time resolved patterns using only the first two eigenvectors from the SVD analysis.**



**Figure 6. Calorimetric recording of prohead II (continuous line) and expanded capsid (broken line).**

## References

1. Ackermann, H. W. (2009). Phage classification and characterization. *Methods Mol Biol* **501**, 127-40.
2. Steven, A. C., Heymann, J. B., Cheng, N., Trus, B. L. & Conway, J. F. (2005). Virus maturation: dynamics and mechanism of a stabilizing structural transition that leads to infectivity. *Curr Opin Struct Biol* **15**, 227-36.
3. Johnson, J. E. (2010). Virus particle maturation: insights into elegantly programmed nanomachines. *Curr Opin Struct Biol* **20**, 210-6.
4. Aksyuk, A. A. & Rossmann, M. G. (2011). Bacteriophage assembly. *Viruses* **3**, 172-203.
5. Rao, V. B. & Feiss, M. (2008). The bacteriophage DNA packaging motor. *Annu Rev Genet* **42**, 647-81.
6. Casjens, S. R. (2011). The DNA-packaging nanomotor of tailed bacteriophages. *Nat Rev Microbiol* **9**, 647-57.
7. Buttner, C. R., Chechik, M., Ortiz-Lombardia, M., Smits, C., Ebong, I. O., Chechik, V., Jeschke, G., Dykeman, E., Benini, S., Robinson, C. V., Alonso, J. C. & Antson, A. A. (2012). Structural basis for DNA recognition and loading into a viral packaging motor. *Proc Natl Acad Sci U S A* **109**, 811-6.
8. Gao, S. & Rao, V. B. (2011). Specificity of interactions among the DNA-packaging machine components of T4-related bacteriophages. *J Biol Chem* **286**, 3944-56.
9. Zhao, H., Finch, C. J., Sequeira, R. D., Johnson, B. A., Johnson, J. E., Casjens, S. R. & Tang, L. (2010). Crystal structure of the DNA-recognition component of the bacterial virus Sf6 genome-packaging machine. *Proc Natl Acad Sci U S A* **107**, 1971-6.
10. Orlia, A. S., Prevelige, P. E., Jr., Johnson, J. E. & Cingolani, G. (2011). Three-dimensional structure of a viral genome-delivery portal vertex. *Nat Struct Mol Biol* **18**, 597-603.
11. Orlova, E. V., Dube, P., Beckmann, E., Zemlin, F., Lurz, R., Trautner, T. A., Tavares, P. & van Heel, M. (1999). Structure of the 13-fold symmetric portal protein of bacteriophage SPP1. *Nat Struct Biol* **6**, 842-6.
12. Simpson, A. A., Tao, Y., Leiman, P. G., Badasso, M. O., He, Y., Jardine, P. J., Olson, N. H., Morais, M. C., Grimes, S., Anderson, D. L., Baker, T. S. & Rossmann, M. G. (2000). Structure of the bacteriophage phi29 DNA packaging motor. *Nature* **408**, 745-50.
13. Kondabagil, K. R., Zhang, Z. & Rao, V. B. (2006). The DNA translocating ATPase of bacteriophage T4 packaging motor. *J Mol Biol* **363**, 786-99.
14. Oliveira, L., Cuervo, A. & Tavares, P. (2010). Direct interaction of the bacteriophage SPP1 packaging ATPase with the portal protein. *J Biol Chem* **285**, 7366-73.
15. Smith, D. E., Tans, S. J., Smith, S. B., Grimes, S., Anderson, D. L. & Bustamante, C. (2001). The bacteriophage straight phi29 portal motor can package DNA against a large internal force. *Nature* **413**, 748-52.
16. Tavares, P., Zinn-Justin, S. & Orlova, E. V. (2012). Genome gating in tailed bacteriophage capsids. *Adv Exp Med Biol* **726**, 585-600.

17. Veesler, D. & Cambillau, C. (2011). A common evolutionary origin for tailed-bacteriophage functional modules and bacterial machineries. *Microbiol Mol Biol Rev* **75**, 423-33, first page of table of contents.
18. Duda, R. L., Hempel, J., Michel, H., Shabanowitz, J., Hunt, D. & Hendrix, R. W. (1995). Structural transitions during bacteriophage HK97 head assembly. *J Mol Biol* **247**, 618-35.
19. Lee, K. K., Tsuruta, H., Hendrix, R. W., Duda, R. L. & Johnson, J. E. (2005). Cooperative reorganization of a 420 subunit virus capsid. *J Mol Biol* **352**, 723-35.
20. Conway, J. F., Wikoff, W. R., Cheng, N., Duda, R. L., Hendrix, R. W., Johnson, J. E. & Steven, A. C. (2001). Virus maturation involving large subunit rotations and local refolding. *Science* **292**, 744-8.
21. Gertsman, I., Gan, L., Guttman, M., Lee, K., Speir, J. A., Duda, R. L., Hendrix, R. W., Komives, E. A. & Johnson, J. E. (2009). An unexpected twist in viral capsid maturation. *Nature* **458**, 646-50.
22. Ionel, A., Velazquez-Muriel, J. A., Luque, D., Cuervo, A., Caston, J. R., Valpuesta, J. M., Martin-Benito, J. & Carrascosa, J. L. (2011). Molecular rearrangements involved in the capsid shell maturation of bacteriophage T7. *J Biol Chem* **286**, 234-42.
23. Jiang, W., Li, Z., Zhang, Z., Baker, M. L., Prevelige, P. E., Jr. & Chiu, W. (2003). Coat protein fold and maturation transition of bacteriophage P22 seen at subnanometer resolutions. *Nat Struct Biol* **10**, 131-5.
24. Smith, D. E. (2011). Single-molecule studies of viral DNA packaging. *Curr Opin Virol* **1**, 134-41.
25. Duda, R. L. (1998). Protein chainmail: catenated protein in viral capsids. *Cell* **94**, 55-60.
26. Lander, G. C., Evilevitch, A., Jeembaeva, M., Potter, C. S., Carragher, B. & Johnson, J. E. (2008). Bacteriophage lambda stabilization by auxiliary protein gpD: timing, location, and mechanism of attachment determined by cryo-EM. *Structure* **16**, 1399-406.
27. Rao, V. B. & Black, L. W. (2010). Structure and assembly of bacteriophage T4 head. *Virol J* **7**, 356.
28. Effantin, G., Boulanger, P., Neumann, E., Letellier, L. & Conway, J. F. (2006). Bacteriophage T5 structure reveals similarities with HK97 and T4 suggesting evolutionary relationships. *J Mol Biol* **361**, 993-1002.
29. Huet, A., Conway, J. F., Letellier, L. & Boulanger, P. (2010). In vitro assembly of the T=13 procapsid of bacteriophage T5 with its scaffolding domain. *J Virol* **84**, 9350-8.
30. Saigo, K. (1978). Isolation of high-density mutants and identification of nonessential structural proteins in bacteriophage T5; dispensability of L-shaped tail fibers and a secondary major head protein. *Virology* **85**, 422-33.
31. Steven, A. C., Greenstone, H. L., Booy, F. P., Black, L. W. & Ross, P. D. (1992). Conformational changes of a viral capsid protein. Thermodynamic rationale for proteolytic regulation of bacteriophage T4 capsid expansion, co-operativity, and super-stabilization by soc binding. *J Mol Biol* **228**, 870-84.
32. Medina, E., Nakatani, E., Kruse, S. & Catalano, C. E. (2012). Thermodynamic characterization of viral procapsid expansion into a functional capsid shell. *J Mol Biol* **418**, 167-80.
33. Conway, J. F., Duda, R. L., Cheng, N., Hendrix, R. W. & Steven, A. C. (1995). Proteolytic and conformational control of virus capsid maturation: the bacteriophage HK97 system. *J Mol Biol* **253**, 86-99.
34. Fokine, A., Leiman, P. G., Shneider, M. M., Ahvazi, B., Boeshans, K. M., Steven, A. C., Black, L. W., Mesyanzhinov, V. V. & Rossmann, M. G. (2005). Structural and

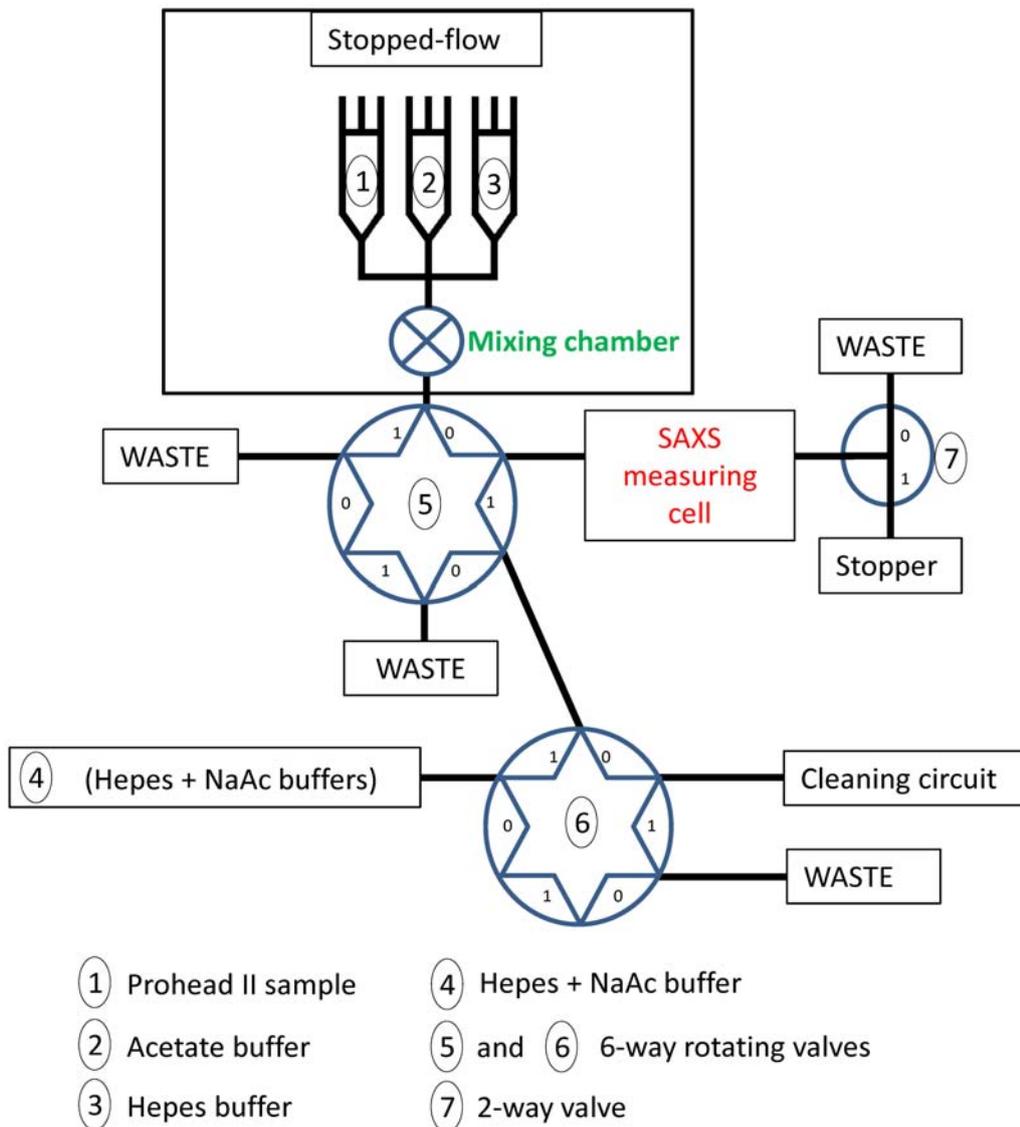
- functional similarities between the capsid proteins of bacteriophages T4 and HK97 point to a common ancestry. *Proc Natl Acad Sci U S A* **102**, 7163-8.
35. Tsuruta, H., Brennan, S., Rek, Z. U., Irving, T. C., Tompkins, W. H. & Hodgson, K. O. (1998). A Wide-Bandpass Multilayer Monochromator for Biological Small-Angle Scattering and Fiber Diffraction Studies. *J. Appl. Cryst.* **31**, 672-682.
  36. Wikoff, W. R., Liljas, L., Duda, R. L., Tsuruta, H., Hendrix, R. W. & Johnson, J. E. (2000). Topologically linked protein rings in the bacteriophage HK97 capsid. *Science* **289**, 2129-33.
  37. Ross, P. D., Cheng, N., Conway, J. F., Firek, B. A., Hendrix, R. W., Duda, R. L. & Steven, A. C. (2005). Crosslinking renders bacteriophage HK97 capsid maturation irreversible and effects an essential stabilization. *EMBO J* **24**, 1352-63.
  38. Fokine, A., Chipman, P. R., Leiman, P. G., Mesyanzhinov, V. V., Rao, V. B. & Rossmann, M. G. (2004). Molecular architecture of the prolate head of bacteriophage T4. *Proc Natl Acad Sci U S A* **101**, 6003-8.
  39. Chen, D. H., Baker, M. L., Hryc, C. F., DiMaio, F., Jakana, J., Wu, W., Dougherty, M., Haase-Pettingell, C., Schmid, M. F., Jiang, W., Baker, D., King, J. A. & Chiu, W. (2011). Structural basis for scaffolding-mediated assembly and maturation of a dsDNA virus. *Proc Natl Acad Sci U S A* **108**, 1355-60.
  40. Parent, K. N., Khayat, R., Tu, L. H., Suhanovsky, M. M., Cortines, J. R., Teschke, C. M., Johnson, J. E. & Baker, T. S. (2010). P22 coat protein structures reveal a novel mechanism for capsid maturation: stability without auxiliary proteins or chemical crosslinks. *Structure* **18**, 390-401.
  41. Ray, K., Oram, M., Ma, J. & Black, L. W. (2009). Portal control of viral prohead expansion and DNA packaging. *Virology* **391**, 44-50.
  42. Cerritelli, M. E., Conway, J. F., Cheng, N., Trus, B. L. & Steven, A. C. (2003). Molecular mechanisms in bacteriophage T7 procapsid assembly, maturation, and DNA containment. *Adv Protein Chem* **64**, 301-23.
  43. Orthaber, D., Bergmann, A. & Glatter, O. (2000). SAXS experiments on absolute scale with Kratky systems using water as a secondary standard. *J. Appl. Cryst.* **33**, 218 - 225.
  44. David, G. & Perez, J. (2009). Combined sampler robot and high-performance liquid chromatography: a fully automated system for biological small-angle X-ray scattering experiments at the Synchrotron SOLEIL SWING beamline. *J. Appl. Crystallogr.* **42** 892-900.
  45. Svergun, D. I. (1992). Determination of the Regularization Parameter in Indirect - Transform Methods Using Perceptual Criteria. *J. Appl. Cryst.* **25**, 495-503.
  46. Fowler, A. G., Foote, A. M., Moody, M. F., Vachette, P., Provencher, S. W., Gabriel, A., Bordas, J. & Koch, M. H. J. (1983). Stopped-flow solution scattering using synchrotron radiation : apparatus, data collection and data analysis. *J. Biochem. Biophys. Methods* **7**, 317 - 329.
  47. Provencher, S. W. & Glöckner, J. (1983). Analysis of components present in kinetics (or titration) curves. *J. Biochem. Biophys. Methods* **7**, 331 - 334.
  48. Fetler, L., Tauc, P., Herve, G., Moody, M. F. & Vachette, P. (1995). X-ray scattering titration of the quaternary structure transition of aspartate transcarbamylase with a bisubstrate analogue: influence of nucleotide effectors. *J Mol Biol* **251**, 243-55.
  49. Chen, L., Hodgson, K. O. & Doniach, S. (1996). A Lysozyme Folding Intermediate Revealed by Solution X-ray Scattering. *J. Mol. Biol.* **261**, 658 -671.
  50. Hagihara, Y., Hoshino, M., Hamada, D., Kataoka, M. & Goto, Y. (1998). Chain-like conformation of heat-denatured ribonuclease A and cytochrome *c* as evidenced by solution X-ray scattering. *Fold. Des.* **3**, 195-201.

51. Pérez, J., Vachette, P., Russo, D., Desmadril, M. & Durand, D. (2001). Heat-induced unfolding of neocarzinostatin, a small all-beta protein investigated by small-angle X-ray scattering. *J Mol Biol* **308**, 721-43.
52. Perez, J., Defrenne, S., Witz, J. & Vachette, P. (2000). Detection and characterization of an intermediate conformation during the divalent ion-dependent swelling of tomato bushy stunt virus. *Cell Mol Biol (Noisy-le-grand)* **46**, 937-48.
53. Petoukhov, M. V., Franke, D., Shkumatov, A. V., Tria, G., Kikhney, A. G., Gajda, M., Gorba, C., Mertens, H. D. T., Konarev, P. V. & Svergun, D. I. (2012). New developments in the ATSAS program package for small-angle scattering data analysis. *J. Appl. Cryst.* **45**, 342-350.

## **SUPPLEMENTAL FIGURES**

**A two-state cooperative expansion converts the procapsid shell of bacteriophage T5 into a highly stable capsid isomorphous to the final virion head**

Olivier Preux, Dominique Durand, Alexis Huet, Aurélie Bertin, Claire Boulogne, Jeannine Drouin-Wahbi, Didier Trévarin, Javier Pérez, Patrice Vachette, Pascale Boulanger



**Figure S1: Schematic representation of the experimental setup for TR- SAXS measurements.**

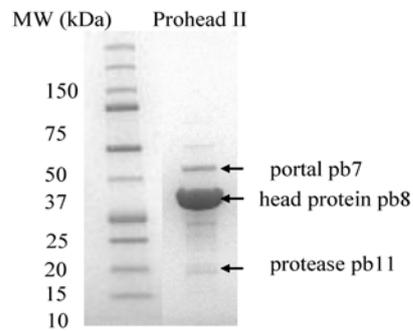
*Syringe 1:* prohead II (1mg/mL) solution in 20mM Hepes buffer pH 7.6, 400mM NaCl.

*Syringe 2:* 160mM sodium acetate buffer adjusted at the appropriate pH in the range [3.90-4.25] to yield after mixing with an equal volume of Prohead II solution (syringe 1) samples at final pH values of 4.30, 4.26, 4.23 and 4.17.

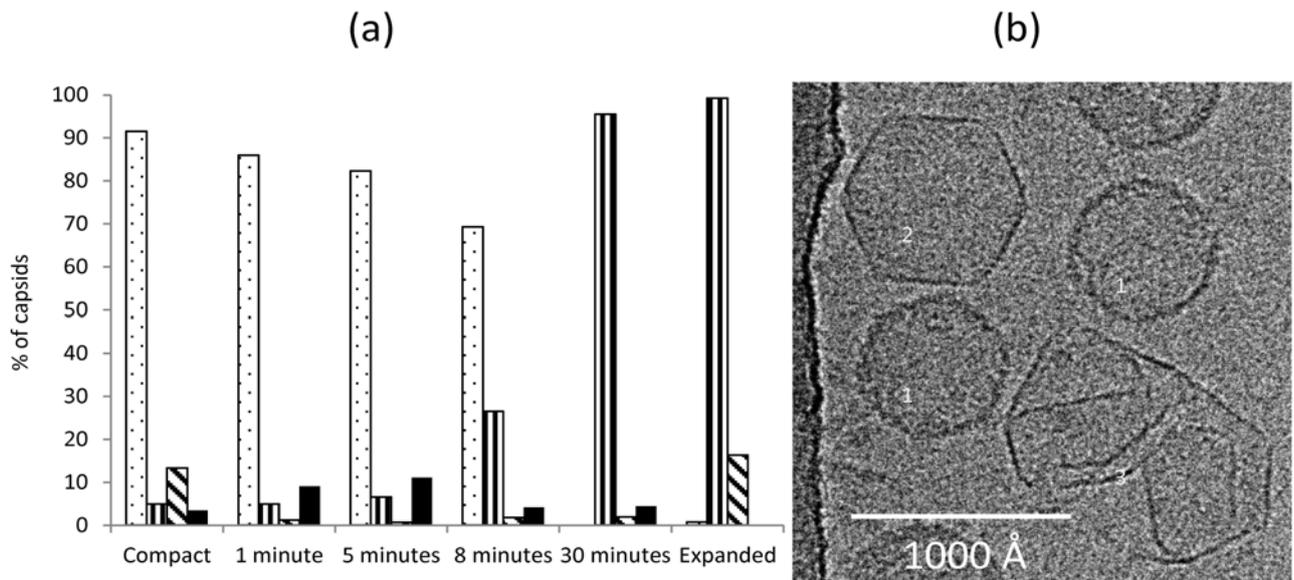
*Syringe 3:* 20mM Hepes buffer pH 7.6, 400mM NaCl (capsid buffer before mixing)

*Syringe 4:* 80mM sodium acetate buffer, 10mM Hepes buffer, 200mM NaCl (capsid buffer after mixing).

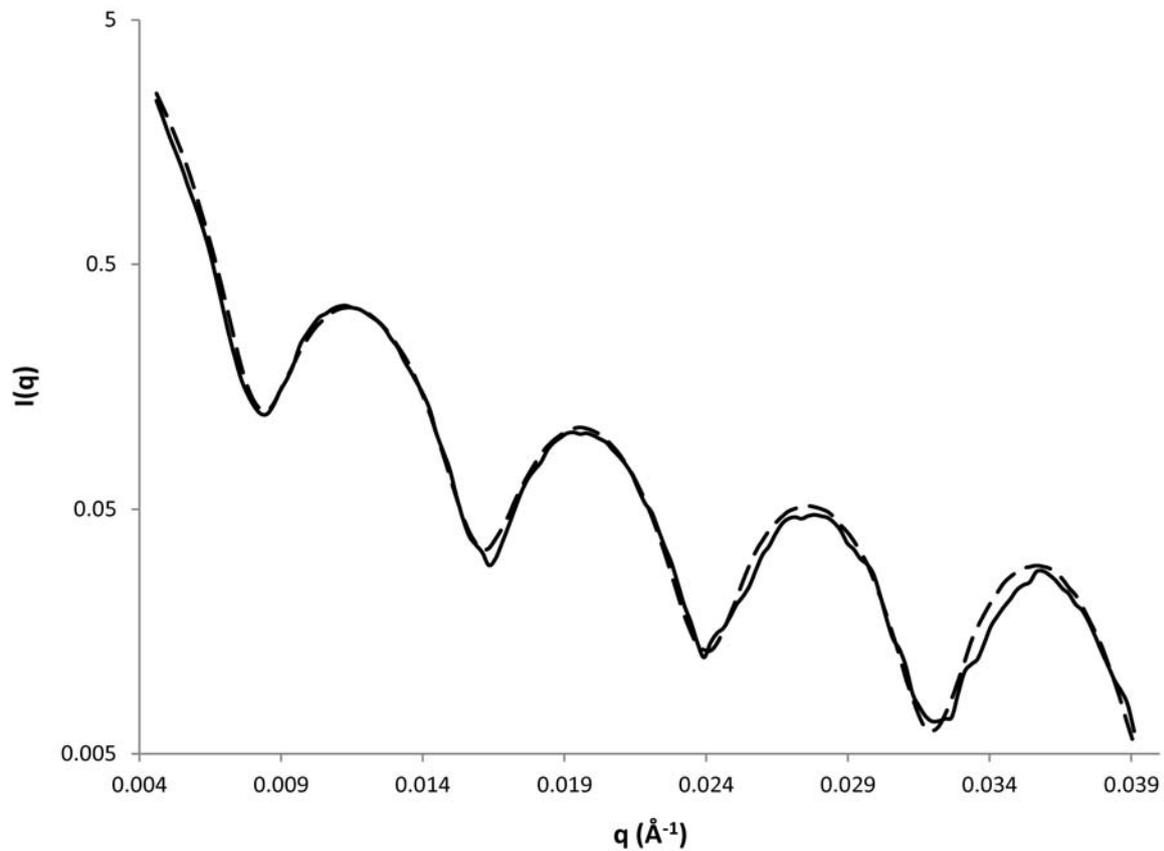
Before starting the experiment, valves (5), (6) and (7) were in position 0, 1 and 0 respectively. A typical experiment was run as follows: **i)** 200 $\mu$ L of prohead II sample and of acetate buffer were mixed and pushed into the SAXS measuring unit. The 2-way valve (7) was in position 0 in order to avoid over pressure. **ii)** Immediately after mixing, valve (7), was rotated in position 1 to keep the mixed solution in the SAXS measuring cell. **iii)** 2 seconds after mixing, we recorded the first frame. **iv)** Valve (5) was then rotated into position 1 and 6 $\mu$ L of a 1:1 mixture of Hepes and acetate buffers (Syringe 4) was used to push the same volume of the acidified prohead II solution contained in the tubing connecting the valve (5) with the SAXS measuring cell. The prohead sample was thus renewed to prevent radiation damage generated by the high-brilliance synchrotron source. The tubing volume (220  $\mu$ L) allowed us to perform 18 cycles of SAXS recording followed by sample renewal. With this number of frames we could monitor the complete expansion process prohead II at pH 4.26. **v)** In parallel with the first cycle of step (iv), 200 $\mu$ L of Hepes buffer (syringe 3) were pushed into the waste to chase away the acetate buffer from mixing chamber, thereby preventing premature acidification of the prohead II sample in syringe 1 (the Biologics stopped-flow apparatus does not contain any internal valve). **vi)** Between each experiment, the system was thoroughly cleaned via valve (6) using a commercial Hellmanex<sup>®</sup> solution and finally rinsed with the mixture of Hepes and sodium acetate buffers (syringe 4).



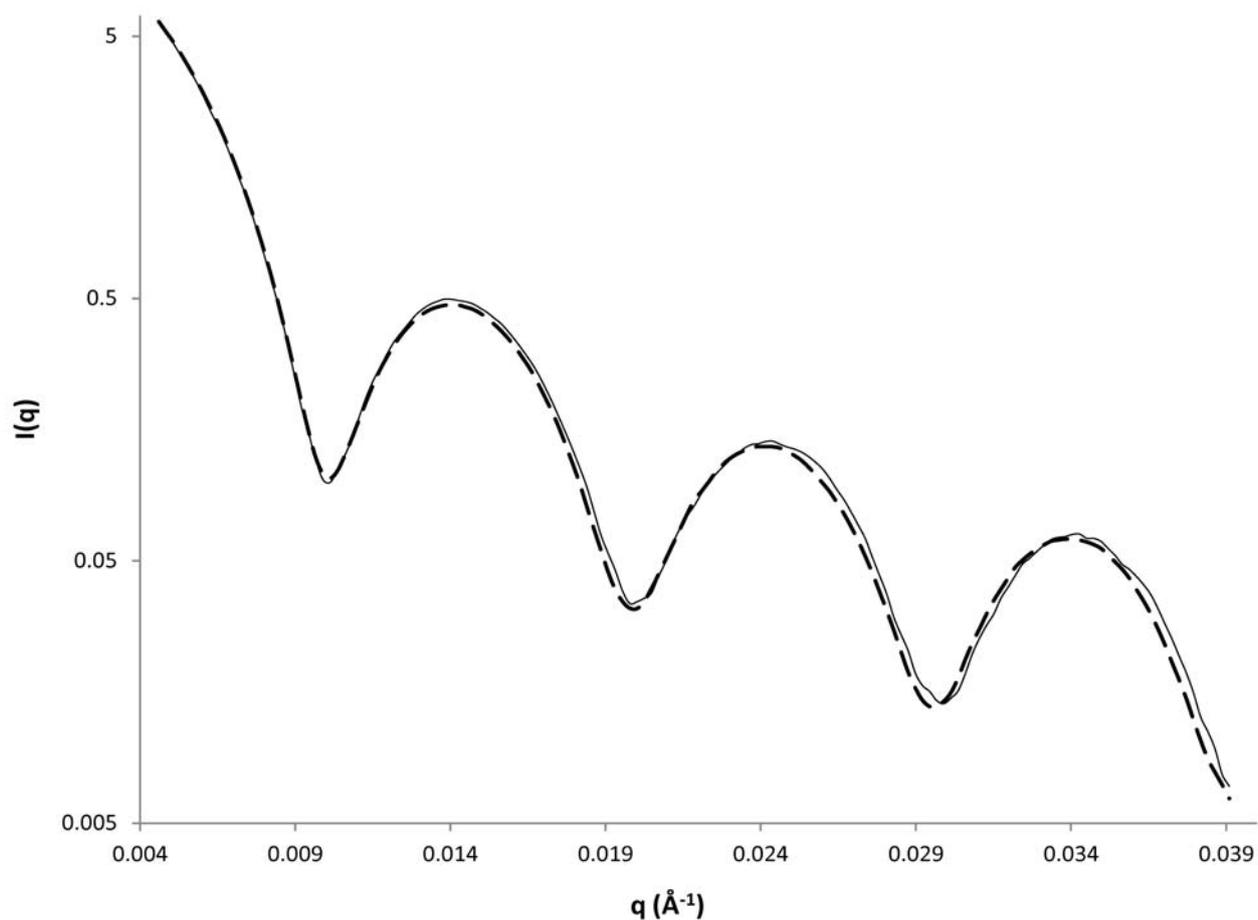
**Figure S2: Protein composition of T5 prohead II.** SDS-PAGE (4-20% Precise Protein Gel, Thermo Scientific) shows the major head protein pb8 (32.8 kDa), the portal pb7 (43.8 kDa) and the head protease pb11 (15.8 kDa). The three prohead II components have undergone proteolytic cleavage by pb11 upon maturation of prohead I into prohead II (manuscript in preparation).



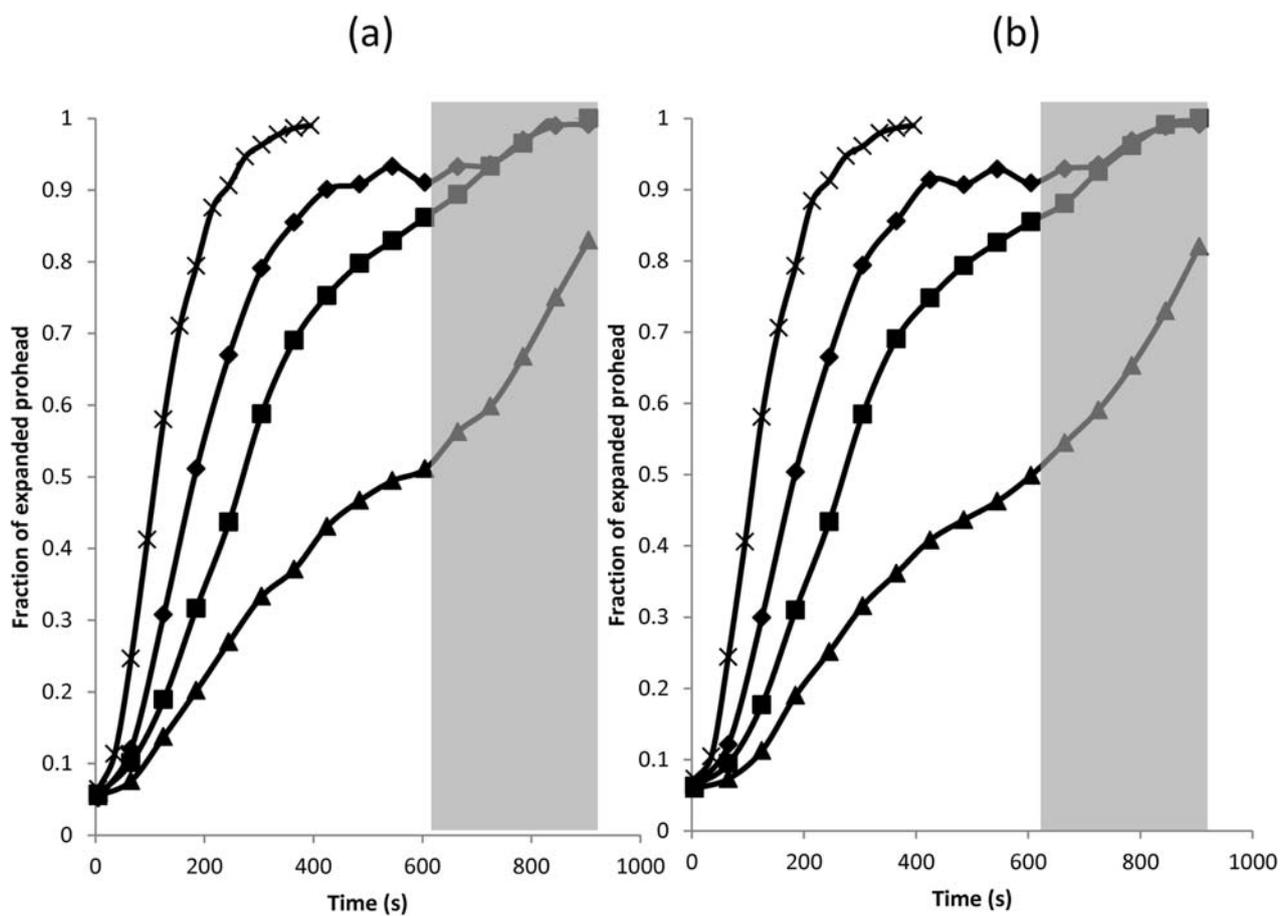
**Figure S3: Cryo EM imaging of acid-induced expansion of proheads II.** **a)** Histogram of the fractions of prohead II (dotted bar), expanded capsids (vertical lines), vesicles (oblique lines) and broken expanded capsids (solid bar) out of the total number of capsids and over acidification time. “Compact” corresponds to the native prohead sample, “expanded” corresponds to the dialyzed capsids. The total number of capsids counted is approx. 2900. **b)** Cryo-electron micrograph of a prohead II sample at pH 4.30 in 200mM NaCl flash-frozen 8 minutes after pH drop. Compact proheads II (1), expanded capsids (2) and broken expanded capsids (3) can be seen.



**Figure S4: Comparison of the SAXS patterns of capsid expanded by dialysis against a 50mM Hepes buffer pH 7.6 (broken line) and by acidification at pH 4.26 (continuous lines).** The latter curve corresponds to the last frame of TR-SAXS experiment. The SAXS pattern of the dialyzed capsids is the average of 20 frames of 1 second exposure while the scattering curve of the acidified capsids is a unique 1 second frame and is therefore noisier. Except for the noise level, the 2 curves show a great similarity. Cryo-EM analysis showed that dialyzed capsids are only composed of expanded capsids; this suggests that the last frame of TR-SAXS measurements after acidification at pH 4.26 also corresponds to expanded capsids.



**Figure S5: Comparison of the SAXS patterns of a prohead II sample (broken line) with the first frame of TR measurements following pH-drop (continuous line).** The two curves are very similar. The very slight shift towards larger angles of the acidified capsid pattern is probably due to the disappearance of vesicles



**Supplementary Figure 6**

**Figure S6. pH dependence of expansion kinetics.** The time-dependent fraction of expanded prohead after acidification at pH 4.30 (triangle), 4.26 (square), 4.23 (diamond) and 4.17 (cross) were computed from (a) the pair distance distribution and (b) the time-dependence of the projections on the second basis vector in SVD analysis as described in Material and methods. The later phase of the kinetics that is altered by dilution artefacts is shaded.

### III- Modèles cinétiques de l'expansion

L'analyse SVD des données SAXS présentée dans l'article montre qu'il n'y a pas d'état d'intermédiaire détectable par SAXS durant l'expansion, c'est-à-dire pas de capsides ayant une forme globale et/ou une taille différente des états compact et expansé. J'ai effectué l'analyse de l'évolution des populations respectives de capsides compactes et de capsides expansées dans le but de déterminer la présence ou l'absence d'état(s) intermédiaire(s) non détectable(s) par SAXS.

Considérons la transition irréversible en une étape, c'est-à-dire sans état intermédiaire, entre l'état compact et l'état expansé:



La constante cinétique est  $k_1=1/\tau_1$ , où  $\tau_1$  est le temps caractéristique de cette étape.

Les concentrations des capsides compactes  $[C]$  et expansées  $[E]$  sont alors égales à :

$$[C] = C_0 e^{-k_1 t}$$

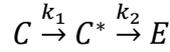
$$[E] = C_0 (1 - e^{-k_1 t})$$

En faisant l'hypothèse d'une transition en une seule étape, la fraction de capsides expansées s'exprime donc par cette formule :

$$G_{exp,t} = 1 - e^{-k_1 t} \quad (7)$$

Or cette expression ne peut pas rendre compte des données expérimentales (figure 33 (a)), notamment du fait que l'évolution rapide ne se déclenche qu'au bout d'une dizaine de secondes.

L'évolution des populations de capsides compactes et expansées ne pouvant pas être décrite par un modèle de transition consistant en une seule barrière d'énergie, nous avons essayé d'en rendre compte avec un modèle incluant un état intermédiaire :



Les constantes cinétiques sont  $k_1=1/\tau_1$  and  $k_2=1/\tau_2$ ,  $\tau_1$  et  $\tau_2$  étant les temps caractéristiques respectifs des étapes 1 et 2.

Ce troisième état, noté  $C^*$ , ne peut pas être différencié de l'état compact par SAXS ou par microscopie électronique, il pourrait consister en une capside dont les capsomères ont subi des réarrangements structuraux locaux.

La dépendance temporelle de la concentration des trois espèces de capsides est entièrement déterminée par les constantes cinétiques  $k_1$  et  $k_2$  (voir l'annexe 3 pour la démonstration de ces formules):

$$[C] = C_0 e^{-k_1 t}$$

$$[C^*] = C_0 \frac{k_1}{k_2 - k_1} [e^{-k_1 t} - e^{-k_2 t}]$$

$$[E] = C_0 \left[ 1 + \frac{k_1}{k_2 - k_1} e^{-k_2 t} - \left( 1 + \frac{k_1}{k_2 - k_1} \right) e^{-k_1 t} \right]$$

La fraction de capsides expansées s'exprime donc par la formule :

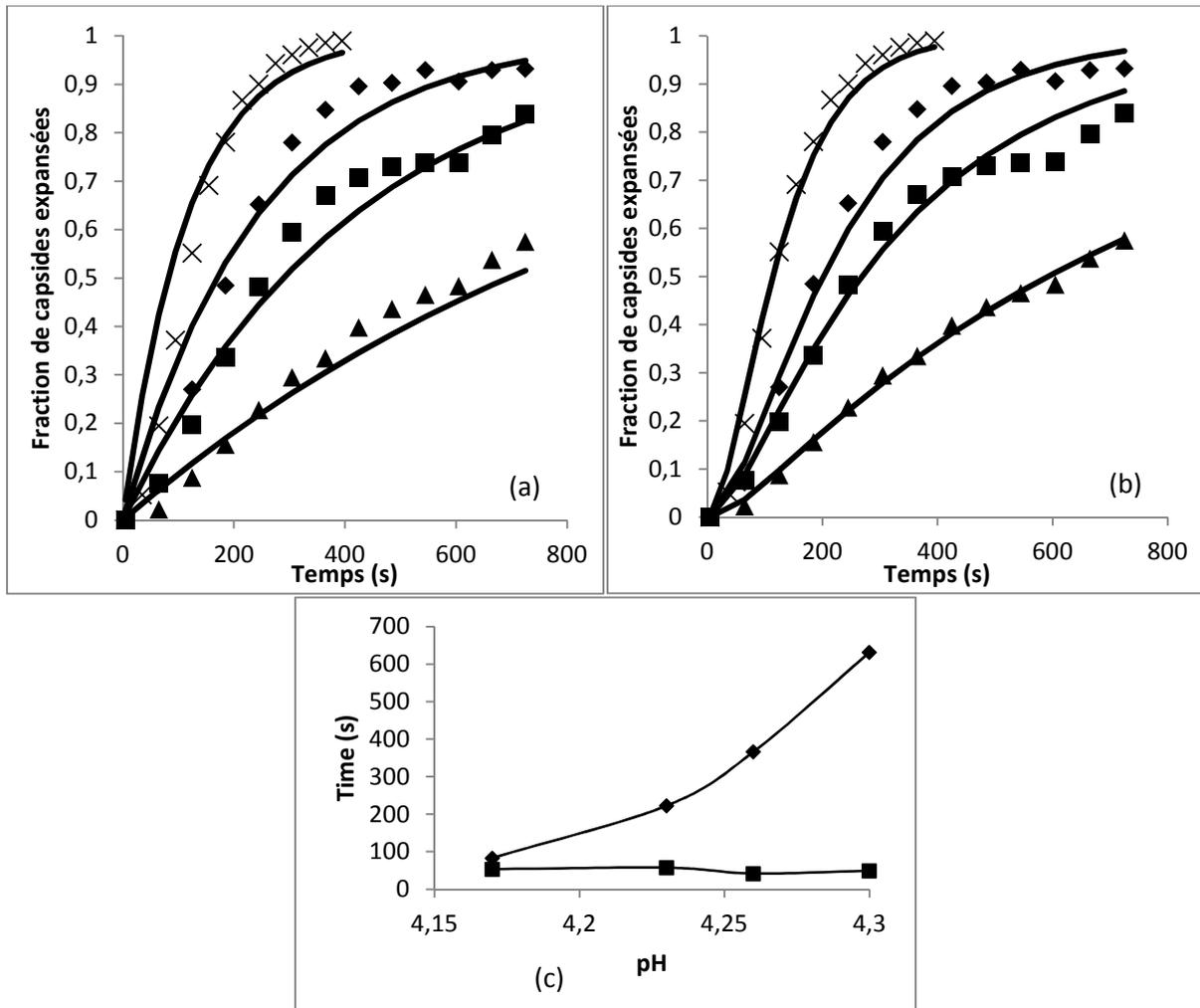
$$G_{exp,t} = 1 + \frac{k_1}{k_2 - k_1} e^{-k_2 t} - \left( 1 + \frac{k_1}{k_2 - k_1} \right) e^{-k_1 t} \quad (8)$$

La figure 33 (b) permet de comparer les données expérimentales de l'évolution de la fraction de capsides expansées aux données théoriques calculées à l'aide de l'équation (8) dont les paramètres  $k_1$  et  $k_2$  ont été ajustés mathématiquement à l'aide du logiciel

KaleidaGraph (Synergy Software). On remarque que la valeur de  $\tau_2$  est fortement dépendante du pH auquel les capsides ont été acidifiées, au contraire de la valeur de  $\tau_1$  (figure 33 (c)), qui est constant, à la précision des mesures près, dans la gamme de pH étudié (4,17-4,30).

Avec ce modèle, le processus d'expansion pourrait être décrit de la manière suivante. La baisse du pH entraîne un changement de l'état de protonation de certains acides aminés, qui pourrait entraîner localement des répulsions électrostatiques responsables de changements structuraux locaux de la protéine majeure de capside, éventuellement une modification de la structure secondaire. De par l'invariance de la valeur de  $\tau_1$  dans la gamme de pH 4.17-4.30, il est probable que tous les acides aminés incriminés dans cette étape soient protonés, donc que leur pKa soit bien supérieur à 4,30. En revanche la dépendance de  $\tau_2$  au pH d'acidification indique que le pKa des acides aminés impliqués dans la seconde étape, consistant en un changement spectaculaire de taille et de forme de la capside, serait du même ordre de grandeur que ce pH d'acidification.

Cependant, comme on peut le voir sur la figure 33 (b), si les valeurs des fractions de capsides expansées calculées avec le modèle comportant un état intermédiaire sont dans la marge d'erreur des données expérimentales obtenues à pH 4.30, elles s'en écartent un peu à pH 4.26 et ne rendent pas compte des données expérimentales à pH 4.23 et 4.17. Les données expérimentales ne permettent donc pas de conclure de manière univoque sur la présence d'un état intermédiaire non détectable par SAXS.



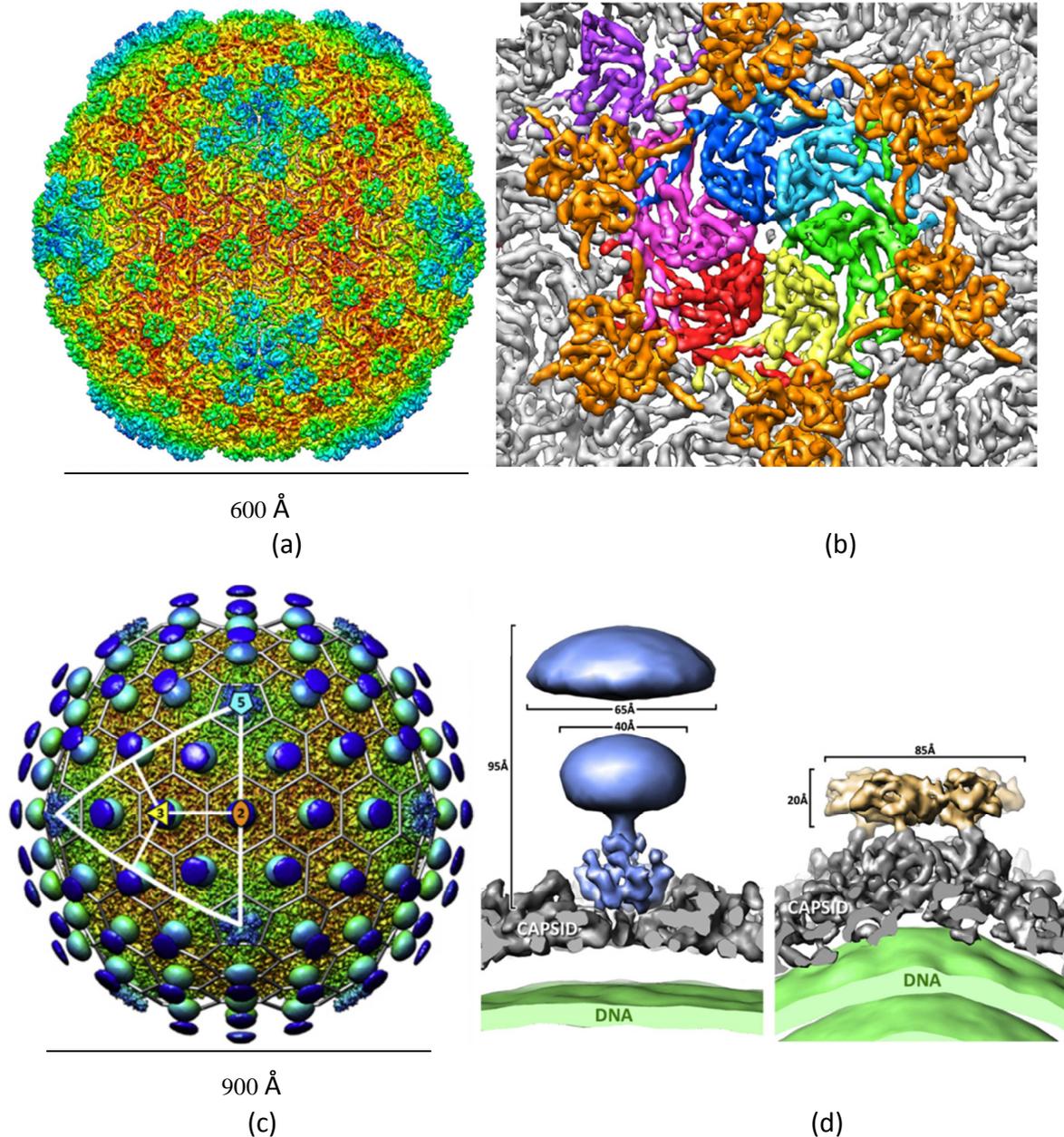
**Figure 33. Etude de la cinétique de l'expansion.** (a) et (b) Fraction de capsides expansées en fonction du temps après acidification à pH 4.17 (croix), 4.23 (losanges), 4.26 (carrés) et 4.30 (triangles). Les lignes continues correspondent aux « fits » obtenus en utilisant les équations (7) (a) et (8) (b). (c) Variation des temps caractéristiques (équation (8)) de la première ( $\tau_1$ , carrés) et de la deuxième ( $\tau_2$ , losanges) étape de l'expansion en fonction du pH de l'acidification.



**ETUDE STRUCTURALE ET FONCTIONNELLE  
DE LA PROTEINE DE DECORATION pb10**

# I- INTRODUCTION

Les capsides des phages icosaédriques sont constituées de protéines majeures formant l'architecture de la capsid, d'une protéine portale formant le port d'entrée et de sortie de l'ADN et de protéines de fermeture formant le connecteur, vertex spécialisé dans l'attachement à la queue. Certains phages possèdent d'autres protéines, dites « accessoires », qui ont la particularité de ne se fixer qu'après l'expansion de la capsid (Lander, Evilevitch et al. 2008). Leur site de fixation est donc généré par les modifications structurales dues à l'expansion. Certaines de ces protéines, comme la protéine gpD du phage  $\lambda$  (Yang, Maluf et al. 2008), sont indispensables pour stabiliser la capsid et lui permettre de résister à la pression interne générée par l'état fortement compacté de l'ADN. D'autres ne sont pas indispensables à l'assemblage complet de la capsid, mais participent à sa stabilité, par exemple la protéine SOC du phage T4 (Qin, Fokine et al. 2010) et la protéine DEC du phage L (Tang, Gilcrease et al. 2006). Ces trois protéines renforcent la solidité de la capsid en se fixant aux interfaces entre trois hexamères (figure 34). Elles « cimentent » ainsi les hexamères entre eux. Certaines protéines se fixent au centre des hexamères, d'autres au centre des pentamères, comme les protéines de décoration du phage SIO-2 (Lander, Baudoux et al. 2012). Il existe des protéines accessoires dont aucun effet stabilisateur n'a pu être mis en évidence. Par exemple, la protéine HOC des phages de la famille de T4 (Fokine, Chipman et al. 2004), la protéine gp12 du phage SPP1 (White, Sherman et al. 2012) et la protéine pb10 du phage T5 (Effantin, Boulanger et al. 2006), qui se fixent au centre des hexamères, ou encore la protéine gp8.5 du phage  $\phi$ 29. Le rôle biologique de ces protéines est pour l'instant inconnu.



**Figure 34. Structure de la capside des bactériophages  $\lambda$  et SIO-2.** (a) Reconstruction tridimensionnelle à partir de clichés de cryo-microscopie électronique de la capside du phage  $\lambda$  (résolution : 6.8 Å). Le diamètre moyen de la capside est de 600 Å. Les densités électroniques sont colorées en fonction de leur distance du centre de la capside, du rouge au bleu. (b) Vue d'une unité asymétrique de la capside. Les protéines majeures de capside formant un hexamère sont coloriées en bleu, cyan, vert, jaune, rouge et rose. La protéine majeure formant un des côtés d'un pentamère est coloriée en violet. Les six protéines gpD sont coloriées en orange. (c) Reconstruction tridimensionnelle à partir de clichés de cryo-microscopie électronique (résolution : 8.5 Å) de la capside du phage SIO-2. En bleu sont représentées les protéines de décoration. (d) A gauche : vue en coupe d'une protéine de décoration se fixant au centre d'un hexamère (en bleu). La forme en champignon ainsi que la discontinuité de la densité électronique est un effet de la reconstruction, elle est due à la flexibilité de la protéine. A droite : vue en coupe d'une protéine de décoration se fixant au centre d'un pentamère (en orange). (a) et (b) : figure adaptée de (Lander, Evilevitch et al. 2008) ; (c) et (d) : figure adaptée de (Lander, Baudoux et al. 2012).

# 1- Les protéines HOC

## 1-1 Présentation

Le bactériophage T4 est un phage infectant *Escherichia coli* qui appartient à l'ordre des *Caudovirales* et à la famille des *Myoviridae*. Sa capsid de symétrie pseudo-icosaédrique est allongée et a pour nombres de triangulation  $T=13$  et  $Q=20$  (Fokine, Chipman et al. 2004). Elle contient 155 copies de HOC (Highly immunogenic Outer Capsid protein), une protéine de 40 kDa se fixant au centre des hexamères formés par la protéine majeure de capsid, gp23 (figure 35 (a)). Cette protéine ou des analogues voisins au niveau de leur séquence protéique est également présente chez tous les phages apparentés à T4. Les deux seuls autres exemples, actuellement décrits, de protéines de décoration occupant le centre des hexamères de la protéine majeure de capsid sont les protéines gp12 de SPP1 et pb10 de T5.

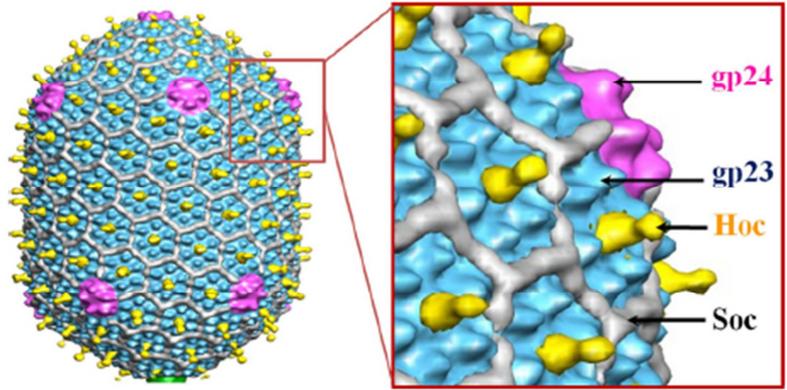
## 1-2 Données structurales des protéines HOC

L'analyse de la séquence de HOC a montré que cette protéine est composée de 4 domaines d'environ 10 kDa chacun. Les trois domaines N-terminaux possèdent un repliement de type immunoglobuline, consistant en 7 brins  $\beta$  repliés en 2 feuillets. Le domaine C-terminal est le mieux conservé parmi les protéines HOC. Les prédictions de structures secondaires indiquent avec une forte probabilité que cette protéine est constituée de quatre hélices (Cole, Barber et al. 2008; Fokine, Islam et al. 2011).

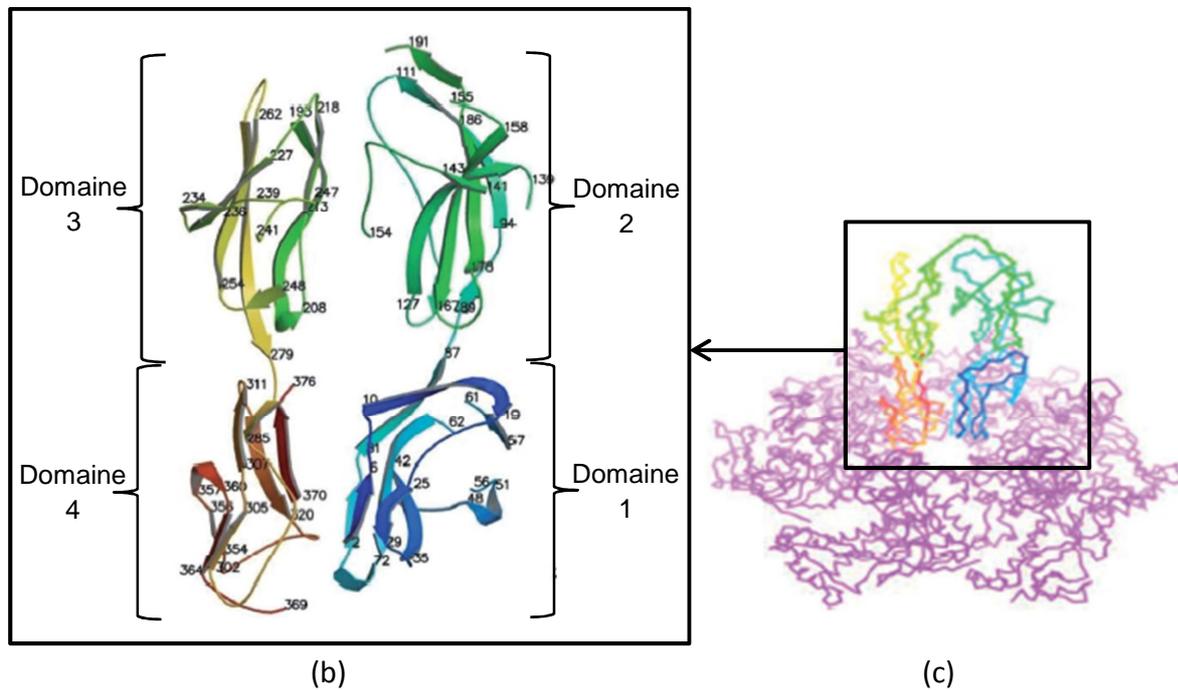
Une analyse biochimique a montré que le domaine C-terminal (domaine 4) de la protéine HOC de T4 contient un site de fixation à la capsid (Sathaliyawala, Islam et al.

2010). D'autres analyses biochimiques suggèrent que le domaine 1 stabilise le domaine 4, ces domaines interagiraient donc ensemble. L'hypothèse avancée par (Sathaliyawala, Islam et al. 2010) est que les deux domaines 1 et 4 formeraient le module de fixation de HOC sur la capsid, ainsi HOC aurait une forme de fer à cheval (figure 35 (b)). Le modèle structural de HOC obtenu par homologie de séquence avec l'Hémoline, une protéine d'insecte de type immunoglobuline, présente une forme similaire. De plus, cette forme est compatible avec la reconstruction tridimensionnelle à partir de clichés de cryo-microscopie électronique de la capsid (Fokine, Chipman et al. 2004) (figure (c)).

La structure atomique des 3 domaines N-terminaux de la protéine HOC du phage RB49 a été résolue par cristallographie (Fokine, Islam et al. 2011) (figure 36 (a)). La structure obtenue a une forme allongée. Elle montre un linker court et met en évidence des interactions stables entre les domaines 2 et 3 (figure 36 (b)). Le modèle par homologie de séquence de la protéine HOC de T4 basé sur la structure de HOC du phage RB49 suggère que le linker est également court chez la HOC de T4. Ces données sont incompatibles avec le modèle précédent en fer à cheval. Le modèle proposé est donc celui d'une protéine allongée (figure 36). La résolution de la structure de la capsid à une échelle subnanométrique permettrait de trancher entre ces deux modèles, à moins que la flexibilité des liaisons entre les domaines 1 et 2 et entre les domaines 3 et 4 n'aboutisse, par moyennage, à une densité électronique correspondant aux domaines 1 et 2 trop faible pour sortir du bruit de fond. Il est à noter que la partie C-terminale, c'est-à-dire le domaine de fixation à la capsid, n'a pas pu être cristallisée. L'explication avancée est que ce domaine soit relié au domaine 2 par un linker flexible, une autre possible serait que ce domaine n'est pas structuré en solution.



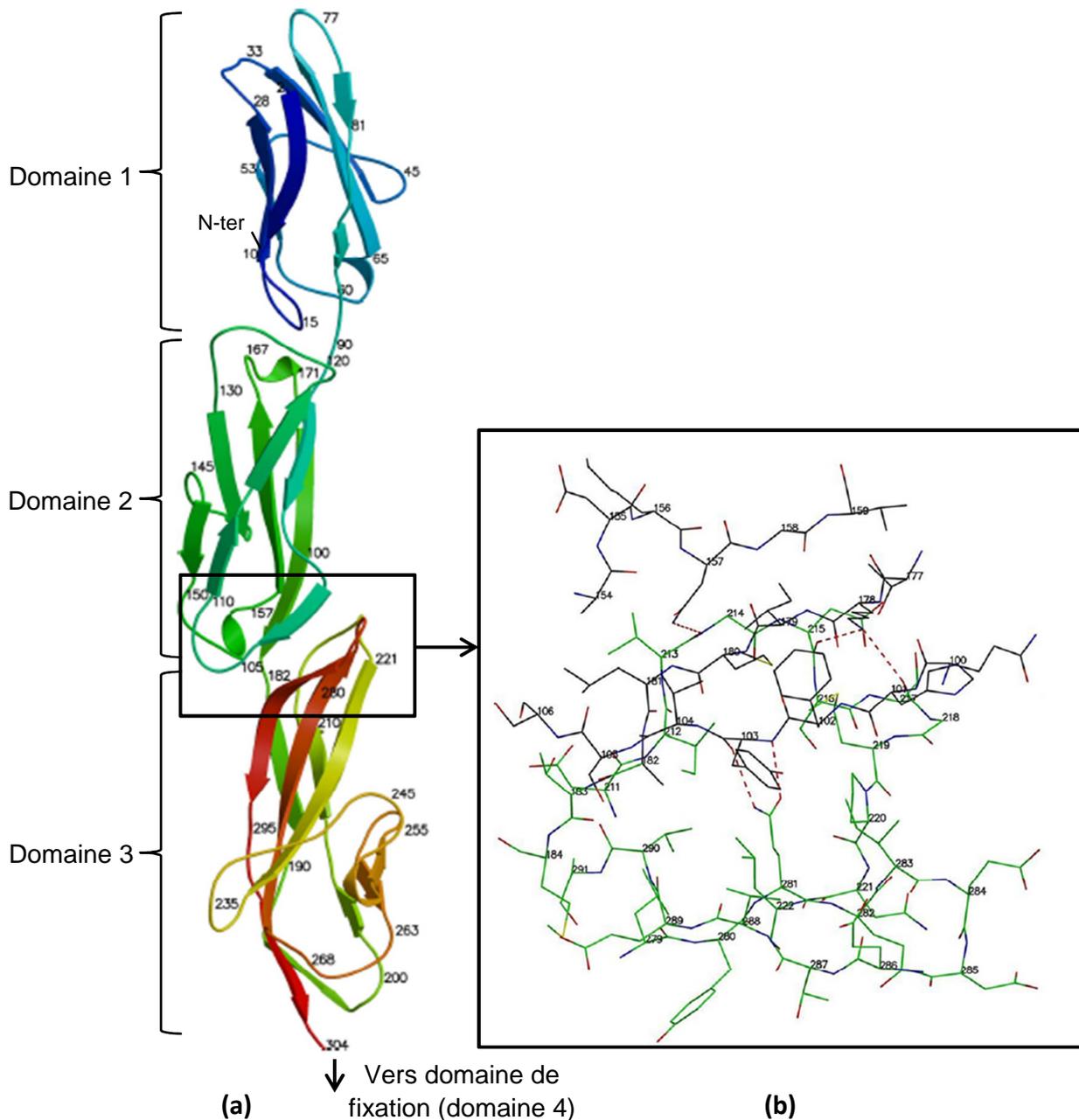
(a)



(b)

(c)

**Figure 35. Structure de la capside et modèle de la protéine HOC du bactériophage T4.** (a) Reconstruction tridimensionnelle à partir de clichés de cryo-microscopie électronique (résolution : 22 Å) de la capside du phage T4, montrant les protéines majeures de tête gp23 et gp24, formant respectivement les hexamères et les pentamères, les protéines accessoires SOC entre les hexamères et HOC au milieu de ceux-ci. (b) Modèle en fer à cheval de HOC obtenu par homologie de séquence avec l'Hémoline (structure obtenue par *Su et al.* (Su, Gastinel et al. 1998)). (c) Interaction de la protéine HOC de T4 avec l'hexamère de la protéine majeure de capside gp23. En rouge, jaune, vert et bleu : modèle en fer à cheval de HOC. En magenta : modèle de gp23 obtenu par homologie de séquence (Fokine, Leiman et al. 2005). (a) : figure adaptée de (Fokine, Chipman et al. 2004)), (b) et (c) : figures adaptée de (Sathaliyawala, Islam et al. 2010).



**Figure 36. Structure de la protéine HOC du bactériophage RB49**, un phage très proche de T4. (a) Seul les 3 domaines N-terminaux ont été cristallisés et sont présentés. Ces domaines présentent un repliement de type immunoglobuline. Le domaine C-terminal, domaine de fixation à la capside pourrait être relié par un linker flexible, ce qui pourrait empêcher la cristallisation de la protéine entière. (b) Carte atomique montrant les interactions entre les domaines 2 et 3 de la protéine HOC de RB49. Les atomes d'oxygène sont en rouge, les atomes d'azote en bleu et les atomes de soufre en orange. Les atomes de carbone du domaine 2 sont en noir, ceux du domaine 3 sont en vert. Figure adapté de (Fokine, Islam et al. 2011).

### 1-3 Rôle biologique de HOC

Cette protéine n'est pas indispensable à l'assemblage ni à l'infectiosité du phage (Ishii and Yanagida 1977). Elle ne se fixe que sur les capsides gonflées (Carrascosa 1978; Rao and Black 1985). Contrairement à SOC, la protéine HOC n'a pas d'effet significatif sur la stabilité de la capside. Son rôle biologique au cours du cycle de vie du virus n'est toujours pas déterminé, cependant plusieurs hypothèses ont été émises.

La protéine HOC pourrait se fixer sur les bactéries *E. coli*. Cette fixation, réversible, se ferait au moyen d'interactions faibles et non spécifiques entre les domaines immunoglobulines et les sucres de la paroi bactérienne (Fraser, Maxwell et al. 2007). Ces interactions pourraient conduire à plusieurs avantages sélectifs. D'une part, elles pourraient augmenter la concentration de virus près de l'hôte, et ainsi augmenter la chance d'infection. D'autre part elles pourraient permettre au phage de se servir d'une bactérie, qui ne soit pas nécessairement son hôte, comme « moyen de transport ». Les bactéries utilisant la chimiotaxie pour se diriger vers des environnements dans lesquelles les conditions physico-chimiques sont favorables (Adler and Tso 1974), cette fixation pourrait être plus efficace que la simple diffusion pour trouver des hôtes.

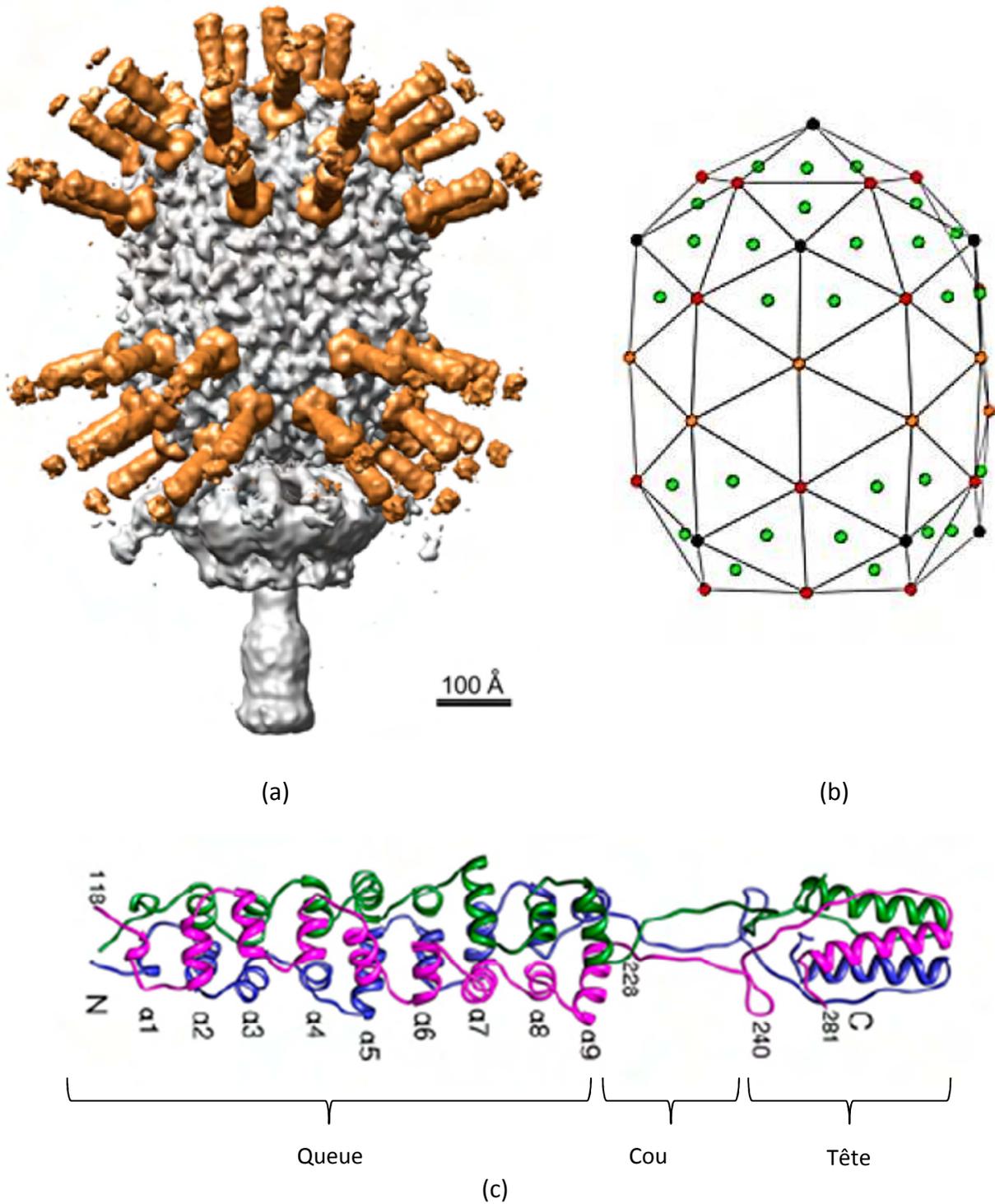
Par ailleurs, il a été montré que les mutants de T4 dépourvus de HOC s'agrègent dans des solutions dépourvues d'ions divalents, contrairement aux phages T4 natifs (Yamaguchi and Yanagida 1980) (Childs 1980). La protéine HOC préviendrait donc cette agrégation, probablement grâce à son point isoélectrique faible ( $pI = 4,17$ ). La gêne stérique induite par les 60 Å de long de HOC pourrait aussi empêcher les interactions entre les protéines majeures gp23 de deux phages voisins.

## 2- Les fibres de la capsid de $\phi 29$

Le bactériophage  $\phi 29$  est un phage caudé de la famille des *Podoviridae* infectant les bactéries à Gram positif *Bacillus subtilis* (Salas, Vasquez et al. 1972). Sa capsid de symétrie pseudo-icosaédrique est oblongue et a pour nombres de triangulation  $T = 3$  et  $Q = 5$  (Figure 37 (a) et (b)). Elle est décorée de 55 fibres constituées de trimères de la protéine gp8.5. Ces fibres, de 150 Å de long, se fixent à l'interface d'un pentamère et de deux hexamères.

La structure cristallographique de la partie C-terminale de ces fibres a été récemment déterminée à une résolution de 1,6 Å (Xiang and Rossmann 2011). Cette partie correspond à la zone exposée au solvant, le domaine de fixation n'ayant pas pu être cristallisé. Elle est longue de 110 Å et est divisée en trois domaines distincts : la tête, le cou et la queue (figure 37 (c)). La queue est composée de 9 hélices  $\alpha$  reliées entre elles par des coudes  $\beta$ . Ces coudes présentent des angles dièdres différents de tous les coudes  $\beta$  précédemment définis.

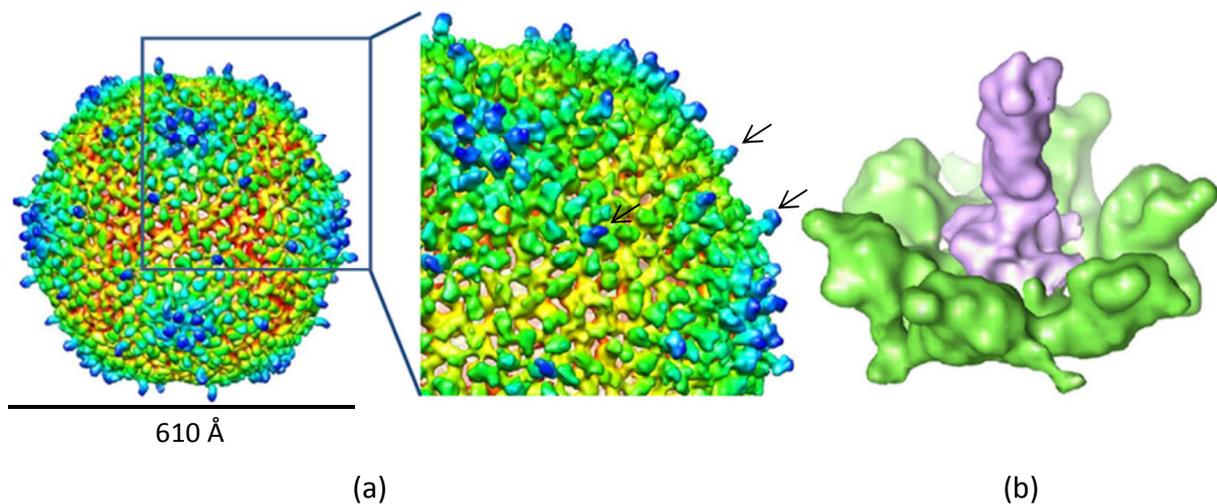
Comme pour la protéine HOC des phages de la famille de T4, le rôle sélectif des fibres de  $\phi 29$  n'est pas connu. Aucun effet sur la stabilité de la capsid n'a pu être démontré. La délétion du gène codant pour ces fibres n'engendre aucune baisse de l'infectiosité de  $\phi 29$  *in vitro*. Cependant des données de microscopie électronique suggèrent que cette fibre aide le phage à reconnaître la surface de la bactérie. Ces données n'ont toutefois été confirmées par aucune analyse biochimique.



**Figure 37. Structure du phage  $\phi 29$  et de la sa fibre.** (a) Reconstruction tridimensionnelle à partir de clichés de cryo-microscopie électronique (résolution : 25 Å) du phage  $\phi 29$ . Les fibres de la capside sont affichées en orange. (b) Représentation schématique de la capside pseudo icosaédrique. Les sites de fixation des fibres sont représentés par des points verts. Les points noirs représentent les centres des pentamères, les points rouges et orange les centres des hexamères. (c) Structure cristallographique à 1,6 Å de résolution de la partie C-terminale d'une fibre, composée de trois copies de la protéine gp8.5. Cette partie a une longueur de 110 Å. Figure adaptée de (Xiang and Rossmann 2011).

### 3- La protéine gp12 du phage SPP1

Le bactériophage SPP1 est un phage caudé de la famille des *Siphoviridae* infectant les bactéries à Gram positif *Bacillus subtilis*. Sa capside de symétrie icosaédrique a pour nombre de triangulation  $T = 7$ . Elle est décorée de 60 trimères de gp12 (figure 38), une protéine de 6.6 kDa. Ces trimères ne se fixent qu'après l'expansion de la capside, au centre des hexamères (Isidro, Henriques et al. 2004) (White, Sherman et al. 2012). L'effet de la fixation de cette protéine sur la stabilité thermique de la capside a été récemment étudié par mesure de la fluorescence pendant un gradient de température. L'intensité de fluorescence est liée à la présence d'ADN libre en solution et permet donc de mesurer la quantité de capside dont l'ADN s'est échappé. Cette étude a montré que la fixation de gp12 ne participait pas à la stabilisation de la capside. Son rôle biologique est donc pour l'instant inconnu.



**Figure 38. Structure de la capside du phage SPP1.** (a) Reconstruction tridimensionnelle de la capside pleine d'ADN de SPP1 à partir de cryo-microscopie électronique (résolution : 11.7 Å). À gauche une vue zoomée de cette capside montrant les zones de densité électronique correspondant aux trimères de gp12 (flèches noires). (b) En magenta : densité électronique correspondant à un trimère de gp12. Cette densité a été obtenue en soustrayant la carte de densité précédente à celle obtenue avec un mutant produisant des capsides non décorées. En vert : densité électronique correspondant à un hexamère de la capside. Figure adaptée de (White, Sherman et al. 2012).

## 4- Applications biotechnologiques des protéines de décoration

Le *phage display* est une méthode d'étude des interactions protéine-protéine ou protéine-ADN. Elle est basée sur la présentation d'un polypeptide à la surface d'un bactériophage (Smith and Petrenko 1997). Pour cela, ce polypeptide est fusionné sur une protéine structurale de ce phage. Les protéines d'enveloppes des phages filamenteux M13 et fd ont notamment été abondamment utilisées pour exposer des banques de peptides comportant de six à huit acides aminés (Smith 1985). Les phages icosaédriques ont également été largement utilisés, notamment la protéine de décoration gpD (420 copies) (Sternberg and Hoess 1995), la protéine majeure de queue gpV (192 copies) (Maruyama, Maruyama et al. 1994) du phage  $\lambda$  et la protéine majeure de capsid gp10 (415 copies) (Castillo, Goodson et al. 2001) du phage T7. Le fait que les protéines du phage utilisées pour exposer le polypeptide soient des protéines structurales indispensables à l'assemblage du phage présente néanmoins une limitation, notamment liée à la taille de ce polypeptide (Smith and Petrenko 1997; Manoutcharian, Gevorkian et al. 2001).

La protéine de décoration HOC du phage T4, les fibres du phage  $\phi$ 29, la protéine de décoration gp12 du phage SPP1 ainsi que la protéine de décoration pb10 du phage T5 ne sont en revanche pas essentielles pour l'assemblage des virions. La fixation de ces protéines peut donc se faire *in vitro* après assemblage des virions ou des capsides seules, et non *in vivo* comme précédemment. Cette méthode permettrait de bénéficier de l'exposition, de la distribution symétrique et de l'abondance des protéines de décoration à la surface des capsides sans être limitée par la taille du polypeptide fusionné.

Ces propriétés rendent les protéines de décoration excellentes candidates pour les fusionner avec des antigènes afin de générer une réponse immunitaire. Il a notamment été montré qu'il était possible de fusionner la protéine HOC avec l'antigène protecteur de l'anthrax, puis de fixer avec une grande affinité ( $K = 69\text{nM}$ ) cette molécule sur des virions dépourvus de protéine HOC (Shivachandra, Rao et al. 2006). Ces virions ainsi marqués déclenchent une réaction immunitaire chez la souris même en l'absence d'adjuvant.

## 5- Présentation de la thématique de recherche

pb10 est la protéine de décoration du phage T5. C'est une protéine de 17 kDa, non essentielle pour l'assemblage et l'infectiosité des virions (Saigo 1978). Son rôle dans le cycle de vie du virus n'est pas connu. Comme les protéines HOC de T4, elle se fixe au centre de chaque hexamère de pb8, il y a donc 120 copies de pb10 par capsid (Effantin, Boulanger et al. 2006). Des analyses bio-informatiques ont montré (Fraser, Yu et al. 2006) que pb10 comprend un domaine immunoglobuline possédant 43% d'identité de séquence avec celui de la protéine HOC de T4, alors qu'il n'y a pas de similarité de séquence entre les autres régions de ces protéines.

Le but de ce projet a été de :

- produire et purifier pb10
- caractériser pb10 en solution
- étudier la fixation *in vitro* de pb10 sur les capsides vides d'ADN. En particulier, de vérifier leur fixation sur les capsides gonflées et non sur les capsides compactes, puis d'analyser l'affinité et la coopérativité de sa fixation sur la capsid gonflée

- étudier son rôle biologique, notamment sa capacité à stabiliser la capsid
- déterminer sa structure. Plusieurs approches ont été effectuées, à basse résolution avec une étude par SAXS et à haute résolution avec une tentative de cristallogénèse de pb10.

## **II- RESULTATS**

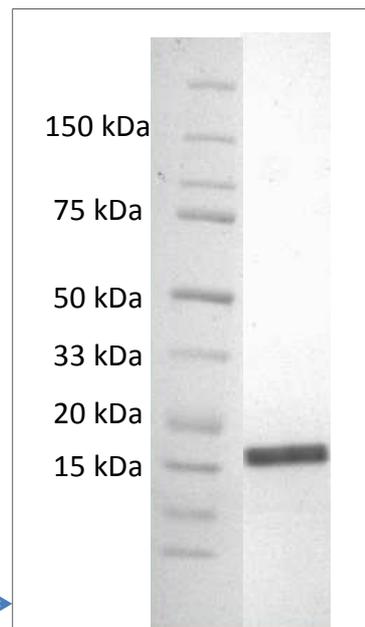
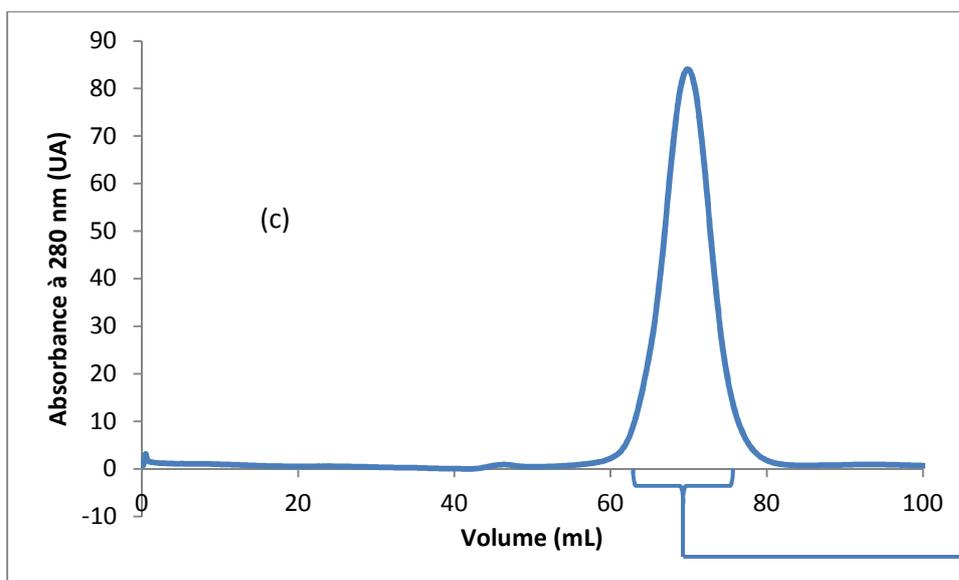
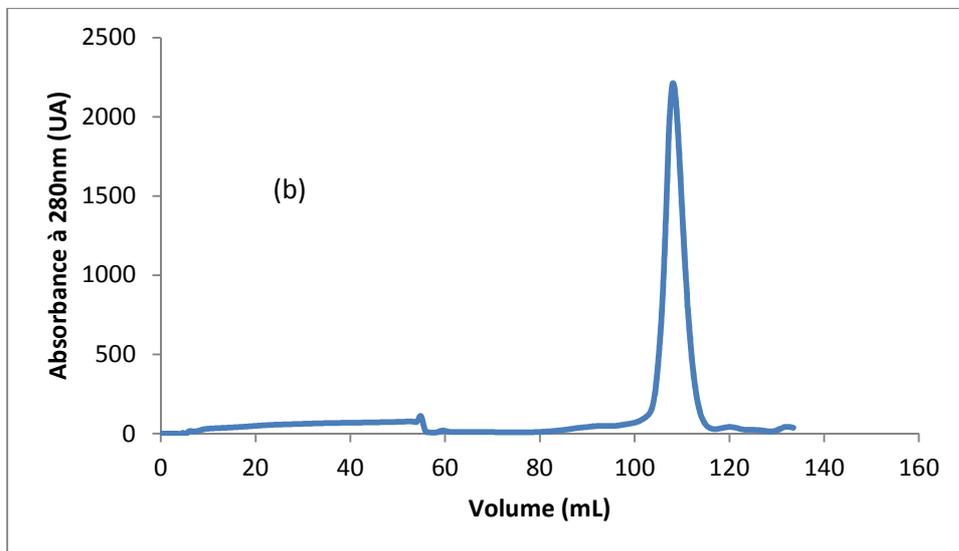
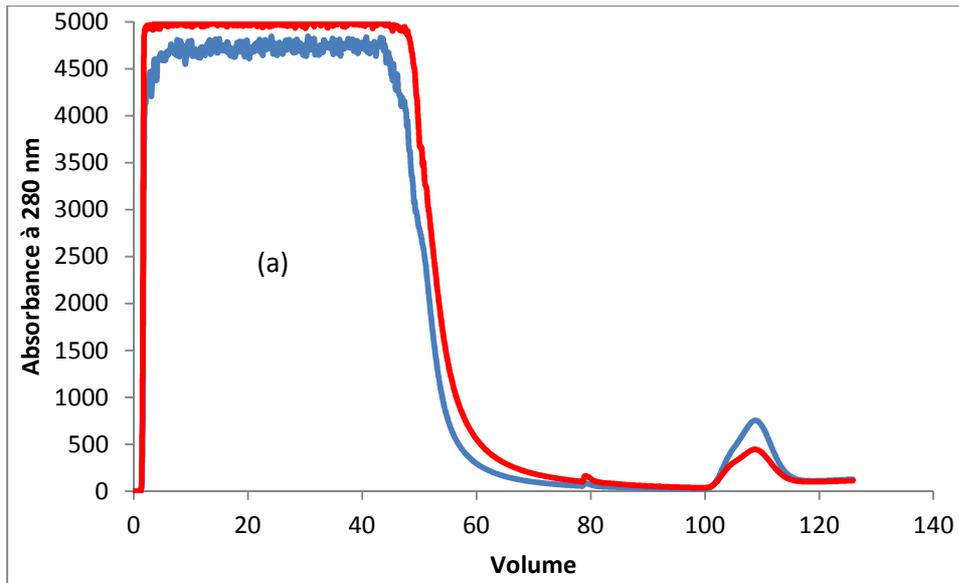
### **1- Clonage, production et purification de pb10**

#### **1-1 Production**

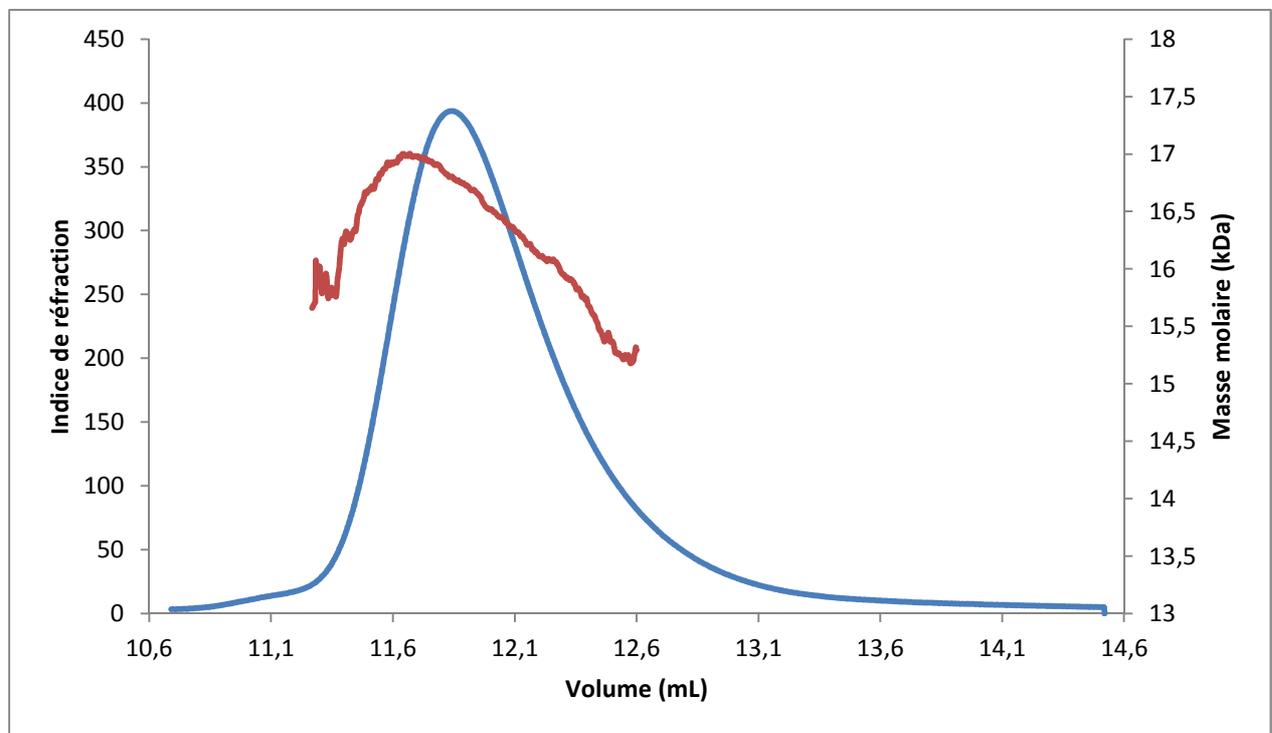
Le gène codant pour pb10 a été cloné dans le vecteur de surexpression bactérien pET28b (Cf. Matériel et Méthodes). Une étiquette codant pour 6 histidines est ajoutée en 3' du gène pb10. La protéine pb10 a été produite dans la souche d'expression *E. coli* BL21. Son niveau d'expression et sa solubilité sont compatibles avec les différentes études envisagées : étude de la fixation sur la capsid par ultracentrifugation et étude structurale par cristallographie et SAXS.

#### **1-2 Purification**

La purification des protéines se déroule comme indiqué dans la partie Matériel et Méthode. Pour obtenir une protéine pure, trois étapes sont nécessaires : chromatographie d'affinité sur un métal immobilisé (IMAC), chromatographie d'échange de cations et chromatographie d'exclusion de taille (figure 39). La protéine est concentrée à une

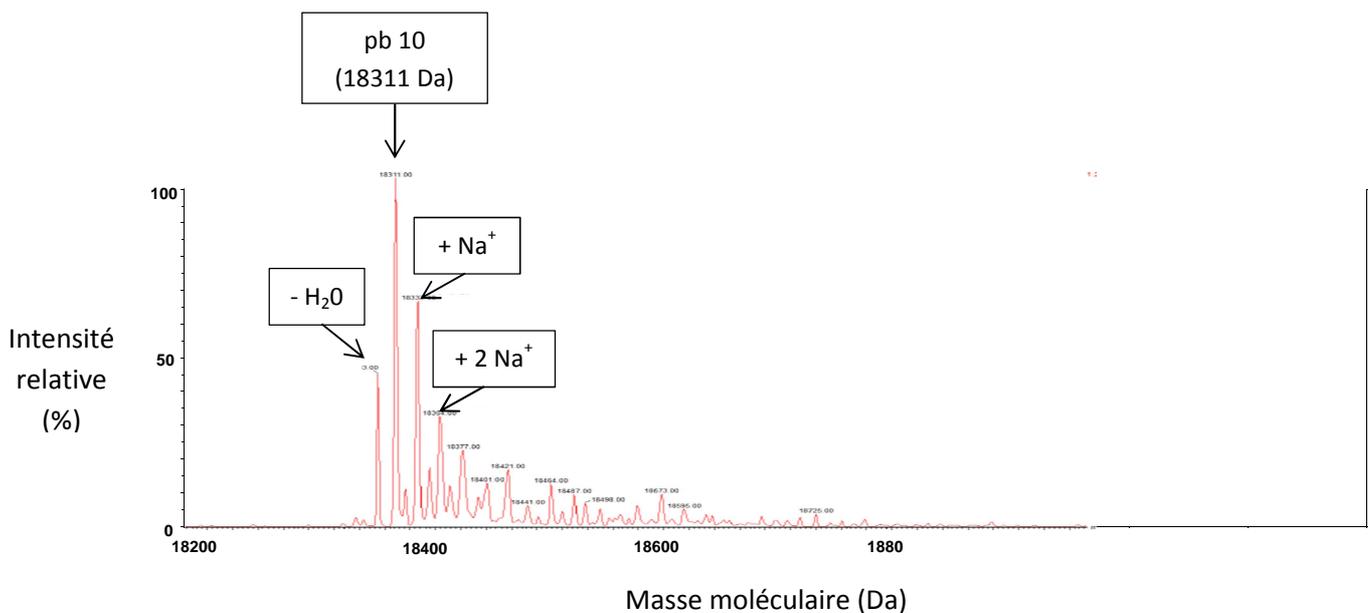


**Figure 39. Purification de pb10.** Après cassage des bactéries à la presse de French, la fraction soluble obtenue par ultracentrifugation de la suspension bactérienne est injectée sur une colonne de chromatographie d'affinité (HisTrap FF, GE Healthcare) (a). La colonne est lavée avec du tampon 50 mM Hepes pH 7.2, 1M NaCl pour éliminer l'ADN fixé non spécifiquement sur la colonne et sur pb10. L'élué se fait par gradient 0-40% (Solution A : 50 mM Hepes pH7.2, 50 mM NaCl, 20mM imidazole, solution B : 50 mM Hepes pH7.2, 50 mM NaCl, 1M imidazole). Les fractions d'intérêt sont ensuite injectées dans une colonne de chromatographie d'échange de cations (HiTrap SP HP) (b) et éluées par un gradient de concentration en NaCl. Les fractions contenant la protéine sont regroupées et concentrées à 5 mL puis injectées sur une colonne de chromatographie d'exclusion de taille Sephadex S75 16/60 HL (GE Healthcare) (c). A l'issue de cette dernière étape, les fractions contenant pb10 sont regroupées et dialysées contre 50 mM Hepes pH7.2, 150 mM NaCl, puis concentrées. La concentration finale varie de 3 à 32 mg/mL selon les expériences effectuées.



**Figure 40. Caractérisation de pb10 par SEC-MALLS.** La masse moléculaire (courbe rouge) est calculée à partir de l'intensité de la lumière diffusée en fonction de l'angle de diffusion et de l'indice de réfraction (courbe bleue). Cette masse moléculaire est estimée à  $16 \pm 2$  kDa, ce qui indique que la protéine est sous la forme monomérique.

concentration variant de 3 à 32 mg/mL selon les expériences effectuées. Le rendement moyen est de 2 mg de pb10 purifiée par litre de culture. La pureté de l'échantillon protéique est ensuite contrôlée par gel SDS PAGE. L'état d'oligomérisation de la protéine est déterminé par SEC-MALLS (Size Exclusion Chromatography - Multi-Angle Laser Light Scattering), technique qui permet de déterminer de manière approchée la masse molaire. Cette analyse nous indique que la protéine est monomérique (figure 40), la masse indiquée,  $16 \pm 2$  kDa, étant de l'ordre de grandeur de la masse prédite par la séquence. L'analyse par spectrométrie de masse (Q-TOF) de pb10 a ensuite montré l'intégrité de cette protéine purifiée (figure 41).



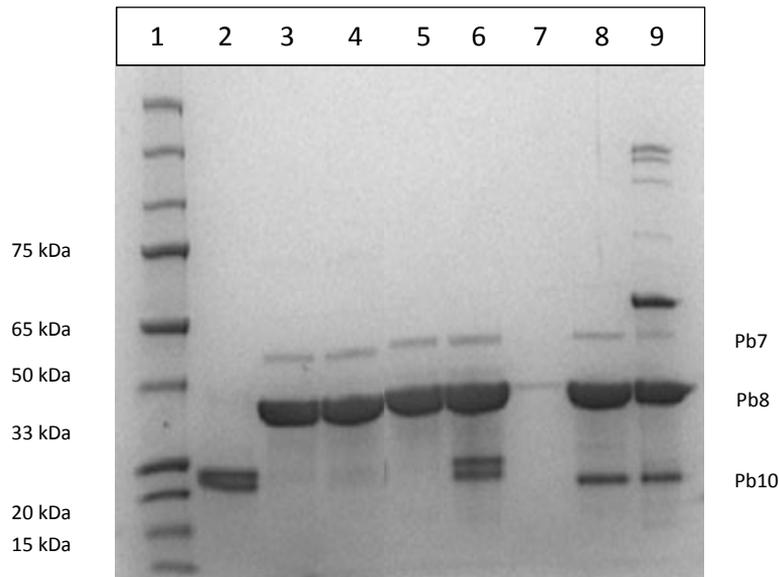
**Figure 41. Spectrométrie de masse (Q-TOF) de la protéine pb10.** L'incertitude absolue sur la mesure est de 2 Da. On observe plusieurs pics. Le pic d'intensité relative de 100 % correspond à pb10 purifiée seule, les autres pics correspondent à la protéine plus ou moins une molécule d'eau ou un ion sodium. La masse de pb10 déterminée par cette méthode, 18311 daltons, est exactement celle prédite à partir de sa séquence : pb10 n'est donc pas protéolysée.

## **2- Etude fonctionnelle de pb10**

### **2-1 Etude de la fixation de pb10 sur les capsides vides d'ADN**

#### **2-1-1 Etude de la fixation de pb10 par ultracentrifugation**

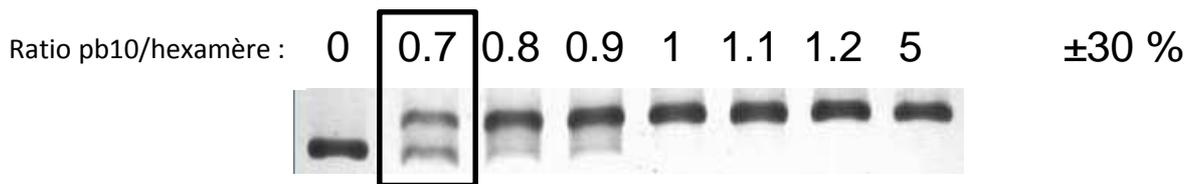
La fixation de pb10 sur les capsides vides d'ADN a été étudiée *in vitro* par ultracentrifugation. Les capsides compactes et gonflées ont été incubées avec la protéine pb10, puis ultra-centrifugées à une vitesse qui permet la sédimentation des capsides mais ne permet pas la sédimentation de la protéine pb10. Lorsqu'elle est incubée avec des capsides gonflées, la protéine pb10 sédimente dans le culot ; en revanche après incubation avec les capsides compactes elle reste dans le surnageant (figure 42). Cette expérience prouve qu'*in vitro* la protéine pb10 se fixe sur les capsides matures vide d'ADN mais ne se fixe pas sur les capsides compactes. Leur fixation est donc un marqueur du processus d'expansion. Cette fixation *in vitro* montre que les protéines pb10 produites avec une étiquette poly-histidine sont fonctionnelles. L'étiquette étant insérée du côté C-terminal, cela renforce l'idée que le site de fixation se situe sur le domaine N-terminal de pb10. C'est également un élément de preuve que l'expansion *in vitro* des capsides conduit au même état final que l'expansion *in vivo*. En effet, *in vivo* on observe également que pb10 ne se fixe pas sur les capsides compactes, car les capsides produites par le mutant AMN5 ne sont pas décorées alors que pb10 est produite par ce mutant.



**Figure 42. Etude de la fixation de la protéine de décoration par ultracentrifugation.** 5 échantillons ont été centrifugés : la protéine pb10 seule, les capsides gonflées (CG), les capsides compactes (CC), le mélange capsides gonflées plus pb10 et le mélange capsides compactes plus pb10. Les différentes pistes sont : 1: marqueur de taille, 2 : surnageant pb10, 3 : culot CC, 4 : culot CC + pb10, 5 : culot CX, 6 : culot CX + pb10, 7 : surnageant CX + pb10. Les deux dernières pistes sont des témoins. La piste 8 contient des têtes pleines produites séparément des queues et la piste 9 contient des fantômes de phage T5. Sur ces deux témoins nous retrouvons les bandes correspondantes à la protéine portale pb7, à la protéine majeure de tête pb8 ainsi qu'à pb10. Cependant, la migration de cette dernière protéine est plus basse compte tenu du fait qu'elle ne possède pas d'étiquette poly-histidine. Nous n'avons pas d'explication pour expliquer la migration en doublet de pb10 purifiée.

## 2-1-2 Etude par migration sur gel d'agarose

La fixation de pb10 sur les capsides gonflées a ensuite été étudiée par migration sur gel natif d'agarose (figure 43), ce qui permet d'analyser des objets macromoléculaires de la taille des capsides en les faisant migrer en fonction de leur taille et de leur charge exposée au solvant. On observe un changement de la vitesse de migration des capsides gonflées ayant été incubées avec pb10 comparativement aux capsides gonflées sans traitement ; les capsides décorées peuvent donc être discriminées des capsides non décorées par cette technique expérimentale. Ce changement de vitesse de migration peut s'expliquer par le changement de la charge exposée au solvant, pb10 se fixant sur la surface des capsides.



**Figure 43. Analyse de la fixation de pb10 sur les capsides gonflées par gel natif d'agarose.** Les capsides ont été incubées avec des concentrations croissantes de pb10, mesurées par le rapport entre le nombre de pb10 et le nombre d'hexamères de pb8 en solution. La concentration en capside est constante et égale à 2  $\mu\text{mol/l}$  (en mole d'unités fixantes, i.e. d'hexamères).

Différentes quantités de protéine pb10 ont été incubées avec les capsides. Les différentes expériences réalisées montrent que toutes les capsides migrent à la vitesse des capsides décorées dès lors que le rapport entre la concentration de pb10 et celle des hexamères est compris entre 1 à 1.6. Cette variabilité est due à l'imprécision dans la détermination de ce rapport (voir matériels et méthodes). D'autre part, les bandes correspondant aux états non décorés et totalement décorés sont beaucoup plus intenses que les bandes correspondant aux états partiellement décorés. Nous allons montrer que cela implique une coopérativité dans la fixation de pb10.

### 2-1-3 Analyse de l'affinité et mise en évidence de la coopérativité de la fixation

En faisant l'hypothèse d'une fixation non coopérative de pb10 sur les capsides, nous allons montrer que nous ne pouvons pas avoir pour le même échantillon 2 bandes correspondant aux états non décorés et totalement décorés.

On considère la réaction de fixation de la protéine pb10 avec un hexamère. Etant non coopérative par hypothèse, cette fixation ne dépend pas de l'état des 119 autres sites de fixation. On appelle  $a$  la concentration en hexamère libre, c'est-à-dire non fixé par pb10,  $b$  la concentration en protéine pb10 libre,  $c$  la concentration en hexamères liés à pb10,  $d$  la

concentration totale de pb10 (liée et libre) en solution et  $h$  la concentration totale d'hexamères.

La constante d'affinité de la réaction est :

$$K = \frac{a * b}{c} \quad (9)$$

De plus la concentration totale en hexamères est égale à la somme de la concentration en hexamère libre et de la concentration en hexamères liés :

$$a + c = h \quad (10)$$

De même pour la concentration totale en pb10 :

$$b + c = d \quad (11)$$

Ce qui donne donc, après avoir remplacé dans l'équation (9)  $a$  et  $b$  par leur expression issue de (10) et (11) :

$$cK = (h - c)(d - c)$$

La concentration en hexamères liés à pb10 est donc :

$$c = \frac{h + d + K - \sqrt{(h + d + K)^2 - 4hd}}{2}$$

Ainsi la probabilité qu'un hexamère soit lié à pb10 vaut :

$$p = \frac{c}{h} = \frac{h + d + K - \sqrt{(h + d + K)^2 - 4hd}}{2h} \quad (12)$$

En utilisant la loi binomiale, nous pouvons déterminer la probabilité  $P_n$  d'avoir  $n$  pb10 sur une capsid :

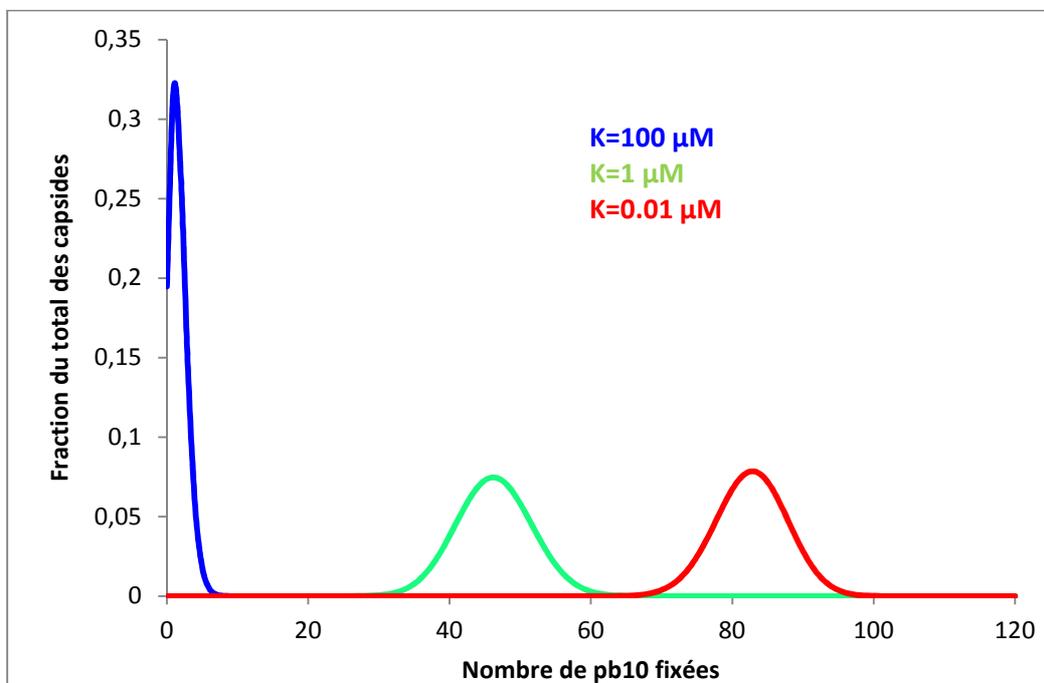
$$P_n = \frac{120!}{n! (120 - n)!} * p^n * (1 - p)^{120-n} \quad (13)$$

Donc le cas d'un grand nombre de capsides, le nombre de capsides ayant fixées  $n$  pb10 est :

$$N_n = \frac{h * V * N_A}{120} * P_n$$

$N_A$  étant le nombre d'Avogadro et  $V$  le volume de l'échantillon.

Considérons le mélange suivant : 2  $\mu\text{mol/l}$  de capsides (en mole d'unités fixantes, i.e. d'hexamères) et 1,4  $\mu\text{mol/l}$  de pb10. La migration sur gel d'agarose natif (figure 43, deuxième piste) nous montre que les états totalement décorés et non décorés sont majoritairement peuplés. Or, quelle que soit la constante d'équilibre de la fixation, le calcul du pourcentage de capsides ayant fixé  $n$  pb10 effectué à partir des équations (12) et (13) (figure 44) nous montre que les états totalement décorés et non décorés ne peuvent pas coexister dans un même échantillon. L'hypothèse de départ était donc fautive, et il est très probable que la fixation de pb10 sur la capside soit coopérative. Des expériences de Résonance Plasmonique de Surface (SPR) sont envisagées pour vérifier cette coopérativité ainsi que pour déterminer la constante d'affinité.



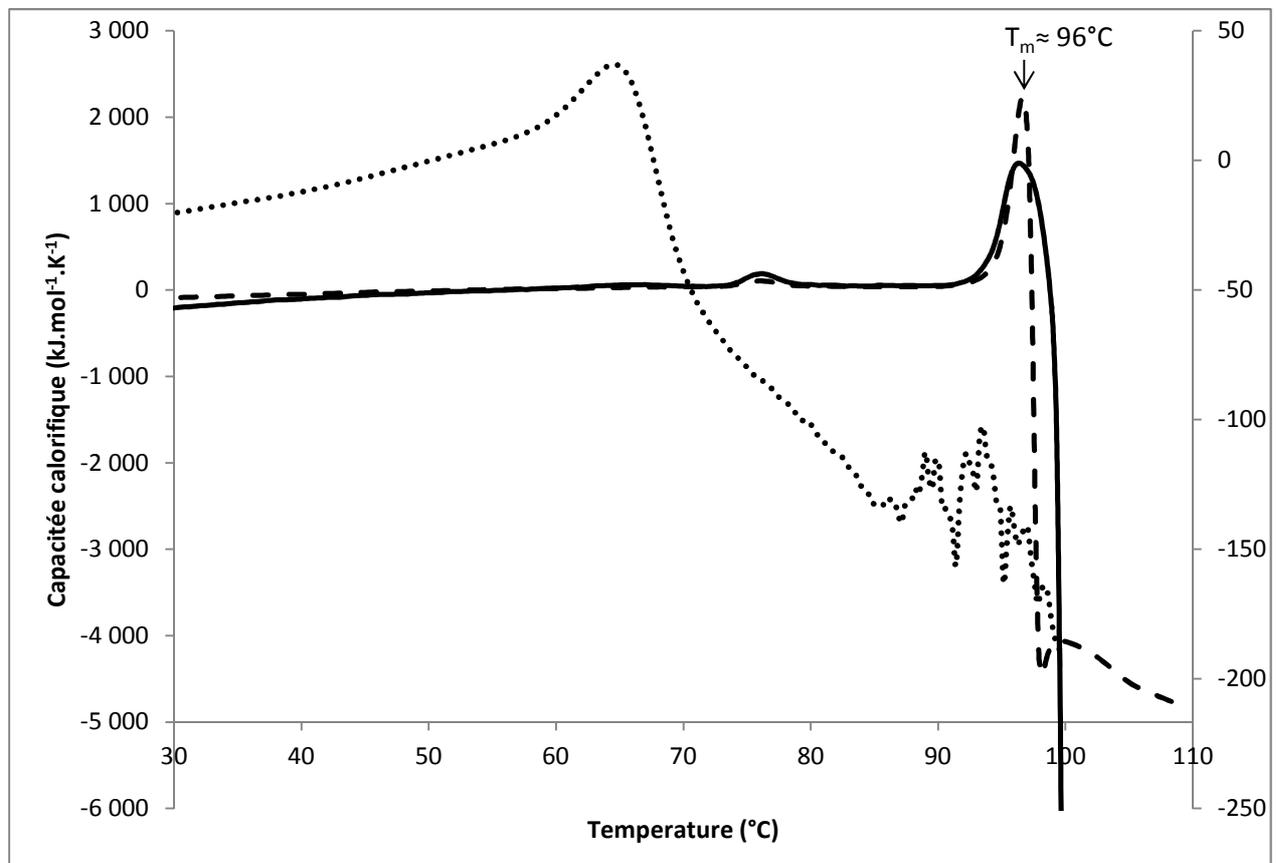
**Figure 44. Calcul de la fraction de capsides ayant fixées  $n$  pb10 en fonction du nombre  $n$  de pb10 fixés.** Ce graphique se lit ainsi : pour une constante de réaction de  $K=0.01 \mu\text{M}$ , 8% des capsides sont décorées par 82 pb10, 0% par 45 pb10 ; pour une constante de  $K=1 \mu\text{M}$ , 7% des capsides sont décorées par 45 pb10, 0% par 82 pb10. Ce calcul a été effectué en prenant

une concentration de capsidite de 2  $\mu\text{mol/l}$  (en mole d'unités fixantes, i.e. d'hexamères) et une concentration en pb10 de 1,4  $\mu\text{mol/l}$ . Il est effectué en utilisant les équations (12) et (13), équations dérivées en faisant l'hypothèse d'une fixation non coopérative.

## **2-2 Etude par calorimétrie différentielle à balayage du rôle de la protéine de décoration**

Chez certains phages des protéines de décoration se fixent aux points de convergences de trois hexamères et améliorent la solidité des capsides en soudant entre eux ces hexamères (Gilcrease, Winn-Stapley et al. 2005) (Qin, Fokine et al. 2010). Chez T5, la protéine de décoration pb10 se fixe aux sommets des hexamères. Son rôle est inconnu. Nous avons cherché à déterminer par calorimétrie différentielle à balayage (DSC) si cette protéine a un rôle stabilisateur, c'est-à-dire s'il faut plus d'énergie pour dissocier une capsidite après la fixation de la protéine pb10 (figure 45). Nous avons donc comparé les courbes de dénaturation thermique des capsides gonflées avec et sans décoration par pb10. Pour ces deux types de capsides on remarque une baisse drastique de la capacité calorifique après 95°C. Cette chute de la capacité calorifique après la température  $T_m$  est révélatrice de l'agrégation de la protéine majeure de tête et donc de la dislocation des capsides à cette température. A cause de cette agrégation, la valeur exacte de la température de transition  $T_m$  n'a pas pu être déterminée. Toutefois elle peut être approximée à la température à laquelle le maximum de la courbe est atteint. Cette température est exceptionnellement importante au regard de celle d'autres capsides de phages : 81°C pour la capsidite du phage HK97 (Conway, Cheng et al. 2007) et 74°C pour la capsidite du phage  $\phi 6$  (Nemecek, Cheng et al. 2011). L'étape d'expansion permet de faire passer cette température de 75 à 96°C. En revanche la fixation de pb10 n'induit pas d'augmentation ni de diminution de la température de transition. Cette

étude montre donc que la décoration de la capside par pb10 n'améliore pas la stabilité thermique de la capside.



**Figure 45. Etude du rôle de la protéine de décoration par calorimétrie différentielle à balayage (DSC).** Courbes de dénaturation thermique de la capside gonflée seule (trait en tiret, échelle de droite), du mélange capsides gonflées + pb10 (trait continu, échelle de droite) et pb10 seule (trait en pointillés, échelle zoomée à gauche). Pour aider à la comparaison, les valeurs de la capacité calorifique des capsides avec et sans décoration sont données en kilojoule par degré Kelvin et par mole d'hexamères de pb8. La baisse drastique de la capacité calorifique après la température  $T_m$  est révélatrice de l'agrégation de la protéine majeur de tête et donc de la dislocation des capsides à cette température. A cause de cette agrégation, la valeur exacte de la température de transition  $T_m$  ne peut pas être déterminée. Toutefois elle peut être approximée à la température à laquelle le maximum de la courbe est atteint. Cette température ne varie pas significativement après fixation de pb10. A cause de cette agrégation la valeur de l'enthalpie de transition  $\Delta H_m$  ne peut pas non plus être déterminée. La petite bosse à 75°C correspond à la dislocation de la petite (<2%) fraction de capsides compactes présente dans l'échantillon.

### 3- Etude structurale de pb10

La cristallographie et la diffusion des rayons X aux petits angles (SAXS) sont deux méthodes complémentaires pour l'étude structurale des protéines. La cristallographie permet l'obtention de la structure à résolution atomique, mais nécessite l'obtention de cristaux ; en revanche l'étude par SAXS s'effectue en solution et a pour but d'obtenir des données structurales de basse résolution, et notamment d'avoir accès à l'enveloppe de cette protéine

#### 3-1 Etude par SAXS

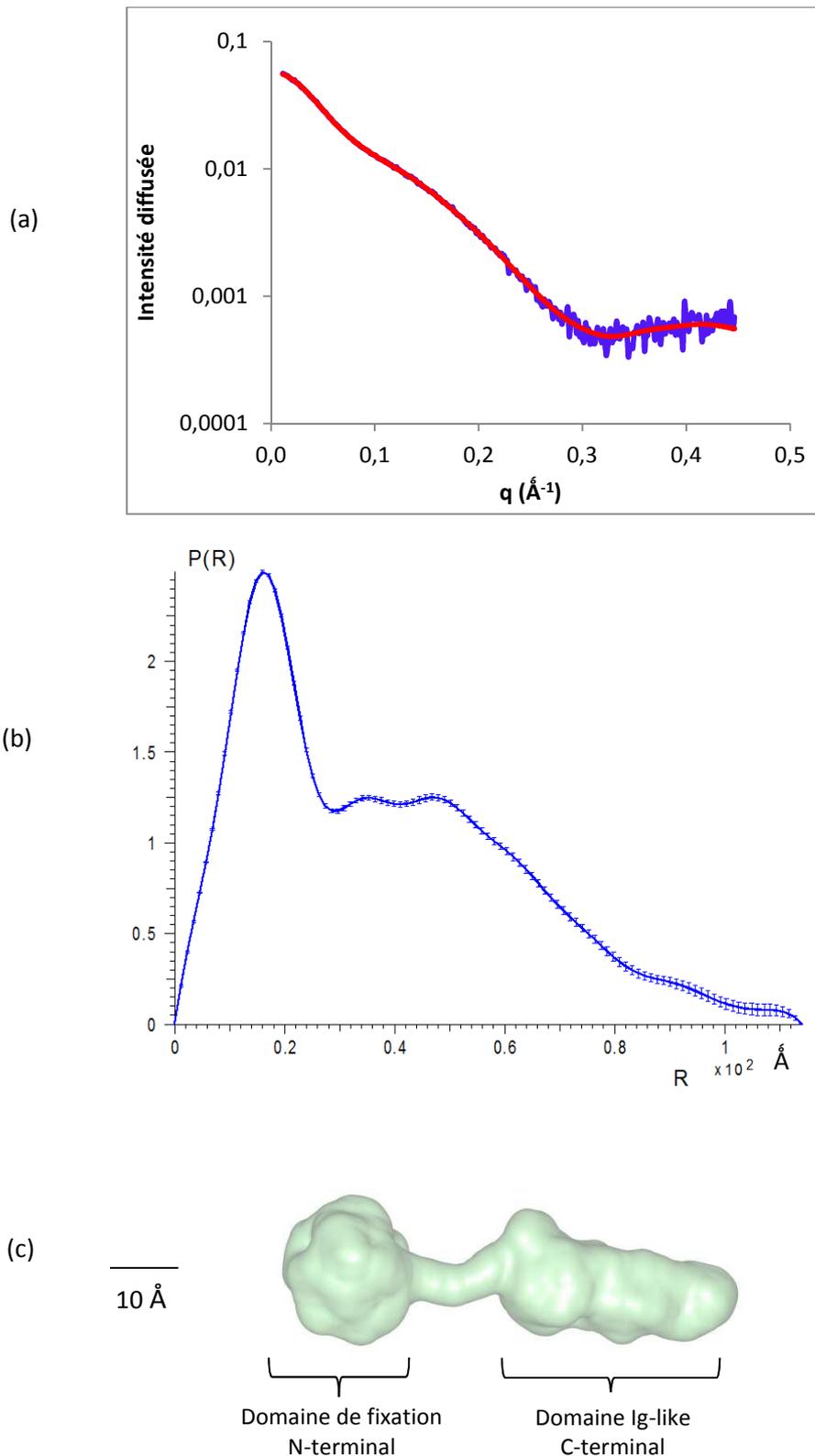
##### 3-1-1 Analyse de la courbe de diffusion et de la fonction de distribution des distances

L'application de la loi de Guinier (cf. introduction) à la courbe de diffusion (figure 46 (a)) conduit à des valeurs du rayon de giration  $R_g$  et de l'extension maximale  $D_{max}$  de pb10 de 29.8 Å et 112 Å respectivement. Ces valeurs sont très nettement supérieures aux valeurs attendues pour une protéine globulaire de 173 résidus ; pour une protéine globulaire les formules empiriques sont :  $R_g \sim 3(n)^{1/3}$  Å , où n est le nombre de résidus, et  $D_{max} = 2,6 R_g$ . Ces formules donnent ici des valeurs de 17 Å pour le  $R_g$  et 43 Å pour le  $D_{max}$ . Il a été vérifié à partir de la valeur de l'intensité diffusée à angle nul  $I(0)$  que la solution contenait exclusivement des monomères et pas d'oligomère : la masse de l'espèce diffusante déduite de  $I(0)$  est exactement égale à la masse théorique de 18500 Da si l'on considère que le volume spécifique de la protéine est de 0.75 cm<sup>3</sup>/g, et une masse même inférieure de 10% si l'on prend pour le volume spécifique la valeur calculée à partir de la séquence ( $v=0.73$  cm<sup>3</sup>/g).

La détermination de la fonction de distribution des distances  $P(r)$  confirme que la protéine n'est pas globulaire (figure 46 (b)). La  $P(r)$  est caractéristique d'un objet allongé mais la présence du pic fin et bien marqué centré en  $r \approx 18\text{\AA}$  impose que la protéine possède un ou des globules de petite taille.

### 3-1-2 Reconstitution de l'enveloppe moléculaire

Ces observations qualitatives sont quantifiées à partir de l'analyse *ab initio* de la courbe de diffusion à l'aide du programme GASBOR, qui assimile la protéine à une chaîne de pseudo-résidus distants de  $3.8\text{\AA}$ . La position de ces pseudo-résidus est obtenue en ajustant la courbe calculée sur la chaîne de pseudo-résidus à la courbe expérimentale, ce qui donne une enveloppe de la protéine. Une vingtaine d'enveloppes ont ainsi été obtenues, et leur similarité est tout à fait remarquable : le NSD (Normal Spatial Discrepancy) qui chiffre les écarts entre les différentes enveloppes, est de 0.8, ce qui est une valeur très faible, prouvant la grande similitude des enveloppes. L'enveloppe typique présentée sur la figure 46 (c) montre que la protéine est constituée de deux domaines, l'un de forme allongée, l'autre quasi-sphérique, reliés entre eux par un court linker. Le nombre de pseudo-résidus situés dans chacun des domaines et dans le linker est le même pour les 20 enveloppes :  $75 \pm 3$  pour le domaine sphérique, 5 ou 6 pour le linker,  $93 \pm 3$  pour le domaine allongé. Le rayon de giration calculé par CRY SOL sur les pseudo-résidus du domaine quasi-sphérique donne une valeur de  $13\text{\AA}$ , très proche de la valeur estimée  $3(n)^{1/3} = 12.7\text{\AA}$  pour un globule de 75 résidus.



**Figure 46. Analyse structurale par SAXS de la protéine pb10.** (a) En bleu : courbe SAXS expérimentale de pb10. En rouge : courbe ayant permis de calculer l'enveloppe moléculaire présentée en (c). (b) Fonction de distribution des distances interatomiques de pb10. (c) Modèle à basse résolution d'enveloppe moléculaire de pb10 calculé à l'aide du logiciel GASBOR.

### **3-1-3 Analyse des résultats de SAXS à la lumière de données bio-informatiques**

Ces résultats sont à rapprocher de ce qu'on sait sur la protéine en termes de séquence : à partir du résidu 80, la séquence de la protéine est homologue à un domaine immunoglobuline. La forme déduite de la courbe de SAXS laisse donc supposer que la protéine comporte un domaine N-terminal globulaire, puis un linker court au niveau des résidus 75-80, puis un domaine C-terminal allongé de type immunoglobuline. Il est intéressant de remarquer que la plupart des prédictors de désordre ne donnent pas de déstructuration dans la protéine sauf dans une petite région (moins de 20 résidus) centrée sur les résidus 75-80.

#### **Détermination du domaine de fixation à la capsid**

Il est à noter que les domaines immunoglobulines sont très présents chez les protéines des phages, aussi bien dans les protéines de capsid que de queue (Fraser, Yu et al. 2006). De nombreuses études ont toutes montrées que ces domaines étaient exposés à la surface des phages (Fokine, Chipman et al. 2004) (Morais, Choi et al. 2005) (Katsura 1981). Il est donc vraisemblable que le domaine immunoglobuline de pb10 soit le domaine exposé à la surface, le domaine N-terminal serait donc le domaine de fixation.

L'alignement de séquence avec les pb10 des phages SPC35, EPS7 et H8, phages apparentés à T5, montre que le domaine N-terminal est très bien conservé, au contraire du domaine C-terminal (figure 47). Les structures des protéines majeures de capsid étant très bien conservées, il s'agit d'un autre argument confirmant que le site de fixation se situe sur le domaine N-terminal de pb10.

```

T5 |AAU05286.1| MIDYSGLRITFGEKLPESHIFFFATVAAHKYVPSYAFRLRRELGLSSAHTNR 50
H8 |AER23811.1| MIDYSGLRITFGEKLPESHIFFFATVAAHKYVPSYAFRLRRELGLSSAHTNR 50
ESPS7 |YP_001837085. MIDYDGLKAIFGEKLPESHIFFFATVAAHKFVFNATLRKEFGMTTAHTNR 50
SPC35 |YP_004306621. MIDYSGLRITFGEKLPESHIFFFATVAAHKYVPSYAFRLRRELGLSSAHTNR 50
*** . ** : ***** : ** * ** : * : *****

T5 |AAU05286.1| KVVKKFVEAYGKAIPPAPAPPLTSLKDLTASMSVEEGAALTLVATGG 100
H8 |AER23811.1| KVVKKFVEAYGKAIPPAP---PLTSLKDLTASMSVDTGAALTLVATGG 97
ESPS7 |YP_001837085. KVVNKFKEVYEVVAPVVP---ALAFKLNLAKTAVDTGAAINLGVTVGG 97
SPC35 |YP_004306621. KVVKKFVEAYGKAGPADVVPAPTLQVTPGENSVTYT---ITDGTNTGSP 97
***:** * * : * . : : . . . . .

T5 |AAU05286.1| TGPYTYAWTKDGSPIPDASGATYTKPTAAAEADAG-SYKVTVTDSKQVSKD 149
H8 |AER23811.1| TGPYTYAWTKNGSPIPGASGATYTKPTAAAEADAG-SYKVTVTDSKRVSKD 146
ESPS7 |YP_001837085. TAPYTYAWTKDGSPLSGVTGPNFNKATAAAEADAG-TYKVVVTVDSKSTSIT 146
SPC35 |YP_004306621. ITQYKVYYTDGTTTPKTHNSGSSKTGTISTLEAKYTFQSVAVNGAGESPK 147
* . : * . : * . : * . : : * : : : . . . *

T5 |AAU05286.1| STTCAVT-VNPTVPGG 164
H8 |AER23811.1| STACTVT-VNLGVGG- 160
ESPS7 |YP_001837085. SNECVTT-INPAPEA- 160
SPC35 |YP_004306621. STAVKATPTAPGVGG- 162
* . . * .

```

**Figure 47. Conservation des résidus du domaine N-terminal des protéines de décoration chez les phages apparentés à T5.** Alignement de séquences réalisées avec clustalW. De haut en bas les pb10 des phages : T5, H8, EPS7 et SPC35. La partie encadrée en noir correspond au linker putatif.

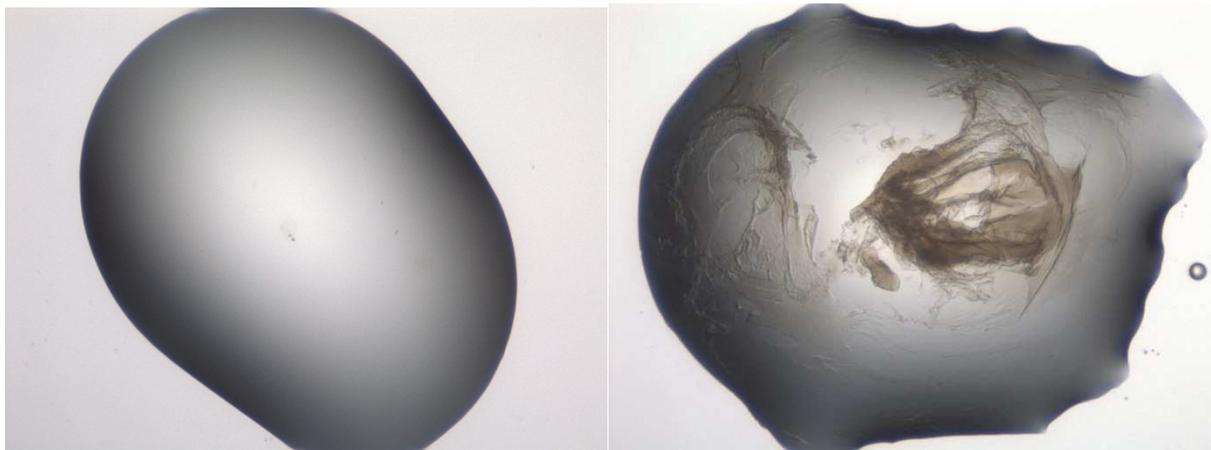
## Modélisation structurale

A part chez les pb10 des phages apparentés à T5, on ne trouve aucune séquence homologue au domaine N-terminal. En revanche le domaine C-terminal a de forte homologie avec de très nombreux domaines immunoglobuline, et notamment avec le domaine 1 de la HOC du phage RB49. L'utilisation du serveur HHpred (Soding, Biegert et al. 2005) (figure 48 (a)) nous indique avec une probabilité de 99.5% qu'ils adoptent le même repliement. A l'aide du serveur Modeler (Eswar, Webb et al. 2007), nous avons donc modélisé par homologie de séquence la structure de domaine C-terminal de pb10 en utilisant pour patron la structure cristallographique de HOC (figure 48 (b)).



pour tous les essais de cristallogenèse : 50mM Hepes pH 7.2 et 150 mM NaCl) et 100 nL de solution de cristallisation. Les trois concentrations protéiques utilisées étaient : 2,6, 5,3 et 10,6 g/L. Aucun cristal n'a été observé lors de ces premiers essais de cristallogenèse. La quasi-totalité des gouttes étaient claires (figure 49) ce qui indique que la concentration en protéine était insuffisante pour atteindre la zone de sursaturation nécessaire à l'obtention de cristaux. Les essais suivants ont été effectués avec une concentration en pb10 de 21g/l, puis de 32 g/L (pour le dernier essai). Les kits Classic, Sulfate d'ammonium, Protein Complex, MPD, JCSG, PEG et PEG II ont été utilisés, les gouttes étant déposées par le robot Mosquito, ce qui permet de travailler avec un volume de protéine réduit (20uL par kit). La présence de gouttes précipitées indique que la concentration protéique est adéquate, mais nous n'avons pas observé de cristaux.

Deux explications peuvent être avancées à cette absence de cristallisation de pb10 : le linker entre les deux domaines pourrait permettre la rotation respective des deux domaines par rapport à son axe, ou bien le domaine de fixation est replié en *molten globule* et ne se structure que pendant la fixation à la capsid.



**Figure 49. Cristallogenèse de pb10.** Exemple de gouttes couchées : claire (à gauche) et précipitée (à droite), visualisées par le robot de visualisation Rock Imager 182.

### **III- Conclusion et perspectives**

#### **1- Récapitulatif des résultats obtenus**

La protéine pb10 a été produite et purifiée en quantité suffisante pour effectuer des études structurales. Les analyses en MALS et en spectrométrie de masse Q-TOF indiquent que la protéine purifiée est monomérique en solution et ont permis de vérifier son intégrité.

L'étude fonctionnelle in vitro a permis de prouver que pb10 ne se fixe que sur les capsides gonflées et est donc un marqueur de l'expansion. Cette étude a également permis de mettre en évidence la coopérativité au sein d'une capside de la fixation de pb10, ainsi qu'une forte affinité de cette fixation. L'étude par calorimétrie différentielle de la fixation n'a en revanche montré aucun effet stabilisateur de cette fixation sur la capside ; on ne peut cependant pas exclure un éventuel effet stabilisateur vis-à-vis de la présence de détergent ou d'un pH fortement acide ou basique, ou un effet protecteur contre la protéolyse des protéines majeures de capside.

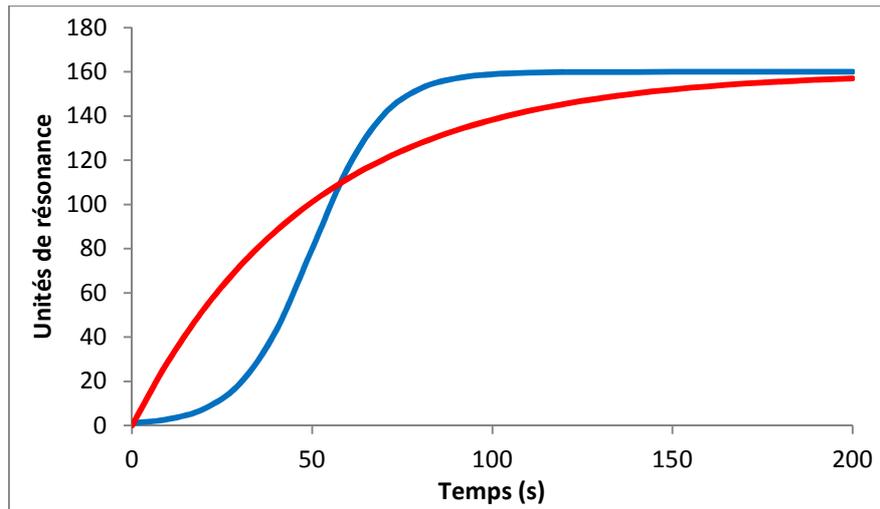
L'analyse par SAXS montre que la protéine est allongée et est constituée de deux domaines séparés par un linker de 5 ou 6 acides aminés. Le modèle à basse résolution d'enveloppe moléculaire de pb10 calculé à partir des données SAXS montre également deux domaines : le premier globulaire, le deuxième de forme allongée. L'analyse de la séquence de pb10 montre que les 90 résidus C-terminaux ont une forte homologie à un domaine immunoglobuline. Le domaine allongé correspondrait au domaine C-terminal et le domaine globulaire au domaine N-terminal. Le domaine C-terminal est vraisemblablement exposé à la surface de la capside, le site de fixation serait donc situé dans le domaine N-terminal. Ces données sont cohérentes avec le fait que l'étiquette poly-histidine C-terminale de pb10

n'empêche pas sa fixation sur la capsid. Nous avons également déterminé par homologie avec la séquence de HOC du phage RB49 un modèle du domaine C-terminal. En revanche l'étude n'a pas permis la cristallisation de pb10.

## **2- Fixation de pb10 sur la capsid**

L'analyse de la mobilité électrophorétique des capsides décorées sur gel d'agarose suggère fortement une fixation coopérative de pb10 sur la capsid. Cette coopérativité impliquerait un réarrangement conformationnel de la capsid après la fixation d'une ou de quelques molécules de pb10. A ce jour les changements de conformation des sous unités des protéines de capsid induits par une protéine de décoration n'ont jamais été décrits. Si la résolution le permet, ce changement de conformation pourra être mis en évidence par cryo-microscopie électronique, en comparant les reconstructions des capsides vides expansées avant et après décoration.

Nous prévoyons de mesurer l'affinité de pb10 pour les capsides et de mettre en évidence la coopérativité de la fixation par Résonance Plasmonique de Surface (SPR). En effet cette méthode permet de mesurer la quantité totale de pb10 fixée sur les capsides en fonction du temps. Ainsi une courbe exponentielle est typique d'une fixation non coopérative (figure 50), alors qu'une courbe sigmoïdale traduit une fixation coopérative. En outre, cette technique permettra d'avoir accès à la constante d'équilibre de la fixation de pb10. En effet l'étude par migration sur gel d'agarose de la fixation de pb10 ne permet pas de déterminer la valeur de cette constante, mais permet en revanche de déterminer une borne supérieure, qui est de 100 nM.



**Figure 50. Signaux théoriques de SPR.** En rouge : signal de type exponentiel, typique d'une fixation non coopérative. En bleu : signal de type sigmoïdal, typique d'une fixation coopérative.

### 3- Données structurales

L'absence de cristallisation de pb10 est à mettre en relation avec les échecs de cristallisation de toutes les autres protéines de décoration de phages. Pour les protéines HOC du phage T4 et gp8.5 du phage  $\phi$ 29 seuls les domaines exposés au solvant ont été cristallisés ; les structures des domaines de fixation ne sont donc pas connues. Deux explications sont possibles : la première est que le linker entre les deux domaines pourrait être flexible, la deuxième est que le domaine N-terminal n'est pas structuré en solution et adopte un repliement stable lors de sa fixation à la capsid. Ce comportement, qui a été observé pour l'extrémité N-terminale de la protéine D du phage  $\lambda$  (Lander, Evilevitch et al. 2008), pourrait être général pour plusieurs types de protéines de décoration. Les parties du gène de pb10 correspondant aux deux domaines identifiés par SAXS viennent d'être clonées et surexprimées au laboratoire afin d'analyser leur propriétés. Le domaine immunoglobuline est fortement produit et soluble, alors que le domaine de fixation isolé est très faiblement produit et difficilement purifiable, ce qui renforce l'idée qu'il n'est pas structuré quand il n'est pas

fixé. L'ensemble de ces résultats nous conduiront à étudier la structure de pb10 par RMN. En amont de la détermination de la structure, nous pourrons évaluer si certains résidus de la protéine correspondent à une région non structurée. Par ailleurs, en comparant les spectres HSQC de la protéine entière et du fragment Immunoglobuline enrichis en  $^{15}\text{N}$ , nous pourrons également déterminer si les deux domaines de pb10 sont indépendants, ou si au contraire il existe une région d'interaction entre les deux. Si les deux domaines sont indépendants, il sera alors judicieux de tenter de déterminer la structure des deux domaines séparés, à condition de pouvoir optimiser la production du domaine N-terminal. Notons que les structures de nombreux domaines immunoglobulines ont été résolues par cristallographie et que la faible taille du domaine C-terminal de pb10 (90 acides aminés, ~10 kDa) laisse envisager la cristallisation de ce domaine en vue de la détermination de sa structure.

# IV- MATERIEL ET METHODES

## 1- Clonage, production et purification de pb10

### 1-1 Clonage

La séquence codante de pb10 (GenBank GI:51512087) a été amplifiée par PCR et clonée par les techniques classiques de restriction/ligation (NcoI-XhoI) dans le plasmide d'expression pET28<sup>b</sup> (Novagen- Résistance Kanamycine), en phase avec une séquence 6His en position C-terminale.

### 1-2 Surexpression en système bactérien

Le plasmide de surexpression est introduit dans des bactéries compétentes de la souche *E. coli* BL21. Les bactéries transformées sont sélectionnées grâce à la résistance à la kanamycine portée par le vecteur. Les cellules ainsi obtenues sont cultivées une nuit à 37°C dans 30 mL de milieu LB complété par de la kanamycine (500mU). Cette préculture est transférée dans 170 mL de milieu 2YT avec 500mU de kanamycine. Arrivée à saturation cette préculture est ensuite versée dans un bioréacteur contenant 4,8 L de milieu 2YT. La température est de 37°C, l'agitation 800rpm, le débit d'air 5L/minutes. La croissance des bactéries est contrôlée par mesure de la densité optique à 600 nm. Lorsque l'absorbance à 600 nm atteint 1, l'expression de la protéine est induite grâce à l'addition d'IPTG 0,5mM. Après 3h d'induction, les bactéries sont centrifugées 30 min à 5000g. Le culot contenant les cellules

est repris dans 200 mL de tampon (50mM Hepes pH7,2 ; 500mM NaCl ; 20mM Imidazole ; 2mM  $\beta$ -mercaptoéthanol), et stocké à -20°C jusqu'à la purification.

### **1-3 Préparation des extraits cellulaires**

Après décongélation, les cellules sont lysées par Presse de French à la pression de 1200psi. Le lysat est centrifugé pendant 30 min à 40 000g. Le surnageant, composé de la fraction cytoplasmique soluble contenant pb10, est récupéré pour la suite de la purification. Ces étapes sont effectuées à 4°C.

### **1-4 Purification des protéines par chromatographie**

La purification de pb10 nécessite les 3 étapes suivantes : Chromatographie d'affinité sur un métal immobilisé (IMAC), Chromatographie d'échange de cation et Chromatographie d'exclusion de taille (SEC). Toutes ces chromatographies sont réalisées sur un système automatisé FPLC Purifier ou Xpress (GE Healthcare).

#### **1-4-1 Chromatographie d'affinité sur un métal immobilisé (IMAC)**

La protéine pb10 porte une étiquette hexa-histidine à son extrémité C-terminale, ce qui permet une première étape de purification sur une colonne d'affinité, colonne composée d'une résine greffée avec des ions  $Ni_2^+$ . Après injection de la fraction cytoplasmique, la colonne est lavée avec du tampon 50 mM Hepes pH 7.2, 1M NaCl pour éliminer l'ADN fixé non spécifiquement sur la colonne et sur pb10. L'élution se fait par gradient d'imidazole (Solution A : 50 mM Hepes pH7.2, 50 mM NaCl, 20mM imidazole, solution B : 50 mM Hepes pH7.2, 50 mM NaCl, 1M imidazole).

### **1-4-2 Chromatographie d'échange de cation**

La chromatographie d'échange de cation sépare les protéines selon leur charge par passage à travers une résine chargée électriquement. Le point isoélectrique de pb10 avec son étiquette hexa-histidine est de 7.9, ce qui permet une chromatographie d'échange de cation. L'éluion s'effectue par gradient 0-400mM NaCl, dans du tampon 50mM Hepes pH 7.2.

### **1-4-3 Chromatographie d'exclusion de taille (SEC)**

La chromatographie d'exclusion sépare les protéines selon leur taille par passage à travers un tamis moléculaire de billes poreuses. La solution issue de l'étape chromatographique précédente, contenant la protéine pb10, est concentrée par ultrafiltration à 5500g sur membrane de porosité limitée à des molécules de 3kDa (Millipore, Amicon Ultra), jusqu'à atteindre un volume de 5mL. Elle est alors injectée sur le tamis moléculaire (S75 16/60 HL, GE Healthcare), à un débit de 1ml/min. Des fractions de 1,5mL sont collectées en sortie de colonne. Un détecteur d'UV mesure l'absorbance à 280nm en sortie de colonne, ce qui permet d'identifier les fractions où la protéine est présente.

### **1-4-4 Concentration et contrôle de la qualité de l'échantillon protéique purifié**

La protéine pb10 est ensuite concentrée à la concentration voulue par ultrafiltration sur membrane. La concentration de la protéine est déterminée par mesure de l'absorbance à 280nm. Pour cela, on utilise les valeurs théoriques du coefficient d'extinction molaire calculées à partir de la séquence protéique ( $\epsilon = 24410 \text{ M}^{-1} \text{ cm}^{-1}$ ) et du poids moléculaire ( $M = 18474.8 \text{ g/mol}$ ). De plus, l'examen du spectre d'absorbance entre 240 et 340 nm, et notamment l'absorbance à 260nm, permet de vérifier l'absence de contamination par des acides nucléiques.

## **1-5 Contrôle de l'état d'oligomérisation et de l'intégrité de la protéine pb10 purifiée**

### **1-5-1 SEC-MALS**

100 µL de solution de pb10 (concentration : 5.2 mg/ml, tampon : 50 mM HEPES pH7.2, 150 mM NaCl) ont été injectée à un débit constant de 0.5 ml/min à l'aide du système GPCmax chromatographic (Malvern) dans une colonne Superdex<sup>TM</sup> 200 HR 10/30 (GE Healthcare) préalablement équilibrée avec le tampon de la protéine. L'élution a été suivie à l'aide du détecteur TDA302 (Viscotek), comprenant un refractomètre différentiel, deux détecteurs de diffusion de la lumière situé à des angles de 7° et 90° et d'un viscomètre à pression différentielle. Les données ont été collectées et traitées avec le programme OmniSEC (Viscotek).

### **1-5-2 Spectrométrie de masse (Q-TOF)**

Les mesures de la masse moléculaire de pb10 ont été effectuées avec un spectromètre quadripôle à électrospray (Q/TOF *Premier*, Waters) équipé d'un appareil Nanomate (Advion). 0.5 µL de solution de pb10 (concentration : 5 µM) ont été mélangé avec un volume équivalent de solution d'acétonitrile (concentration finale : 50% acétonitrile, 1% acide formique). Le débit était de 100 à 500 nl/min. L'énergie de collision était fixée à 6 eV, le gaz de collision utilisé étant l'argon. Les données ont été collectées et traitées avec le logiciel The Mass Lynx 4.1. La déconvolution des données a été effectuée en appliquant l'algorithme MaxEnt1. L'incertitude des mesures est de  $\pm 2$  Da.

## **2- Etudes Fonctionnelles et Structurales**

### **2-1 Etude par ultracentrifugation de la fixation de pb10 sur les capsides vides d'ADN**

20  $\mu$ L de capsides compactes et gonflées (concentration : 2,4 mg/ml, tampon 20 mM TRIS pH7.6, 200mM NaCl) ont été incubées 1h avec la 10  $\mu$ L de solution de protéines pb10 (ratio de 2 protéines pb10 par hexamère, tampon 50 mM HEPES pH7.2, 150mM NaCl), puis ultra-centrifugées à 70 000 g pendant ½ heure à 4°C. Les culots ont été resuspendus dans 30  $\mu$ L de tampon 20mM TRIS pH 7.6. Les surnageants et les culots ont été déposés sur gel SDS-PAGE 4-20% biorad.

### **2-2 Etude par migration sur gel d'agarose natif**

20  $\mu$ L de capsides gonflées (concentration : 2 mg/ml, tampon 20 mM TRIS pH7.6, 200mM NaCl) ont été incubées 1h avec 2.6, 3.0, 3.4, 3.8, 4.2, 4.6 et 19  $\mu$ L de solution de protéines pb10 (ratio de 0.7, 0.8, 0.9, 1, 1.1, 1.2 et 5 protéines pb10 par hexamère) et complété à 40  $\mu$ L avec du tampon 50 mM HEPES pH7.2, 150 mM NaCl.

Les gels d'agarose natifs 1% sont coulés avec une solution de Tris Acétate 80mM pH 8.4, également utilisée comme tampon de migration. Les échantillons sont chargés avec 10% de glycérol et 0.005% de bleu de bromophénol avant la migration qui s'effectue en trois heures à 50V. Les gels sont révélés par immersion dans du bleu de coomassie Biosafe (BIORAD) pendant quelques heures.

## 2-3 Etude par calorimétrie différentielle à balayage

L'étude de l'effet de la fixation de pb10 sur la stabilité des capsides par calorimétrie différentielle à balayage (DSC) a été menée en utilisant le calorimètre MicroCal modèle VP-DSC. Le balayage a été effectué à la vitesse de 1K par minute, de 20 à 100°C. Les capsides gonflées ainsi que pb10 ont été dialysées dans un tampon 50 mM Phosphate et 200 mM NaCl. Toutes les solutions ont été dégazées avant l'enregistrement de spectres. Toutes les mesures ont été précédées d'une mesure de la ligne de base avec le tampon de dialyse. La capacité calorifique du solvant a été soustraite des mesures des échantillons avant analyse.

## 2-4 Etude structurale de pb10 par diffusion des rayons X aux petits angles

Les expériences de SAXS ont été menées en utilisant un appareil à anode tournante (Cu Ka, longueur d'onde :  $\lambda = 1.54 \text{ \AA}$ ), le Nanostar (Bruker, Karlsruhe, Germany). Les rayons X diffusés ont été collectés à l'aide d'un détecteur 2D (Vantec) situé à 662 mm de l'échantillon. La gamme angulaire était de  $0.011 < q < 0.40 \text{ \AA}^{-1}$ , avec  $q = 4\pi \sin\theta/\lambda$ ,  $2\theta$  étant l'angle de diffusion. Le temps d'acquisition d'une courbe était de 1 heure, et pour chaque échantillon comme pour chaque tampon 10 courbes ont été enregistrées. La moyenne de ces courbes ainsi que la soustraction du tampon correspondant ont été effectuées à l'aide du logiciel PRIMUS (Konarev, Volkov et al. 2003).

Le tampon utilisé lors de ces expériences était : 50mM HEPES pH7.6, 150mM NaCl et les trois concentrations protéiques : 2,5, 4,8 et 9,6 g/L. Les trois courbes étaient parfaitement superposables excepté aux très petits angles, cela indique que les interactions entre protéines sont très faibles dans ces conditions de tampon et de concentration. La courbe

utilisée pour cette analyse contient les données à 2.5 g/L aux petits angles ainsi que les données à 9.6 g/l aux grands angles.

## **2-5 Essais de cristallogenèse de la protéine pb10**

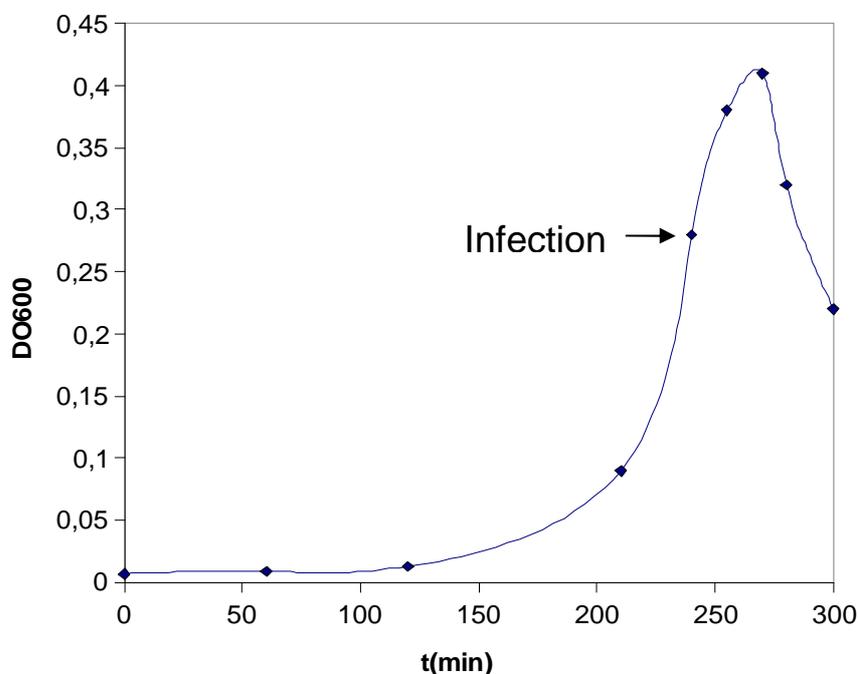
La cristallisation des protéines a été réalisée sur la plateforme de cristallogenèse de l'IBBMC (Quevillon-Cheruel, Liger et al. 2004; Leulliot, Tresaugues et al. 2005). Les robots de pipetage (Cartesian Nanodrop, Tecan Geneis, Mosquito) permettent de tester un grand nombre de conditions de manière automatisée, dans un temps réduit, en utilisant un volume de solution protéique faible (20 à 30  $\mu$ L pour une plaque 96 puits).

La recherche des conditions de cristallogenèse initiales repose sur le principe de la diffusion de vapeur, en utilisant des kits commerciaux (kits Qiagen : Classics, Classics Lite, Nucleix, ProComplex, MBI, MBII, Anions, Cations, MPD, AmSO<sub>4</sub>, PEGs suite, JCSG+). Les essais sont effectués en utilisant des plaques 96 puits ayant trois supports de gouttes assises par puits. Les 96 réservoirs sont remplis avec 150  $\mu$ L de solutions provenant des kits commerciaux par le robot de pipetage Tecan. Puis, les robots de pipetage de type Cartesian ou Mosquito déposent 0,1 $\mu$ L de ces mêmes solutions dans les trois supports de goutte assise, et 0,1 $\mu$ L de la solution protéique. Les trois supports permettent notamment de tester trois concentrations croissantes de protéine pour une même condition de cristallisation. Les plaques sont ensuite scellées et stockées à 18°C. L'aspect des gouttes et son évolution sont observés quotidiennement *via* le robot de stockage et de visualisation (Rock Imager 182, Formulatrix) pendant 18 jours.

### 3- Production, purification et expansion des capsides

#### 3-1 Culture des bactéries et production des capsides

Les capsides sont produites par infection d'une culture de bactéries *E.coli* F (Lanni 1958) à 37°C dans le milieu riche LB complété avec 1mM MgCl<sub>2</sub> et 1mM CaCl<sub>2</sub>. La croissance bactérienne est suivie par mesure de l'absorbance à 600nm. Quand celle-ci atteint la valeur de 0.3, on ajoute une suspension de bactériophage T5 stAmN5 en quantité nécessaire pour avoir une multiplicité d'infection de 5 phages par bactérie, ce qui assure l'infection de 99% des bactéries. L'infection des bactéries par ce mutant conduit à la production de capsides vides et de queues qui ne peuvent s'assembler les unes aux autres. L'incubation se poursuit jusqu'à ce que l'absorbance décroisse jusqu'à 0.1-0.2, indiquant une lyse de la totalité des bactéries (Figure 51).



**Figure 51: Cinétique de la lyse bactérienne après infection avec le phage T5 stAMN5.** La lyse est provoquée à absorbance ~0.3 à une multiplicité de 5 phages pour une bactérie, la population bactérienne est estimée à l'aide de la mesure de l'absorbance à 600nm.

### **3-2 Purification des capsides vides**

Le lysat obtenu est centrifugé à 5000g pendant 30 minutes pour éliminer les débris bactériens les plus denses, et le surnageant est précipité par addition de 7.5% PEG 6000 et 0.5 mol/l NaCl (4°C pendant 12h). Il est ensuite sédimenté par centrifugation à 5000g pendant 30 minutes, puis le culot est re-suspendu dans 2 mL de tampon 20 mM Tris HCl pH 8, 200mM NaCl. Les composants ainsi concentrés, (capsides vides, queues et autres assemblages macromoléculaires issus de la lyse) sont séparés par ultracentrifugation zonale sur gradient discontinu de glycérol (10, 20, 30 et 40 % (p/v)) à 50000g pendant 2 heures. De haut en bas du gradient 12 fractions de 1 mL sont récupérées et analysées en gel natif d'agarose 1%. Les fractions contenant les capsides sont rassemblées et purifiées par deux passages de chromatographie d'échange d'ions sur colonne HitrapQ (GE Healthcare). Les fractions contenant les capsides sont ensuite dialysées dans le tampon le plus adapté aux techniques utilisées, mais pour conserver les capsides compactes plusieurs jours, une concentration de 500 mM de NaCl est nécessaire. Les capsides sont ensuite concentrées par ultrafiltration à 5000g sur membrane de porosité limitée à des molécules de 30kDa (Vivaspin). Une concentration de 1 M de NaCl prévient l'expansion des capsides lors de cette étape.

### **3-3 Expansion des capsides in vitro**

L'expansion des capsides in vitro s'effectue par dialyse pendant 48 heures contre un tampon ne contenant pas de NaCl. L'état des capsides (compactes ou gonflées) est ensuite contrôlé par diffusion dynamique de la lumière (DLS) ou par cryo-microscopie électronique.

### **3-4 Normalisation de la concentration en capside**

A cause de la diffusion liée à la taille de la capsid, nous ne pouvons pas utiliser directement l'absorbance à 280 nm mesurée  $A_{280}^{mes}$  pour déterminer la concentration de l'échantillon de capsid. Le rapport théorique  $r_{th}$  entre l'absorbance à 280 nm et l'absorbance à 260 nm de la capsid est calculé à l'aide de la séquence de pb8. L'absorbance à 280 nm utilisée pour le calcul de la concentration vaut :

$$A_{280} = r_{th} * A_{280}^{mes} * \left( \frac{1}{r_{mes}} - \frac{b}{c} \right) * \frac{1}{1 - \frac{b * r_{th}}{c}}$$

avec  $b = 280^4$ ,  $c = 260^4$  et  $r_{mes} = \frac{A_{280}^{mes}}{A_{260}^{mes}}$ . Le raisonnement conduisant à cette équation est explicité en annexe 1.



## **Conclusion générale**

Le travail réalisé au cours de ma thèse a permis d'apporter de nouveaux éléments de compréhension à la maturation de la procapside du bactériophage T5, notamment aux étapes d'expansion et de décoration.

La première partie de ma thèse concerne l'étude de l'expansion *in vitro* de la capside de T5. Nous avons pu purifier la procapside de T5 et caractériser sa taille (720 Å de diamètre) et sa forme par SAXS et par cryo-microscopie électronique. J'ai pu ensuite déterminer des conditions physico-chimiques permettant son expansion *in vitro*, c'est-à-dire son passage de l'état procapside à celui d'une capside ayant la forme et la taille de la capside du virion infectieux. Une analyse par calorimétrie différentielle à balayage (DSC) des états compact et expansé de la capside a montré que l'expansion permettait une très importante stabilisation de la capside, stabilisation qui contrairement à celle observée lors de l'expansion de la capside du phage HK97, est obtenue sans pontage covalent entre les capsomères, par un « simple » réarrangement structural des protéines majeures de capside.

J'ai étudié la cinétique de ce processus par diffusion dynamique de la lumière, permettant de déterminer l'ordre de grandeur du temps caractéristique de cette étape. Une étude par SAXS couplée avec un stopped-flow a permis de montrer la grande sensibilité de ce temps caractéristique vis-à-vis du pH auquel sont acidifiées les capsides. De plus, l'analyse par décomposition en valeurs propres des séries de courbes SAXS au cours du temps a montré que l'étape d'expansion est hautement coopérative et qu'il n'y a pas d'état intermédiaire dont la forme ou la taille diffère significativement des états compact et gonflé. Cependant, l'étude de l'évolution des populations de capsides compactes et gonflées au cours du temps suggère la présence d'un état intermédiaire qui ne peut pas être distingué par SAXS, mais de nouvelles études sont nécessaires pour pouvoir confirmer cette assertion.

J'ai également tenté de déterminer, par cristallographie, la structure atomique de la forme expansée de la capsidie vide. J'ai obtenu des cristaux de capsidie dont la diffraction, d'une résolution trop faible ( $>20 \text{ \AA}$ ), ne permet de résoudre la structure. L'étude sera poursuivie en essayant de produire et de faire diffracter les capsides pleines d'ADN. Cette approche cristallographique est effectuée en parallèle à la reconstruction, à une résolution pseudo-atomique, de la structure des capsides compacte et expansée, à l'aide d'images de cryo-microscopie électronique, reconstruction effectuée par Alexis Huet (University of Pittsburgh, données pas encore publiées).

La deuxième partie de ma thèse est centrée sur l'étude de la décoration de la capsidie de T5. L'étude fonctionnelle *in vitro* a permis de prouver que la protéine de décoration, pb10, ne se fixe que sur les capsides expansées et est donc un marqueur de l'expansion. Cette étude a également permis de mettre en évidence la coopérativité au sein d'une capsidie de la fixation de pb10, ainsi que la forte affinité de cette fixation. Cette fixation sera ultérieurement étudiée par SPR pour déterminer précisément la constante d'affinité de la fixation. Comme lors de la fixation des protéines HOC, gp12 et gp8.5 sur les capsides des phages T4, SPP1 et  $\phi 29$ , aucun effet stabilisateur de la fixation de pb10 sur la capsidie de T5 n'a pu être montré par DSC. Des études ultérieures devront être effectuées pour mettre en évidence un éventuel effet stabilisateur vis-à-vis de la présence de détergent ou d'un pH fortement acide ou basique, ou encore un effet protecteur contre la protéolyse des protéines majeures de capsidie. Par ailleurs, d'autres pistes sont à explorer pour comprendre le rôle de cette famille de protéines. Il faudrait évaluer leurs éventuelles propriétés de protéines d'« adhérence » évoquées dans la littérature (Fraser, Maxwell et al. 2007), mais non démontrées. Un mutant de T5 dépourvu de pb10, tout récemment obtenu, développe sur un tapis bactérien des « plages de lyse » très

petites par rapport à celle du phage sauvage, indiquant que pb10 favorise une étape de l'infection.

L'analyse structurale de pb10 par SAXS a permis la modélisation à basse résolution de l'enveloppe moléculaire de pb10 ; ce modèle mettant en évidence le fait que pb10 est constituée de deux domaines, le premier globulaire et le deuxième de forme allongée, séparés par un linker de 5 ou 6 acides aminés. L'analyse de la séquence de pb10 montre que les 90 résidus C-terminaux ont une forte homologie à un domaine immunoglobuline, ce domaine correspondrait au domaine allongé déterminé par le modèle moléculaire issu des données SAXS, le domaine N-terminal correspondrait donc au domaine globulaire. Le domaine C-terminal est vraisemblablement exposé à la surface de la capsid, le site de fixation serait donc situé dans le domaine N-terminal. Nous avons également déterminé par homologie avec la séquence de HOC du phage RB49 un modèle du domaine C-terminal. En revanche, les tentatives de cristallisation de pb10 ont échoué. La présence du linker entre les deux domaines et la structuration hypothétique de pb10 en *molten globule* sont deux explications possibles de cette absence de cristallisation, commune aux autres protéines de décoration HOC et gp8.5, dont le domaine de fixation à la capsid n'a pas pu être cristallisé. La détermination de la structure de pb10 pourra être envisagée par cristallisation des deux domaines séparément ou par RMN.

# **ANNEXES**

## Annexe 1 : Calcul de la concentration d'un échantillon de capside vide.

A cause de la diffusion liée à la taille de la capside, nous ne pouvons pas utiliser directement l'absorbance à 280 nm mesurée pour déterminer la concentration de l'échantillon de capside. Nous appelons  $A_{280}^{mes}$  et  $A_{260}^{mes}$  les absorbances mesurées respectivement à 280 nm et 260 nm,  $r_{mes} = \frac{A_{280}^{mes}}{A_{260}^{mes}}$  le rapport entre ces deux absorbances,  $A_{280}$  et  $A_{260}$  les absorbances propres à 280 et 260 nm des capsides, i.e. les absorbances des tryptophanes, tyrosines et des phénylalanines de la capside. On appelle  $r_{th}$  le rapport entre ces deux absorbances :

$$r_{th} = \frac{A_{280}}{A_{260}} \quad (A1-1)$$

Nous savons que la diffusion est inversement proportionnelle à la longueur d'onde à la puissance 4. On en déduit :

$$A_{280} + \frac{a}{b} = A_{280}^{mes} \quad (A1-2)$$

et

$$\frac{A_{280} + \frac{a}{b}}{A_{260} + \frac{a}{c}} = r_{mes} \quad (A1-3)$$

avec  $a$  un coefficient quelconque,  $b = 280^4$  et  $c = 260^4$

L'équation (A1-2) peut être réécrite sous la forme :

$$a = b(A_{280}^{mes} - A_{280}) \quad (A1-4)$$

En combinant (A1-2), (A1-3) et (A1-4) :

$$\frac{A_{280}^{mes}}{A_{260} + \frac{b}{c}(A_{280}^{mes} - A_{280})} = r_{mes}$$

Puis en utilisant (A1-1) dans cette équation :

$$\frac{A_{280}^{mes}}{\frac{A_{280}}{r_{th}} + \frac{b}{c}(A_{280}^{mes} - A_{280})} = r_{mes}$$

On en déduit :

$$\frac{A_{280}}{r_{th}} + \frac{b}{c}(A_{280}^{mes} - A_{280}) = \frac{A_{280}^{mes}}{r_{mes}}$$

Et donc :

$$A_{280} + \frac{b * r_{th}}{c} A_{280}^{mes} - \frac{b * r_{th}}{c} A_{280} = \frac{A_{280}^{mes} * r_{th}}{r_{mes}}$$

Ainsi :

$$A_{280}(1 -) = \frac{A_{280}^{mes} * r_{th}}{r_{mes}} - \frac{A_{280}^{mes} * b * r_{th}}{c}$$

On en conclut que l'absorbance des acides aminés de la capsidie vaut donc

$$A_{280} = r_{th} * A_{280}^{mes} * \left( \frac{1}{r_{mes}} - \frac{b}{c} \right) * \frac{1}{1 - \frac{b * r_{th}}{c}}$$

## Annexe 2 : La cristallogénèse des protéines

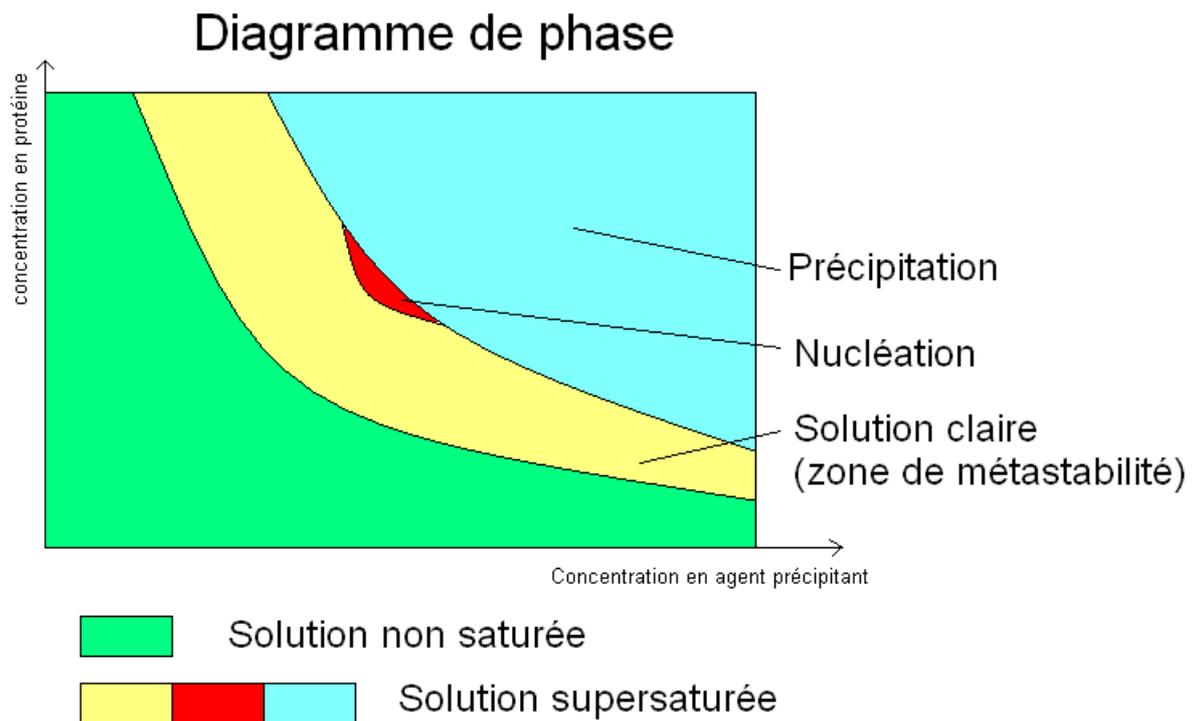
La cristallographie est une méthode couramment utilisée en biologie structurale pour déterminer la structure tridimensionnelle à une résolution atomique ( $\sim 2 \text{ \AA}$ ). Comme le SAXS elle repose sur l'interaction entre un faisceau de rayons X ayant une longueur d'onde de l'ordre de l'angström et les électrons composant l'objet étudié, mais contrairement à cette technique elle nécessite l'obtention de cristaux, via une phase appelée cristallogénèse.

### 1- Théorie

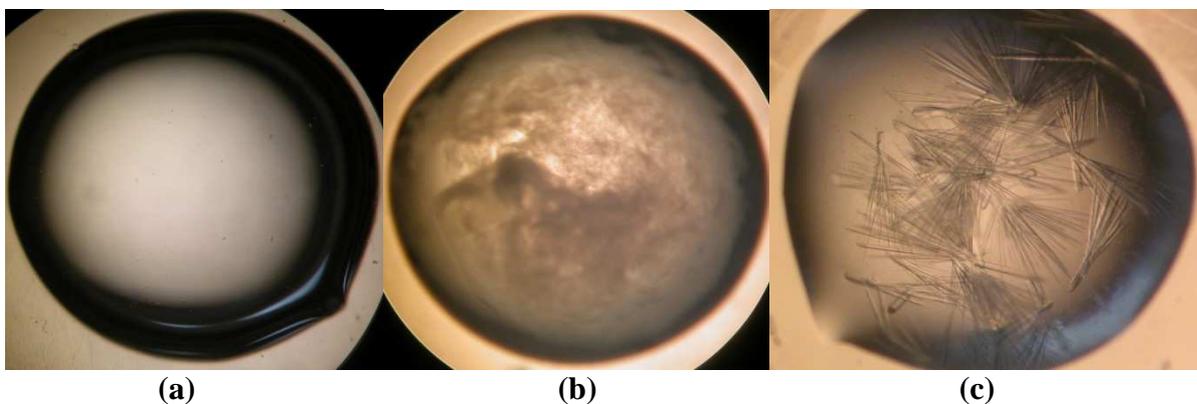
Le cristal est un état solide comportant des invariances par translation dans les trois directions de l'espace, consistant en un arrangement ordonné et répété d'un motif élémentaire. Le but de la cristallogénèse est de former des cristaux à partir d'une solution de protéine. Pour cela, on diminue la solubilité de la protéine, en stabilisant les interactions intermoléculaires. Les protéines peuvent alors se regrouper en noyaux appelés point de nucléation. Il n'est pas entropiquement favorable de former un point de nucléation, donc son apparition peut prendre du temps. A partir d'un point de nucléation, dans une solution dite sursaturée, c'est-à-dire contenant une quantité de protéine supérieure au seuil de saturation, le cristal va rapidement grandir, jusqu'à épuisement de la protéine soluble disponible.

La méthode pour réduire la solubilité de la protéine et ainsi atteindre l'état de sursaturation est l'ajout de sels et/ou de polymères, appelés agents précipitants. Les sels, par exemple, en captant l'eau contribuent à rapprocher les protéines. L'état physico-chimique d'une protéine est fonction de sa concentration et de celle des agents précipitants, et peut être visualisé en utilisant un diagramme appelé diagramme de phase (figure A1). A faible

concentration, le mélange protéine/précipitant est clair, tandis qu'à plus forte concentration la protéine est sous forme précipitée, c'est-à-dire dans un état solide amorphe et désordonné (figure A2). Dans une zone restreinte (zone de nucléation), des points de nucléation peuvent apparaître et conduire à la formation de cristaux.



**Figure A1. Diagramme de phase d'une protéine.** Ce diagramme est représenté en deux dimensions, mais il en possède de nombreuses autres si on considère tous les paramètres : les différents agents précipitants, le pH, la force ionique...



**Figure A2.** Exemples de phase d'une solution de protéine. (a) Une solution claire, située dans la zone du diagramme de phase dans laquelle la solution est non saturée ou est métastable. (b) Une solution située dans la zone de précipitation. (c) Exemple de cristaux de protéines.

## 2- Stratégie expérimentale

Il n'y a pas de prédiction théorique du diagramme des phases, le travail de recherche des conditions de cristallisation est donc expérimental. On peut faire varier :

- la concentration de la protéine
- le nombre d'agents précipitants
- les combinaisons d'agents précipitants
- les différentes concentrations en agent précipitant

Pour simplifier la recherche, on classe les agents précipitants en différentes catégories:

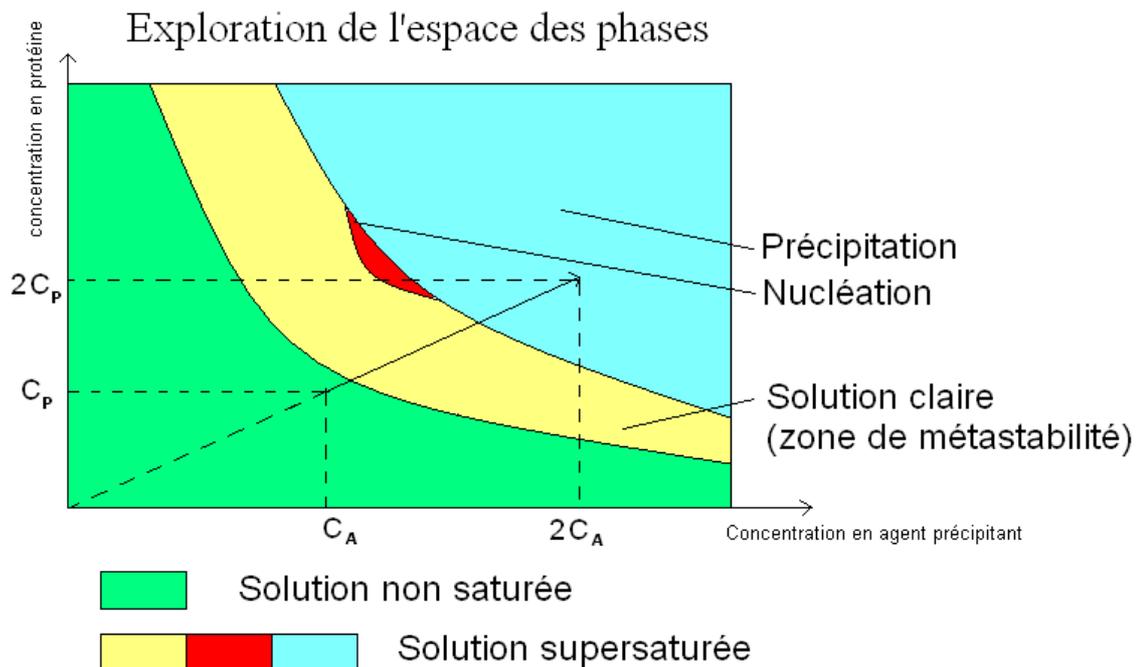
- les sels (force ionique)
- les précipitants, qui déterminent la limite solution claire/solution qui précipite
- les tampons, qui déterminent le pH de la solution
- les additifs, par exemple de l'alcool, des solvants ou des détergents, qui contribuent à la cristallisation selon différents mécanismes souvent mal connus (effet sur la solubilité, changement de la vitesse de diffusion des molécules...)

### 2-1 La méthode du criblage

La méthode qui vient en premier à l'esprit pour déterminer la zone de nucléation est un quadrillage systématique de l'espace des phases. Mais la zone de nucléation étant petite et le nombre de paramètres important, cette méthode consomme une quantité très importante de protéine et n'est donc pas utilisée. Des kits commerciaux contenant des mélanges prêts à l'emploi de précipitants, sels, tampons et additifs sont généralement utilisés pour leur

reproductibilité. Ces solutions ont été choisies pour explorer de manière efficace l'espace des paramètres.

## 2-2 La méthode de diffusion de la vapeur

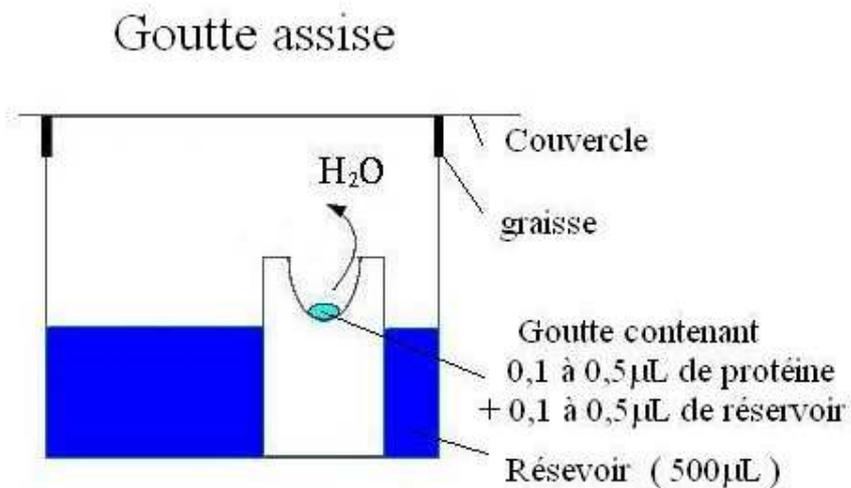


**Figure A3. Diagramme de phase montrant l'exploration de l'espace des phases par la méthode de diffusion de la vapeur.**

Afin de mieux explorer l'espace des phases, on utilise la méthode de diffusion de vapeur (figure A3). Il y a deux techniques qui utilisent la méthode de diffusion de la vapeur : la technique de la goutte suspendue et celle de la goutte assise.

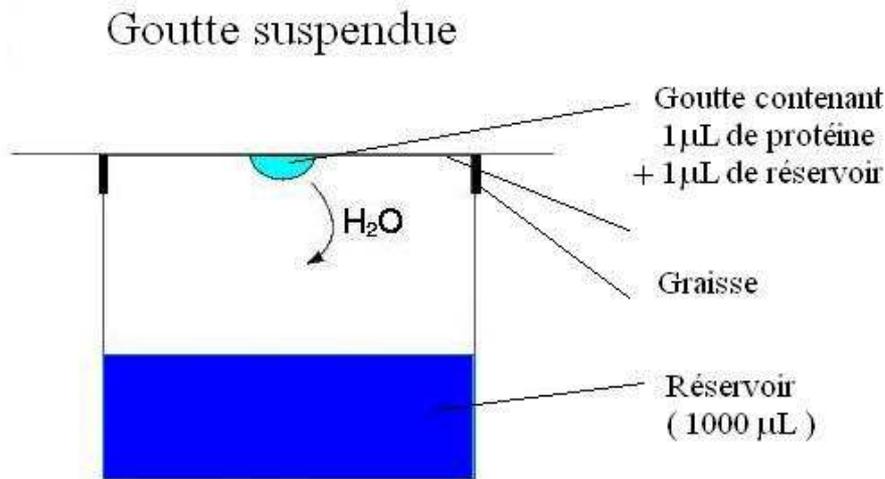
La préparation des boîtes qui utilisent la technique de la goutte assise (figure A4) est effectuée à partir d'un robot. Chaque condition ne nécessite que peu de solution de protéine

(entre 0,1 et 0,5  $\mu\text{L}$ ) et la robotisation permet d'effectuer de nombreux essais. Mais du fait de la faible quantité de protéine utilisée, les cristaux sont petits et rarement exploitables. Cette technique est donc essentiellement utilisée pour la recherche des conditions de cristallisation et assez peu pour l'obtention de cristaux qui sont capables de diffracter.



**Figure A4. Schéma d'une préparation utilisant la technique de la goutte assise.**

Une fois les conditions de cristallisation déterminées, la technique de la goutte suspendue (figure A5) est utilisée pour optimiser ces conditions et obtenir des cristaux capables de diffracter les rayons X. Cette technique est entièrement manuelle et les gouttes sont plus grosses, elle nécessite donc plus de temps et de protéine.

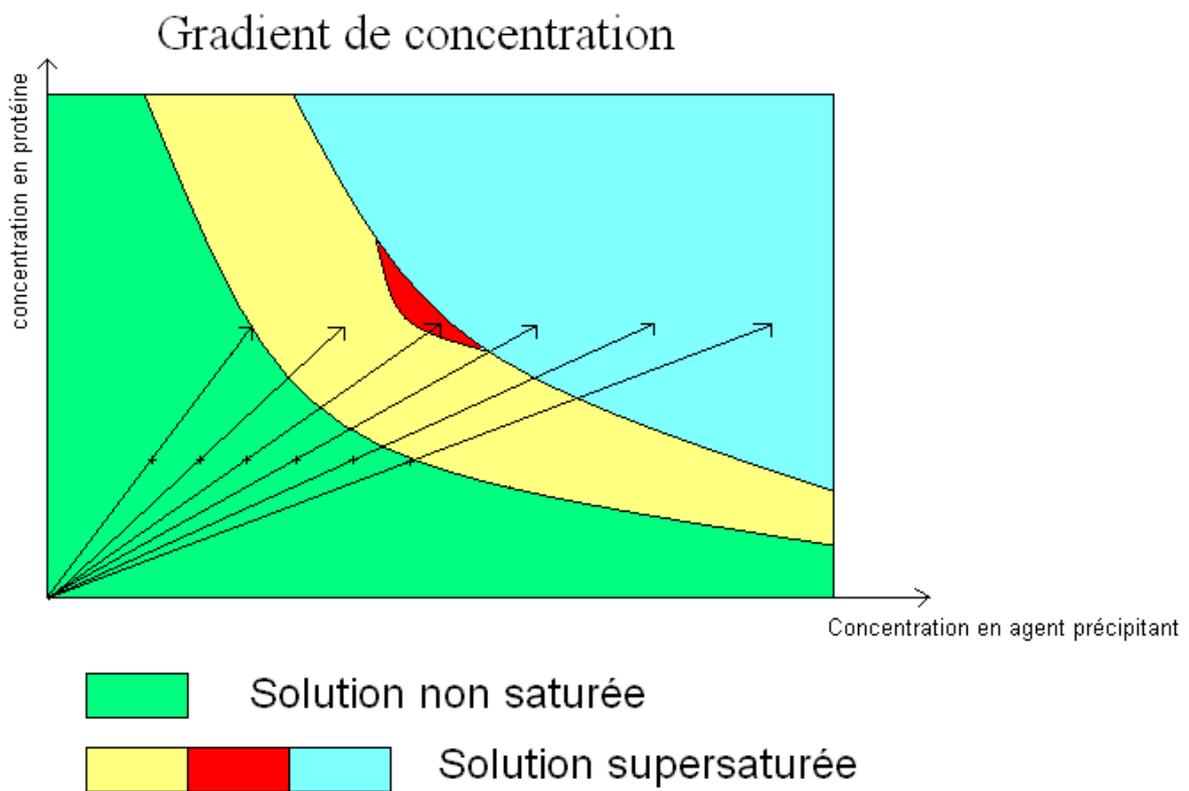


**Figure A5. Schéma d'une préparation utilisant la technique de la goutte suspendue.**

Dans les deux cas, la concentration initiale en agent précipitant dans la goutte est la moitié de la concentration dans le réservoir, car on a mélangé la même quantité de solution du réservoir avec une solution de protéine. Mais au cours du temps, l'évaporation de l'eau de la goutte augmente sa concentration. Il y a aussi évaporation de l'eau du puits, mais le volume de celui-ci n'évolue presque pas et on peut donc considérer sa concentration comme constante. Le système étant hermétiquement fermé, il n'y a donc pas d'échange avec l'extérieur. Au bout d'un certain temps, il y a équilibre entre la goutte et le réservoir : ils sont aux mêmes concentrations en agent précipitant. Le volume de la goutte a été divisé par deux, donc la concentration en agent précipitant et en protéine a été multipliée par deux. On a donc exploré l'espace des phases: si au cours de l'équilibration on passe dans la zone de nucléation, des cristaux de protéines peuvent se former.

## 2-3 L'optimisation des conditions de cristallogenèse

Dans le but d'optimiser les conditions de cristallisation, on prépare des "gradients" de concentration en agent précipitant ou de pH (figure A6). Cela consiste à tester une série de conditions de cristallisations en faisant varier la concentration d'un agent précipitant ou du pH.



**Figure A6. Diagramme de phase représentant l'exploration de l'espace des phases lors de l'optimisation avec un gradient de concentration.**

# Annexe 3 : Calcul des concentrations des différentes espèces lors d'une transition irréversible

## 1- Transition irréversible en une étape

Considérons la transition irréversible en une étape, c'est-à-dire sans état intermédiaire, entre un état initial noté  $C$  et un état final noté  $E$ , qui a pour constante cinétique  $k_1$  :



La concentration de l'espèce  $[C]$  peut être déduit de l'équation différentielle du premier ordre suivante :

$$-\frac{d[C]}{dt} = k_1[C] \quad (\text{A3-1})$$

Avec comme condition initiale :

$$[C(t = 0)] = C_0$$

La solution de cette équation différentielle est :

$$[C] = C_0 e^{-k_1 t}$$

La concentration en espèce finale  $E$  peut s'en déduire :

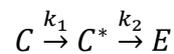
$$[E] = C_0 - [C] = C_0(1 - e^{-k_1 t})$$

En faisant l'hypothèse d'une transition en une seule étape, la fraction de capsides expansées s'exprime donc par cette formule :

$$G_{exp,t} = 1 - e^{-k_1 t} \quad (\text{A3-2})$$

## 2- Transition irréversible en deux étapes

Considérons la transition irréversible en deux étapes, avec un état intermédiaire noté  $C^*$ , ayant pour constantes cinétiques  $k_1$  et  $k_2$  :



Les concentrations des espèces C,  $C^*$  et E peuvent être déduites des équations suivantes :

$$-\frac{d[C]}{dt} = k_1[C] \quad (\text{A3-3})$$

$$\frac{d[C^*]}{dt} = k_1[C] - k_2[C^*] \quad (\text{A3-4})$$

$$[E] = C_0 - [C] - [C^*] \quad (\text{A3-5})$$

Avec comme conditions initiales :

$$[C(t = 0)] = C_0$$

$$[C^*(t = 0)] = [E(t = 0)] = 0$$

La solution de l'équation (A3-3) est (cf. section précédente) :

$$[C] = C_0 e^{-k_1 t}$$

L'équation (A3-4) peut donc être réécrite :

$$\frac{d[C^*]}{dt} + k_2[C^*] = C_0 k_1 e^{-k_1 t} \quad (\text{A3-6})$$

La solution de l'équation différentielle (A3-6) est la somme d'une solution particulière de cette équation et de la solution de l'équation homogène (sans second membre) associée. La solution de l'équation homogène associée est de la forme :

$$[C^*]_0 = \alpha e^{-k_2 t}$$

où  $\alpha$  est une constante quelconque.

Recherchons maintenant une solution particulière de l'équation (A3-6), de la forme

$$[C^*]_1 = Q(t) e^{-k_1 t}$$

où  $Q(t)$  est un polynôme du premier ordre :

$$Q(t) = vt + w$$

En remplaçant  $[C^*]$  par  $[C^*]_1$  dans l'équation A3-6, nous obtenons :

$$\frac{d(Q(t)e^{-k_1t})}{dt} + k_2Q(t)e^{-k_1t} = C_0k_1e^{-k_1t} \quad (\text{A3-7})$$

Soit :

$$Q'(t)e^{-k_1t} + (k_2 - k_1)Q(t)e^{-k_1t} = C_0k_1e^{-k_1t}$$

On en déduit :

$$Q'(t) + (k_2 - k_1)Q(t) = C_0k_1$$

Puis, en remplaçant  $Q(t)$  par son expression :

$$v + (k_2 - k_1)(vt + w) = C_0k_1$$

Il apparait donc un système de deux équations :

$$0 + (k_2 - k_1)v = 0$$

$$v + (k_2 - k_1)w = C_0k_1$$

Donc :

$$v = 0$$

$$(k_2 - k_1)w = C_0k_1$$

Enfin :

$$Q(t) = w = \frac{C_0k_1}{k_2 - k_1}$$

Les solutions de l'équation (A3-6) sont de la forme

$$[C^*] = [C^*]_0 + [C^*]_1$$

En remplaçant  $[C^*]_0$  et  $[C^*]_1$  par leur expression, on obtient :

$$[C^*] = \alpha e^{-k_2t} + \frac{C_0k_1}{k_2 - k_1} e^{-k_1t}$$

Or, à  $t=0$  la concentration en espèce  $C^*$  est nulle ; on peut donc en déduire la valeur de  $\alpha$  :

$$[C^*(t = 0)] = 0 = \alpha + \frac{C_0 k_1}{k_2 - k_1}$$

$$\alpha = -\frac{C_0 k_1}{k_2 - k_1}$$

La solution de l'équation (A3-6) est donc :

$$[C^*] = C_0 \frac{k_1}{k_2 - k_1} (e^{-k_1 t} - e^{-k_2 t})$$

La concentration en espèce  $E$  se déduit des concentrations en espèces  $C$  et  $C^*$  à l'aide de l'équation (A3-5) :

$$[E] = C_0 - C_0 e^{-k_1 t} - C_0 \frac{k_1}{k_2 - k_1} (e^{-k_1 t} - e^{-k_2 t})$$

$$[E] = C_0 \left[ 1 - e^{-k_1 t} - \frac{k_1}{k_2 - k_1} e^{-k_1 t} + \frac{k_1}{k_2 - k_1} e^{-k_2 t} \right]$$

$$[E] = C_0 \left[ 1 + \frac{k_1}{k_2 - k_1} e^{-k_2 t} - \left( 1 + \frac{k_1}{k_2 - k_1} \right) e^{-k_1 t} \right]$$

En résumé, les concentrations des espèces  $C$ ,  $C^*$  et  $E$  ont pour expression :

$$[C] = C_0 e^{-k_1 t}$$

$$[C^*] = C_0 \frac{k_1}{k_2 - k_1} (e^{-k_1 t} - e^{-k_2 t})$$

$$[E] = C_0 \left[ 1 + \frac{k_1}{k_2 - k_1} e^{-k_2 t} - \left( 1 + \frac{k_1}{k_2 - k_1} \right) e^{-k_1 t} \right]$$

# Bibliographie

- Abrescia, N. G., J. J. Cockburn, et al. (2004). "Insights into assembly from structural analysis of bacteriophage PRD1." Nature **432**(7013): 68-74.
- Abrescia, N. G., J. M. Grimes, et al. (2008). "Insights into virus evolution and membrane biogenesis from the structure of the marine lipid-containing bacteriophage PM2." Mol Cell **31**(5): 749-761.
- Abrescia, N. G. A., D. H. Bamford, et al. (2012). "Structure Unifies the Viral Universe." Annual Review of Biochemistry, Vol 81 **81**: 795-822.
- Ackermann, H. W. (2007). "5500 Phages examined in the electron microscope." Arch Virol **152**(2): 227-243.
- Ackermann, H. W. (2009). "Phage classification and characterization." Methods Mol Biol **501**: 127-140.
- Adler, J. and W. W. Tso (1974). "'Decision'-making in bacteria: chemotactic response of Escherichia coli to conflicting stimuli." Science **184**(4143): 1292-1294.
- Agirrezabala, X., J. Martin-Benito, et al. (2005). "Maturation of phage T7 involves structural modification of both shell and inner core components." EMBO J **24**(21): 3820-3829.
- Agirrezabala, X., J. A. Velazquez-Muriel, et al. (2007). "Quasi-atomic model of bacteriophage t7 procapsid shell: insights into the structure and evolution of a basic fold." Structure **15**(4): 461-472.
- Aksyuk, A. A. and M. G. Rossmann (2011). "Bacteriophage assembly." Viruses **3**(3): 172-203.
- Arnold, H. P., Q. She, et al. (1999). "The genetic element pSSVx of the extremely thermophilic crenarchaeon Sulfolobus is a hybrid between a plasmid and a virus." Mol Microbiol **34**(2): 217-226.
- Baker, M. L., W. Jiang, et al. (2005). "Common ancestry of herpesviruses and tailed DNA bacteriophages." J Virol **79**(23): 14967-14970.
- Bensasson, R. V. (1993). Excited states and free radicals in biology and medicine : contributions from flash photolysis and pulse radiolysis. Oxford ;, Oxford University Press.
- Bergmann, A., G. Fritz, et al. (2000). "Solving the generalized indirect Fourier transformation (GIFT) by Boltzmann simplex simulated annealing (BSSA)." Journal of Applied Crystallography **33**: 1212-1216.
- Bhardwaj, A., N. Walker-Kopp, et al. (2009). "An Evolutionarily Conserved Family of Virion Tail Needles Related to Bacteriophage P22 gp26: Correlation between Structural Stability and Length of the alpha-Helical Trimeric Coiled Coil." J Mol Biol **391**(1): 227-245.
- Bjornsti, M. A., B. E. Reilly, et al. (1983). "Morphogenesis of bacteriophage phi 29 of Bacillus subtilis: oriented and quantized in vitro packaging of DNA protein gp3." J Virol **45**(1): 383-396.
- Borsheim, K. Y., G. Bratbak, et al. (1990). "Enumeration and Biomass Estimation of Planktonic Bacteria and Viruses by Transmission Electron-Microscopy." Appl Environ Microbiol **56**(2): 352-356.
- Breitbart, M. (2012). "Marine viruses: truth or dare." Ann Rev Mar Sci **4**: 425-448.

- Brussaard, C. P. (2004). "Optimization of procedures for counting viruses by flow cytometry." *Appl Environ Microbiol* **70**(3): 1506-1513.
- Camacho, A., F. Jimenez, et al. (1979). "Order of assembly of the lower collar and the tail proteins of Bacillus subtilis bacteriophage phi 29." *J Virol* **29**(2): 540-545.
- Cardarelli, L., R. Lam, et al. (2010). "The crystal structure of bacteriophage HK97 gp6: defining a large family of head-tail connector proteins." *J Mol Biol* **395**(4): 754-768.
- Cardarelli, L., L. G. Pell, et al. (2010). "Phages have adapted the same protein fold to fulfill multiple functions in virion assembly." *Proc Natl Acad Sci U S A* **107**(32): 14384-14389.
- Carrascosa, J. L. (1978). "Head maturation pathway of bacteriophages T4 and T2. IV. In vitro transformation of T4 head-related particles produced by mutants in gene 17 to capsid-like structures." *J Virol* **26**(2): 420-428.
- Casjens, S. and J. King (1974). "P22 morphogenesis. I: Catalytic scaffolding protein in capsid assembly." *J Supramol Struct* **2**(2-4): 202-224.
- Casjens, S. R. (2011). "The DNA-packaging nanomotor of tailed bacteriophages." *Nat Rev Microbiol* **9**(9): 647-657.
- Caspar, D. L. and A. Klug (1962). "Physical principles in the construction of regular viruses." *Cold Spring Harb Symp Quant Biol* **27**: 1-24.
- Castillo, J., B. Goodson, et al. (2001). "T7 displayed peptides as targets for selecting peptide specific scFvs from M13 scFv display libraries." *J Immunol Methods* **257**(1-2): 117-122.
- Cerritelli, M. E., J. F. Conway, et al. (2003). "Molecular mechanisms in bacteriophage T7 procapsid assembly, maturation, and DNA containment." *Virus Structure* **64**: 301-+.
- Chen, D. H., M. L. Baker, et al. (2011). "Structural basis for scaffolding-mediated assembly and maturation of a dsDNA virus." *Proc Natl Acad Sci U S A* **108**(4): 1355-1360.
- Childs, J. D. (1980). "Effect of hoc protein on the electrophoretic mobility of intact bacteriophage T4D particles in polyacrylamide gel electrophoresis." *J Mol Biol* **141**(2): 163-173.
- Choi, K. H., M. C. Morais, et al. (2006). "Determinants of bacteriophage phi 29 head morphology." *Structure* **14**(11): 1723-1727.
- Cole, C., J. D. Barber, et al. (2008). "The Jpred 3 secondary structure prediction server." *Nucleic Acids Res* **36**(Web Server issue): W197-201.
- Conway, J. F., N. Cheng, et al. (2007). "A thermally induced phase transition in a viral capsid transforms the hexamers, leaving the pentamers unchanged." *J Struct Biol* **158**(2): 224-232.
- Conway, J. F., R. L. Duda, et al. (1995). "Proteolytic and Conformational Control of Virus Capsid Maturation - the Bacteriophage-Hk97 System." *J Mol Biol* **253**(1): 86-99.
- Curtis, T. P., W. T. Sloan, et al. (2002). "Estimating prokaryotic diversity and its limits." *Proc Natl Acad Sci U S A* **99**(16): 10494-10499.
- d'Herelle, F. (1917). "Sur un microbe invisible antagoniste des bacilles dysentériques." *Comptes rendus Acad Sci Paris* **165**: 373-375.
- Dierkes, L. E., C. L. Peebles, et al. (2009). "Mutational Analysis of a Conserved Glutamic Acid Required for Self-Catalyzed Cross-Linking of Bacteriophage HK97 Capsids." *J Virol* **83**(5): 2088-2098.
- Dodd, I. B., K. E. Shearwin, et al. (2005). "Revisited gene regulation in bacteriophage lambda." *Current Opinion in Genetics & Development* **15**(2): 145-152.
- Dokland, T. and H. Murialdo (1993). "Structural transitions during maturation of bacteriophage lambda capsids." *J Mol Biol* **233**(4): 682-694.
- Dokland, T., S. F. Wang, et al. (2002). "The structure of P4 procapsids produced by coexpression of capsid and external scaffolding proteins." *Virology* **298**(2): 224-231.

- Duckworth, D. H. (1976). "Who discovered bacteriophage?" *Bacteriol Rev* **40**(4): 793-802.
- Duda, R. L., J. Hempel, et al. (1995). "Structural transitions during bacteriophage HK97 head assembly." *J Mol Biol* **247**(4): 618-635.
- Duda, R. L., K. Martincic, et al. (1995). "Genetic-Basis of Bacteriophage Hk97 Prohead Assembly." *J Mol Biol* **247**(4): 636-647.
- Duda, R. L., K. Martincic, et al. (1995). "Bacteriophage-Hk97 Head Assembly." *Fems Microbiology Reviews* **17**(1-2): 41-46.
- Dussoix, D. and W. Arber (1962). "Host specificity of DNA produced by Escherichia coli. II. Control over acceptance of DNA from infecting phage lambda." *J Mol Biol* **5**: 37-49.
- Effantin, G., P. Boulanger, et al. (2006). "Bacteriophage T5 structure reveals similarities with HK97 and T4 suggesting evolutionary relationships." *J Mol Biol* **361**(5): 993-1002.
- Eswar, N., B. Webb, et al. (2007). "Comparative protein structure modeling using MODELLER." *Curr Protoc Protein Sci* **Chapter 2**: Unit 2 9.
- Evans, C. and C. P. Brussaard (2012). "Regional variation in lytic and lysogenic viral infection in the southern ocean and its contribution to biogeochemical cycling." *Appl Environ Microbiol* **78**(18): 6741-6748.
- Fane, B. A. and P. E. Prevelige (2003). "Mechanism of scaffolding-assisted viral assembly." *Virus Structure* **64**: 259-+.
- Fauquet CM, M. M., Maniloff J, Desselberger U, Ball LA (2005). *Virus Taxonomy, VIIIth Report of the ICTV*. London, Elsevier/Academic Press.
- Finan, K., J. P. Torella, et al. (2012). "T7 RNA Polymerase Functions In Vitro without Clustering." *Plos One* **7**(7).
- Fokine, A., P. R. Chipman, et al. (2004). "Molecular architecture of the prolate head of bacteriophage T4." *Proc Natl Acad Sci U S A* **101**(16): 6003-6008.
- Fokine, A., M. Z. Islam, et al. (2011). "Structure of the three N-terminal immunoglobulin domains of the highly immunogenic outer capsid protein from a T4-like bacteriophage." *J Virol* **85**(16): 8141-8148.
- Fokine, A., P. G. Leiman, et al. (2005). "Structural and functional similarities between the capsid proteins of bacteriophages T4 and HK97 point to a common ancestry." *Proc Natl Acad Sci U S A* **102**(20): 7163-7168.
- Fraser, J. S., K. L. Maxwell, et al. (2007). "Immunoglobulin-like domains on bacteriophage: weapons of modest damage?" *Curr Opin Microbiol* **10**(4): 382-387.
- Fraser, J. S., Z. Yu, et al. (2006). "Ig-like domains on bacteriophages: a tale of promiscuity and deceit." *J Mol Biol* **359**(2): 496-507.
- Fuller, D. N., D. M. Raymer, et al. (2007). "Measurements of single DNA molecule packaging dynamics in bacteriophage lambda reveal high forces, high motor processivity, and capsid transformations." *J Mol Biol* **373**(5): 1113-1122.
- Gan, L., J. A. Speir, et al. (2006). "Capsid conformational sampling in HK97 maturation visualized by X-ray crystallography and cryo-EM." *Structure* **14**(11): 1655-1665.
- Gertsman, I., L. Gan, et al. (2009). "An unexpected twist in viral capsid maturation." *Nature* **458**(7238): 646-U114.
- Gilcrease, E. B., D. A. Winn-Stapley, et al. (2005). "Nucleotide sequence of the head assembly gene cluster of bacteriophage L and decoration protein characterization." *J Bacteriol* **187**(6): 2050-2057.
- Glatter, O. (1977). "A new method for the evaluation of small-angle scattering data." *Journal of Applied Crystallography* **10**(5): 415-421.
- Golmohammadi, R., K. Valegard, et al. (1993). "The refined structure of bacteriophage MS2 at 2.8 Å resolution." *J Mol Biol* **234**(3): 620-639.

- Guinier, A. (1938). "The diffusion of x-rays under the extremely weak angles applied to the study of fine particles and colloidal suspension." Comptes Rendus Hebdomadaires Des Seances De L Academie Des Sciences **206**: 1374-1376.
- Guinier, A. and G. Fournet (1955). Small-angle scattering of X-rays, Wiley.
- Guixa-Boixereu, N., D. Vaque, et al. (2002). "Viral distribution and activity in Antarctic waters." Deep-Sea Research Part II-Topical Studies in Oceanography **49**(4-5): 827-845.
- Hankin, E. H. (1896). "L'action bactéricide des eaux de la Jumna et du Gange sur le vibron du choléra." Annales de l'Institut Pasteur **10**: 511-523.
- Hanlon, G. W. (2007). "Bacteriophages: An appraisal of their role in the treatment of bacterial infections." International Journal of Antimicrobial Agents **30**(2): 118-128.
- Hendrix, R. W. and R. L. Duda (1998). "Bacteriophage HK97 head assembly: A protein ballet." Advances in Virus Research, Vol 50 **50**: 235-288.
- Hendrix, R. W. and J. E. Johnson (2012). "Bacteriophage HK97 Capsid Assembly and Maturation." Viral Molecular Machines **726**: 351-363.
- Hershey, A. D. and M. Chase (1952). "Independent functions of viral protein and nucleic acid in growth of bacteriophage." J Gen Physiol **36**(1): 39-56.
- Huet, A., J. F. Conway, et al. (2010). "In vitro assembly of the T=13 procapsid of bacteriophage T5 with its scaffolding domain." J Virol **84**(18): 9350-9358.
- Huet, A., J. F. Conway, et al. (2010). "In Vitro Assembly of the T=13 Procapsid of Bacteriophage T5 with Its Scaffolding Domain." J Virol **84**(18): 9350-9358.
- Ionel, A., J. A. Velazquez-Muriel, et al. (2011). "Molecular rearrangements involved in the capsid shell maturation of bacteriophage T7." J Biol Chem **286**(1): 234-242.
- Ishii, T. and M. Yanagida (1977). "The two dispensable structural proteins (soc and hoc) of the T4 phage capsid; their purification and properties, isolation and characterization of the defective mutants, and their binding with the defective heads in vitro." J Mol Biol **109**(4): 487-514.
- Isidro, A., A. O. Henriques, et al. (2004). "The portal protein plays essential roles at different steps of the SPP1 DNA packaging process." Virology **322**(2): 253-263.
- Jardine, P. J., M. C. McCormick, et al. (1998). "The bacteriophage T4 DNA packaging apparatus targets the unexpanded prohead." J Mol Biol **284**(3): 647-659.
- Jiang, W., Z. L. Li, et al. (2003). "Coat protein fold and maturation transition of bacteriophage P22 seen at subnanometer resolutions." Nat Struct Biol **10**(2): 131-135.
- Johnson, J. E. and W. Chiu (2007). "DNA packaging and delivery machines in tailed bacteriophages." Curr Opin Struct Biol **17**(2): 237-243.
- Katsura, I. (1981). "Structure and function of the major tail protein of bacteriophage lambda. Mutants having small major tail protein molecules in their virion." J Mol Biol **146**(4): 493-512.
- Kellenberger, E. (1990). "Form Determination of the Heads of Bacteriophages." European Journal of Biochemistry **190**(2): 233-248.
- Kelley, L. A. and M. J. E. Sternberg (2009). "Protein structure prediction on the Web: a case study using the Phyre server." Nature Protocols **4**(3): 363-371.
- King, J. and S. Casjens (1974). "Catalytic head assembling protein in virus morphogenesis." Nature **251**(5471): 112-119.
- Konarev, P. V., V. V. Volkov, et al. (2003). "PRIMUS: a Windows PC-based system for small-angle scattering data analysis." Journal of Applied Crystallography **36**: 1277-1282.
- Kozin, M. B. and D. I. Svergun (2001). "Automated matching of high- and low-resolution structural models." Journal of Applied Crystallography **34**: 33-41.

- Krupovic, M., P. Forterre, et al. (2010). "Comparative Analysis of the Mosaic Genomes of Tailed Archaeal Viruses and Proviruses Suggests Common Themes for Virion Architecture and Assembly with Tailed Viruses of Bacteria." J Mol Biol **397**(1): 144-160.
- Kubista, M., J. M. Andrade, et al. (2006). "The real-time polymerase chain reaction." Mol Aspects Med **27**(2-3): 95-125.
- Kutateladze, M. and R. Adamia (2010). "Bacteriophages as potential new therapeutics to replace or supplement antibiotics." Trends in Biotechnology **28**(12): 591-595.
- Kuwamoto, S., S. Akiyama, et al. (2004). "Radiation damage to a protein solution, detected by synchrotron X-ray small-angle scattering: dose-related considerations and suppression by cryoprotectants." Journal of Synchrotron Radiation **11**: 462-468.
- Lander, G. C., A. C. Baudoux, et al. (2012). "Capsomer dynamics and stabilization in the T = 12 marine bacteriophage SIO-2 and its procapsid studied by CryoEM." Structure **20**(3): 498-503.
- Lander, G. C., A. Evilevitch, et al. (2008). "Bacteriophage lambda stabilization by auxiliary protein gpD: timing, location, and mechanism of attachment determined by cryo-EM." Structure **16**(9): 1399-1406.
- Lander, G. C., R. Khayat, et al. (2009). "The P22 Tail Machine at Subnanometer Resolution Reveals the Architecture of an Infection Conduit." Structure **17**(6): 789-799.
- Lanni, Y. T. (1958). "Lysis inhibition with a mutant of bacteriophage T5." Virology **5**(3): 481-501.
- Lech, K., R. Brent, et al. (2001). "Lambda as a cloning vector." Curr Protoc Mol Biol **Chapter 1**: Unit1 10.
- Lee, K. K., L. Gan, et al. (2008). "Virus capsid expansion driven by the capture of mobile surface loops." Structure **16**(10): 1491-1502.
- Lee, K. K., H. Tsuruta, et al. (2005). "Cooperative reorganization of a 420 subunit virus capsid." J Mol Biol **352**(3): 723-735.
- Letellier, L., P. Boulanger, et al. (2004). "Main features on tailed phage, host recognition and DNA uptake." Front Biosci **9**: 1228-1339.
- Leulliot, N., L. Tresaugues, et al. (2005). "High-throughput crystal-optimization strategies in the South Paris Yeast Structural Genomics Project: one size fits all?" Acta Crystallogr D Biol Crystallogr **61**(Pt 6): 664-670.
- Lhuillier, S., M. Gallopin, et al. (2009). "Structure of bacteriophage SPP1 head-to-tail connection reveals mechanism for viral DNA gating." Proc Natl Acad Sci U S A **106**(21): 8507-8512.
- Livingston, K. (1992). "Phage and the Origins of Molecular-Biology - Cairns,J, Stent,Gs, Watson,Jd." Science **257**(5077): 1775-1775.
- Manoutcharian, K., G. Gevorkian, et al. (2001). "Phage displayed biomolecules as preventive and therapeutic agents." Curr Pharm Biotechnol **2**(3): 217-223.
- Marie, D., C. P. D. Brussaard, et al. (1999). "Enumeration of marine viruses in culture and natural samples by flow cytometry." Appl Environ Microbiol **65**(1): 45-52.
- Maruyama, I. N., H. I. Maruyama, et al. (1994). "Lambda foo: a lambda phage vector for the expression of foreign proteins." Proc Natl Acad Sci U S A **91**(17): 8273-8277.
- Marvin, D. A. (1990). "Model-building studies of Inovirus: genetic variations on a geometric theme." Int J Biol Macromol **12**(2): 125-138.
- Maxwell, K. L., A. A. Yee, et al. (2002). "The solution structure of the bacteriophage lambda head-tail joining protein, gpFII." J Mol Biol **318**(5): 1395-1404.
- Maxwell, K. L., A. A. Yee, et al. (2001). "The solution structure of bacteriophage lambda protein W, a small morphogenetic protein possessing a novel fold." J Mol Biol **308**(1): 9-14.

- McKenna, R., D. Xia, et al. (1992). "Atomic structure of single-stranded DNA bacteriophage phi X174 and its functional implications." Nature **355**(6356): 137-143.
- Michaud, G., A. Zachary, et al. (1989). "Membrane-Associated Assembly of a Phage-T4 DNA Entrance Vertex Structure Studied with Expression Vectors." J Mol Biol **209**(4): 667-681.
- Moore, P. (1980). "Small-angle scattering. Information content and error analysis." Journal of Applied Crystallography **13**(2): 168-175.
- Morais, M. C., K. H. Choi, et al. (2005). "Conservation of the capsid structure in tailed dsDNA bacteriophages: the pseudoatomic structure of phi29." Mol Cell **18**(2): 149-159.
- Morais, M. C., S. Kanamaru, et al. (2003). "Bacteriophage phi29 scaffolding protein gp7 before and after prohead assembly." Nat Struct Biol **10**(7): 572-576.
- Murialdo, H. (1991). "Bacteriophage lambda DNA maturation and packaging." Annual Review of Biochemistry, Vol 81 **60**: 125-153.
- Murphy, K. C. (2012). "Phage recombinases and their applications." Advances in Virus Research, Vol 50 **83**: 367-414.
- Nemecek, D., N. Cheng, et al. (2011). "Stepwise expansion of the bacteriophage varphi6 procapsid: possible packaging intermediates." J Mol Biol **414**(2): 260-271.
- Olia, A. S., A. Bhardwaj, et al. (2007). "Role of gene 10 protein in the hierarchical assembly of the bacteriophage P22 portal vertex structure." Biochemistry **46**(30): 8776-8784.
- Olia, A. S., S. Casjens, et al. (2007). "Structure of phage P22 cell envelope-penetrating needle." Nat Struct Mol Biol **14**(12): 1221-1226.
- Olia, A. S., P. E. Prevelige, Jr., et al. (2011). "Three-dimensional structure of a viral genome-delivery portal vertex." Nat Struct Mol Biol **18**(5): 597-603.
- Oppenheim, A. B., O. Kobiler, et al. (2005). "Switches in bacteriophage lambda development." Annual Review of Genetics **39**: 409-429.
- Orlova, E. V., B. Gowen, et al. (2003). "Structure of a viral DNA gatekeeper at 10 angstrom resolution by cryo-electron microscopy." Embo Journal **22**(6): 1255-1262.
- Ortmann, A. C. and C. A. Suttle (2005). "High abundances of viruses in a deep-sea hydrothermal vent system indicates viral mediated microbial mortality." Deep-Sea Research Part I-Oceanographic Research Papers **52**(8): 1515-1527.
- Parker, M. H. and P. E. Prevelige (1998). "Electrostatic interactions drive scaffolding/coat protein binding and procapsid maturation in bacteriophage P22." Virology **250**(2): 337-349.
- Pascal, J. M. (2008). "DNA and RNA ligases: structural variations and shared mechanisms." Curr Opin Struct Biol **18**(1): 96-105.
- Petoukhov, M. V., D. Franke, et al. (2012). "New developments in the ATSAS program package for small-angle scattering data analysis." Journal of Applied Crystallography **45**: 342-350.
- Proctor, L. M. and J. A. Fuhrman (1990). "Viral Mortality of Marine-Bacteria and Cyanobacteria." Nature **343**(6253): 60-62.
- Qin, L., A. Fokine, et al. (2010). "Structure of the small outer capsid protein, Soc: a clamp for stabilizing capsids of T4-like phages." J Mol Biol **395**(4): 728-741.
- Quevillon-Cheruel, S., D. Liger, et al. (2004). "The Paris-Sud yeast structural genomics pilot-project: from structure to function." Biochimie **86**(9-10): 617-623.
- Rao, V. B. and L. W. Black (1985). "DNA packaging of bacteriophage T4 proheads in vitro. Evidence that prohead expansion is not coupled to DNA packaging." J Mol Biol **185**(3): 565-578.
- Rao, V. B. and M. Feiss (2008). "The bacteriophage DNA packaging motor." Annual Review of Genetics **42**: 647-681.

- Ray, K., M. Oram, et al. (2009). "Portal control of viral prohead expansion and DNA packaging." *Virology* **391**(1): 44-50.
- Rohwer, F. (2003). "Global phage diversity." *Cell* **113**(2): 141.
- Ross, P. D., N. Cheng, et al. (2005). "Crosslinking renders bacteriophage HK97 capsid maturation irreversible and effects an essential stabilization." *EMBO J* **24**(7): 1352-1363.
- Saigo, K. (1978). "Isolation of high-density mutants and identification of nonessential structural proteins in bacteriophage T5; dispensability of L-shaped tail fibers and a secondary major head protein." *Virology* **85**(2): 422-433.
- Salas, M., C. Vasquez, et al. (1972). "Head fibers of bacteriophage phi 29." *Virology* **50**(1): 180-188.
- Sanger, F., G. M. Air, et al. (1977). "Nucleotide sequence of bacteriophage phi X174 DNA." *Nature* **265**(5596): 687-695.
- Sathaliyawala, T., M. Z. Islam, et al. (2010). "Functional analysis of the highly antigenic outer capsid protein, Hoc, a virus decoration protein from T4-like bacteriophages." *Mol Microbiol* **77**(2): 444-455.
- Serwer, P. (2003). "Models of bacteriophage DNA packaging motors." *J Struct Biol* **141**(3): 179-188.
- Shivachandra, S. B., M. Rao, et al. (2006). "In vitro binding of anthrax protective antigen on bacteriophage T4 capsid surface through Hoc-capsid interactions: a strategy for efficient display of large full-length proteins." *Virology* **345**(1): 190-198.
- Smith, D. E. (2011). "Single-molecule studies of viral DNA packaging." *Curr Opin Virol* **1**(2): 134-141.
- Smith, G. P. (1985). "Filamentous fusion phage: novel expression vectors that display cloned antigens on the virion surface." *Science* **228**(4705): 1315-1317.
- Smith, G. P. and V. A. Petrenko (1997). "Phage Display." *Chem Rev* **97**(2): 391-410.
- Soding, J., A. Biegert, et al. (2005). "The HHpred interactive server for protein homology detection and structure prediction." *Nucleic Acids Res* **33**(Web Server issue): W244-248.
- Sternberg, N. and R. H. Hoess (1995). "Display of peptides and proteins on the surface of bacteriophage lambda." *Proc Natl Acad Sci U S A* **92**(5): 1609-1613.
- Steven, A. C., J. B. Heymann, et al. (2005). "Virus maturation: dynamics and mechanism of a stabilizing structural transition that leads to infectivity." *Curr Opin Struct Biol* **15**(2): 227-236.
- Su, X. D., L. N. Gastinel, et al. (1998). "Crystal structure of hemolin: a horseshoe shape with implications for homophilic adhesion." *Science* **281**(5379): 991-995.
- Sun, S., K. Kondabagil, et al. (2008). "The Structure of the Phage T4 DNA Packaging Motor Suggests a Mechanism Dependent on Electrostatic Forces." *Cell* **135**(7): 1251-1262.
- Suttle, C. A. (2005). "Viruses in the sea." *Nature* **437**(7057): 356-361.
- Suttle, C. A. (2007). "Marine viruses--major players in the global ecosystem." *Nat Rev Microbiol* **5**(10): 801-812.
- Suttle, C. A., and J. A. Fuhrman (2010). Enumeration of virus particles in aquatic or sediment samples by epifluorescence microscopy. *Manual of Aquatic Viral Ecology*. M. G. W. S. W. Wilhelm, and C. A. Suttle, Association for the Sciences of Limnology and Oceanography: 145-153.
- Svergun, D. I. (1992). "Determination of the Regularization Parameter in Indirect-Transform Methods Using Perceptual Criteria." *Journal of Applied Crystallography* **25**: 495-503.
- Svergun, D. I. and M. H. J. Koch (2003). "Small-angle scattering studies of biological macromolecules in solution." *Reports on Progress in Physics* **66**(10): 1735-1782.

- Svergun, D. I., M. V. Petoukhov, et al. (2001). "Determination of domain structure of proteins from X-ray solution scattering." *Biophys J* **80**(6): 2946-2953.
- Svergun, D. I., A. V. Semenyuk, et al. (1988). "Small-Angle-Scattering-Data Treatment by the Regularization Method." *Acta Crystallographica Section A* **44**: 244-250.
- Tang, L., E. B. Gilcrease, et al. (2006). "Highly discriminatory binding of capsid-cementing proteins in bacteriophage L." *Structure* **14**(5): 837-845.
- Tavares, P., S. Zinn-Justin, et al. (2012). "Genome gating in tailed bacteriophage capsids." *Viral Molecular Machines* **726**: 585-600.
- Twort, F. W. (1915). "An investigation on the nature of ultra-microscopic viruses." *Lancet* **2**: 1241-1243.
- Valpuesta, J. M. and J. L. Carrascosa (1994). "Structure of Viral Connectors and Their Function in Bacteriophage Assembly and DNA Packaging." *Q Rev Biophys* **27**(2): 107-155.
- Veesler, D., J. Quispe, et al. (2012). "Maturation in Action: CryoEM Study of a Viral Capsid Caught during Expansion." *Structure*.
- Weinbauer, M. G. and F. Rassoulzadegan (2004). "Are viruses driving microbial diversification and diversity?" *Environ Microbiol* **6**(1): 1-11.
- Wen, K., A. C. Ortmann, et al. (2004). "Accurate estimation of viral abundance by epifluorescence microscopy." *Appl Environ Microbiol* **70**(7): 3862-3867.
- White, H. E., M. B. Sherman, et al. (2012). "Capsid structure and its stability at the late stages of bacteriophage SPP1 assembly." *J Virol* **86**(12): 6768-6777.
- White, H. E., M. B. Sherman, et al. (2012). "Capsid Structure and Its Stability at the Late Stages of Bacteriophage SPP1 Assembly." *J Virol* **86**(12): 6768-6777.
- Wikoff, W. R., L. Liljas, et al. (2000). "Topologically linked protein rings in the bacteriophage HK97 capsid." *Science* **289**(5487): 2129-2133.
- Winick, H. and S. Doniach (1980). *Synchrotron radiation research*. New York, Plenum Press.
- Xiang, Y. and M. G. Rossmann (2011). "Structure of bacteriophage phi29 head fibers has a supercoiled triple repeating helix-turn-helix motif." *Proc Natl Acad Sci U S A* **108**(12): 4806-4810.
- Yamaguchi, Y. and M. Yanagida (1980). "Head shell protein hoc alters the surface charge of bacteriophage T4. Composite slab gel electrophoresis of phage T4 and related particles." *J Mol Biol* **141**(2): 175-193.
- Yang, Q., N. K. Maluf, et al. (2008). "Packaging of a unit-length viral genome: the role of nucleotides and the gpD decoration protein in stable nucleocapsid assembly in bacteriophage lambda." *J Mol Biol* **383**(5): 1037-1048.
- Zhao, H. Y., C. J. Finch, et al. (2010). "Crystal structure of the DNA-recognition component of the bacterial virus Sf6 genome-packaging machine." *Proc Natl Acad Sci U S A* **107**(5): 1971-1976.
- Zillig, W., H. P. Arnold, et al. (1998). "Genetic elements in the extremely thermophilic archaeon *Sulfolobus*." *Extremophiles* **2**(3): 131-140.
- Zillig, W., F. Gropp, et al. (1986). "Archaeobacterial Virus Host Systems." *Systematic and Applied Microbiology* **7**(1): 58-66.
- Zlotnick, A., M. M. Suhanovsky, et al. (2012). "The energetic contributions of scaffolding and coat proteins to the assembly of bacteriophage procapsids." *Virology* **428**(1): 64-69.
- Zweig, M. and D. J. Cummings (1973). "Structural proteins of bacteriophage T5." *Virology* **51**(2): 443-453.



Le bactériophage T5 est un virus infectant *E. Coli*. L'assemblage et la maturation de sa capsid comportent plusieurs étapes critiques pour la formation des virions lors du cycle infectieux.

Parmi ces étapes, j'ai étudié les processus d'expansion et de décoration de la capsid de T5. L'expansion implique d'importantes réorganisations conformationnelles des 775 sous-unités de la protéine pb8 composant la capsid, conduisant au doublement du volume de la capsid qui peut alors contenir le génome du phage. J'ai déterminé des conditions physico-chimiques permettant d'induire l'expansion de la capsid *in vitro*, puis j'ai effectué des expériences de SAXS résolues dans le temps montrant que l'expansion est un processus hautement coopératif, qui conduit, en une étape, à un état final remarquablement stable. D'autre part, j'ai réalisé une étude fonctionnelle de la protéine de décoration pb10, montrant que sa fixation est un marqueur de l'expansion. Enfin l'étude structurale de pb10 menée par SAXS a permis de déterminer un modèle à basse résolution de son enveloppe moléculaire.

**Mots clés :** Bactériophage, capsid virale, expansion, protéine de décoration, SAXS.

Bacteriophage T5 is a virus infecting *E. Coli*. Its capsid assembly and maturation include several critical steps leading to the formation of new virions during the infectious cycle.

During my thesis I focused on the expansion and decoration of T5 capsid. Expansion consists in large conformational reorganizations of the 775 capsid protein subunits, yielding a two-fold increase of the capsid volume and allowing it to accommodate the full-length genome. I have ascertained physicochemical conditions that trigger *in vitro* expansion of the capsid, and using a time-resolved SAXS study, I showed that expansion is a highly cooperative two-state process leading to a remarkably stable final state. I have also carried out a functional study of the decoration protein pb10, showing that it only binds to expanded proheads. Finally, a low resolution model of pb10 was determined by SAXS.

**Keywords :** Bacteriophage, virus capsid, expansion, decoration protein, SAXS.

### **Laboratoire de rattachement**

Institut de biochimie et biophysique moléculaire et cellulaire (IBBMC)

Equipe Virologie Bactérienne

UMR CNRS 8619 bâtiment 430 Université Paris XI 91405 Orsay Cedex

ECOLE DOCTORALE : INNOVATION THERAPEUTIQUE

PÔLE : INGENIERIE DES PROTEINES ET CIBLES THERAPEUTIQUES

UNIVERSITÉ PARIS-SUD 11

UFR «FACULTÉ DE PHARMACIE DE CHATENAY-MALABRY »

5, rue Jean Baptiste Clément 92296 CHÂTENAY-MALABRY Cedex